
UNIVERSIDADE ESTADUAL DE MARINGÁ
CENTRO DE CIÊNCIAS EXATAS
PROGRAMA DE PÓS-GRADUAÇÃO EM FÍSICA

LUIZ GUSTAVO DE ANDRADE ALVES

INVARIÂNCIA DE ESCALA E DIFUSÃO ANÔMALA
EM SISTEMAS COMPLEXOS URBANOS E BIOLÓGICOS

Maringá
Março de 2017

UNIVERSIDADE ESTADUAL DE MARINGÁ
CENTRO DE CIÊNCIAS EXATAS
PROGRAMA DE PÓS-GRADUAÇÃO EM FÍSICA

LUIZ GUSTAVO DE ANDRADE ALVES

INVARIÂNCIA DE ESCALA E DIFUSÃO ANÔMALA
EM SISTEMAS COMPLEXOS URBANOS E BIOLÓGICOS

Tese apresentada à Universidade Estadual de
Maringá, como requisito parcial para obtenção do
título de doutor.

Orientador: Prof. Dr. Haroldo Valentin Ribeiro

Maringá

Março de 2017

Dados Internacionais de Catalogação na Publicação (CIP)
(Biblioteca Central - UEM, Maringá, PR, Brasil)

A474i Alves, Luiz Gustavo de Andrade
Invariância de escala e difusão anômala em sistemas complexos urbanos e biológicos / Luiz Gustavo de Andrade Alves -- Maringá, 2017.
131 f. : il., color., figs., tabs.

Orientador: Prof. Dr. Haroldo Valentin Ribeiro.
Tese (doutorado) - Universidade Estadual de Maringá, Centro de Ciências Exatas, Programa de Pós-Graduação em Física, 2017.

1. Sistemas complexos. 2. Física estatística. 3. Difusão anômala. 4. Sistemas urbanos. 5. Sistemas biológicos. I. Ribeiro, Haroldo Valentin, orient. II. Universidade Estadual de Maringá. Centro de Ciências da Exatas. Programa de Pós-Graduação em Física. III. Título.

CDD 21.ed. 530.13

AHS

*Dedico esta tese de doutorado aos meus pais,
Luiz Carlos Alves e
Mirian Cristina de Andrade Alves.*

Agradecimentos

O caminho para uma defesa de tese de doutorado é longo e certamente muitas pessoas contribuíram para que isso acontecesse. Seja a contribuição ao ensinar uma disciplina, ajudar em um exercício, tirar uma dúvida sobre um algoritmo ou apoiar no outro hemisfério da vida, o lado não acadêmico, dando suporte emocional ou financeiro, muitas pessoas contribuíram para que minha jornada no doutorado fosse mais fácil. Obrigado a todos que me ajudaram completar essa etapa da minha vida.

Em especial, gostaria de agradecer ao meu orientador Prof. Dr. Haroldo V. Ribeiro pela confiança em meu trabalho e pelo apoio em meus projetos e parcerias. Também gostaria de agradecer ao Prof. Dr. Luís A. N. Amaral por ter me recebido em meu estágio sanduíche na *Northwestern University* e por me ensinar lições valiosas sobre a ciência e o trabalho científico. Sem dúvida, Haroldo e Luís foram essenciais na minha formação como físico e tiveram uma contribuição direta nos resultados de minha pesquisa.

Um agradecimento especial à Débora B. Scariot que me deu suporte emocional nos melhores e piores momentos dessa jornada. Ela que agiu com muita paciência e carinho para me confortar todos os dias, independente do humor e independente da situação. A caminhada da vida (e do doutorado) ficou muito mais fácil e divertida com a Débora.

Agradeço aos meus pais, Luiz C. Alves e Mirian C. A. Alves, pelo carinho, suporte emocional e financeiro (é difícil a vida do bolsista) e por tudo que me ensinaram desde que nasci, fazendo-me

acreditar nos meus sonhos e lutar por meus objetivos de vida. Meus agradecimentos, também, ao meu irmão Cássio R. A. Alves e à Claudia D. B. Bettio que também contribuíram dando suporte nas minhas decisões durante esses anos de doutorado.

Gostaria também de agradecer a todas as pessoas envolvidas nos trabalhos desenvolvidos durante meu doutorado, em especial os professores Dr. Renio S. Mendes, Dr. Ervin K. Lenzi, Dr. Richard I. Morimoto e Dr. Celso V. Nakamura, que contribuíram com o desenvolvimento dos trabalhos descritos nessa tese. Também gostaria de agradecer aos colaboradores dos artigos da tese, Prof. Dr. Rafael Zola, Prof. Dr. Angel A. Tateishi, Dr. Peter B. Winter, Leonardo N. Ferreira, Renato R. Guimarães, Renée M. Brielmann e Débora B. Scariot.

Agradeço a todos os membros do *Amaral-Lab (Northwestern University)* por esse um ano de convivência e paciência com meu sotaque, pelos cafés e cervejas que tomamos juntos e pela troca de experiência de vida e profissional. Meus agradecimentos ao professor doutor Adam R. Pah, aos doutores Diego F. M. Oliveira, Martin Gerlach, Thomas Stoeger, Peter B. Winter, Nicolás Peláez, Julia Poncela-Casasnovas e Yang Yang, aos colegas de laboratório Adam Hockenberry, Hyojun Ada Lee, João Moreira, Meagan Bechel, Sebastian Bernasek, Sophia Liu, Hanyu Shi, Nicholas Timkovich, Leonardo Ferreira e Helio Tejedor. Agradeço também à Dra. Laura Bott e à Dra. Ambre Sala, membras do *Morimoto-Lab (Northwestern University)*, que contribuíram com os experimentos em *C. elegans*. Agradeço também aos membros do *Complex-Lab (Universidade Estadual de Maringá)* pela convivência e discussões frutíferas. Em especial, ao professor doutor Sergio Picoli, e aos colegas de laboratório Denner S. Vieira e professor doutor Fernando J. Antonio. Por fim, agradeço aos meus amigos, Maike A. F. Santos e Renato F. Souza que outrora fizeram parte da Pós-Graduação em Física da Universidade Estadual de Maringá e que desde então sempre me acompanharam em discussões, muitas das vezes de cunho científico/profissional, outras vezes nem tanto, mas sem nunca perder o humor.

Agradeço às agências CNPq e CAPES por suportarem essa tese com o auxílio financeiro.

Abstract

Statistical physics has proved to be fruitful in the study of systems far from traditional physics. Physicists have applied techniques and concepts from statistical mechanics in the study of data from complex systems of the most diverse areas of knowledge. Several studies focus on time series of financial markets, social and biological systems. Many of these approaches use concepts derived from phase transition physics and anomalous diffusion. In this thesis, we apply the concepts of scale invariance and anomalous diffusion in the study of urban and biological complex systems. In Chapter 1, we investigate a metric that takes into account the nonlinearities in the relationship between urban indicators and population size, where we show that this scale-adjusted metric can be used to quantify and predict urban indicators. In Chapter 2, we show that the spatial correlation in the number of per capita homicides decays exponentially and that this correlation is independent of the population dynamics. We also show that this correlation leads to clusters of cities that can be modelled as a percolation-like transition. In Chapter 3, we make a complete characterization of the diffusion patterns of four species of protozoa. We show that the spread of these protozoan is superdiffusive and that there are long-range correlations in the radial velocities. In Chapter 4, we use a similar approach showing that *C. Elegans* is also characterized by superdiffusion and long-range correlations in the velocities. We further show that the exponents characterizing their dynamics change with ageing and diseases, similar to what was previously found in human physiology. Finally, in Chapter 5, we propose an extension for the comb-model via Langevin-like equations driven by fractional Gaussian noises (long-range correlated). We show that the correlations can affect the diffusive behavior in a non-trivial fashion, resulting in a quite rich diffusive scenario, that can be applied in the context of complex systems such as living cells.

Key words: Complex Systems. Statistical Physics. Anomalous Diffusion. Urban System. Systems Biology.

Resumo

A física estatística tem se mostrado frutífera no estudo de sistemas não comuns da física tradicional. Físicos têm aplicado técnicas e conceitos de mecânica estatística no estudo de dados de sistemas complexos das mais diversas áreas do conhecimento. Vários estudos focam em séries temporais de mercados financeiros, sistemas sociais e biológicos. Essas abordagens, frequentemente, usam conceitos provenientes da física de transição de fase e difusão anômala. Nessa tese, aplicamos os conceitos de invariância de escala e difusão anômala no estudo de sistemas complexos urbanos e biológicos. No Capítulo 1, investigamos uma métrica que leva em consideração as não linearidades na relação entre indicadores urbanos e o tamanho populacional para mostrar que essa métrica ajustada à escala pode ser usada para quantificar e prever indicadores urbanos. No Capítulo 2, mostramos que a correlação espacial no número de homicídios *per capita* decai exponencialmente e que essa correlação é independente da dinâmica populacional. Mostramos também que essas correlações levam a aglomerados de cidades que podem ser modeladas no contexto de transição de fase e percolação. No Capítulo 3, apresentamos uma caracterização completa dos padrões de difusão de quatro espécies de protozoários. Mostramos que as trajetórias desses protozoários têm uma dinâmica superdifusiva e que há correlações de longo alcance nas velocidades radiais. No Capítulo 4, usamos uma abordagem similar para mostrar que a dinâmica dos *C. Elegans* é superdifusiva e que suas velocidades apresentam correlações de longo alcance. Mostramos também que os expoentes que caracterizam essa dinâmica mudam com envelhecimento e doenças, semelhante ao que foi encontrado anteriormente na fisiologia humana. Finalmente, no Capítulo 5, propomos uma extensão para o modelo de pente via equações de Langevin governadas por ruídos gaussianos fracionários (com correlação de longo alcance). Mostramos que as correlações podem afetar o comportamento difusivo de forma não trivial, resultando em um cenário difusivo bastante rico.

Palavras-chave: Sistemas Complexos. Física Estatística. Difusão Anômala. Sistemas Urbanos. Sistemas Biológicos.

Sumário

Introdução	1
1 Métricas ajustadas à escala para prever a evolução de indicadores urbanos e quantificar o desempenho de cidades	7
1.1 Introdução e apresentação do problema	8
1.2 Apresentação dos dados	10
1.3 Resultados e discussão	12
1.4 Conclusões e perspectivas	35
2 Correlações espaciais, aglomeração e transições tipo percolação em crimes de homicídio	36
2.1 Introdução e apresentação do problema	37
2.2 Apresentação dos dados	38
2.3 Resultados e discussão	39
2.4 Conclusões e perspectivas	50
3 Transientes de superdifusão e correlações de longo alcance nos padrões de motilidade dos protozoários tripanosomatídeos flagelados	51
3.1 Introdução e apresentação do problema	52

3.2	Métodos	54
3.2.1	Manutenção e cultura dos parasitas	54
3.2.2	Configuração experimental, aquisição de imagem e rastreamento	54
3.3	Resultados e discussão	58
3.4	Conclusões e perspectivas	74
4	Correlações de longo alcance e dinâmica fractal em <i>C. elegans</i>: mudanças com envelhecimento e estresse	75
4.1	Introdução e apresentação do problema	76
4.2	Métodos e apresentação dos dados	77
4.2.1	<i>C. elegans</i> : cultura e experimentos de motilidade	77
4.3	Resultados e discussão	80
4.4	Conclusões e perspectivas	92
5	Investigando a interação entre mecanismos de difusão anômala via caminhada browniana fracionária em uma estrutura de pente	93
5.1	Introdução e apresentação do problema	94
5.2	Generalização do modelo de pente e correlações de longo alcance	96
5.3	Conclusões e perspectivas	106
	Visão geral	107
	Referências Bibliográficas	110

Introdução

A abordagem de problemas interdisciplinares por meio de métodos de física teórica tem se tornado cada vez mais comum nas últimas décadas [1–6]. Conceitos de física estatística, teoria da informação e dinâmica não linear têm sido usados na modelagem de problemas de biologia [7], economia [8], sociologia [9], fisiologia [10], música [11], literatura [12] entre outras disciplinas. Esta tese enquadrada-se nesse contexto multidisciplinar, com o emprego de ferramentas de física estatística para o estudo de sistemas complexos urbanos e biológicos [13–17].

Sistemas Complexos [5] é um campo da ciência que estuda como as partes de um sistema dão origem a comportamentos coletivos. Exemplos de sistemas complexos incluem sistemas sociais formados por pessoas, o cérebro formado a partir de neurônios e moléculas formadas por átomos. O campo de sistemas complexos abrange todas as disciplinas tradicionais da ciência (*e.g.*, física, química, biologia), assim como engenharia, gestão e administração.

Os temas que abordamos nesta tese são sistemas complexos do ponto de vista que existe um comportamento emergente devido à interação das partes que compõem o sistema. Por exemplo, no contexto de sistemas urbanos, o número de pessoas da cidade parece interferir no número de conexões entre elas, fazendo com que indicadores urbanos *versus* o tamanho da população não seja uma relação linear [14]. No contexto de sistemas biológicos, o movimento dos organismos geram padrões complexos que emergem da superposição de fatores externos (por exemplo, temperatura) e internos (por exemplo, mecanismos reguladores do movimento) [17].

Nesse sentido, o objetivo desta tese é usar métodos tradicionais da física estatística e termodinâmica para o estudo de sistemas complexos urbanos e biológicos, dando enfoque a conceitos e técnicas de invariância de escala e autossimilaridade, análise de flutuação, fractalidade, transições de fases e difusão anômala. Esses conceitos, de certa forma, são o ponto em comum nos problemas aqui abordados, visto que os temas são de disciplinas distintas e buscam respostas bem diferentes. Por exemplo, no caso de cidades, existe uma necessidade de entender melhor os padrões urbanos para criação de novas políticas públicas, prevenção de crimes e poluição, assim como o gerenciamento e alocação de recursos; para o caso de sistemas biológicos, procura-se entender melhor organismos causadores de doenças ou organismos modelos para buscar meios de prevenção de doenças.

Métodos e técnicas inspiradas em física estatística são úteis na busca por características ocultas em sistemas complexos [1–6]. As origens da física estatística estão enraizadas nos trabalhos seminais sobre teoria cinética dos gases do ano de 1738 desenvolvidos pelo físico e matemático suíço Danielis Bernoulli em seu livro *Hydrodynamica* [18]. Bernoulli postulou o argumento de que os gases consistem de um grande número de moléculas movendo-se em todas as direções. Dessa forma, o impacto dessas moléculas sobre uma superfície provoca a pressão do gás que sentimos, assim como o calor seria simplesmente a energia cinética dessas partículas em movimento. Mais tarde, em 1860, esses conceitos possibilitaram o físico escocês James Clerk Maxwell formular a distribuição de Maxwell das velocidades das partículas [19,20]. Alguns anos mais tarde, o então estudante de física da Universidade de Viena, Ludwig Eduard Boltzmann, aprofundou-se nos estudos de Maxwell e, assim, deu uma interpretação estatística para as leis da termodinâmica [21]. Juntamente a Boltzmann, o físico matemático americano Josiah Willard Gibbs formalizou a mecânica estatística [22].

Os avanços da física estatística permitiram explicar e descrever quantitativamente a supercondutividade, superfluidez, turbulência, fenômenos coletivos em sólidos e em plasmas e características estruturais de líquidos [22–27]. A física estatística ajudou no desenvolvimento de cristais líquidos [28] e na construção da teoria de transição de fase e fenômenos críticos [29, 30]. A teoria de transição de fase e fenômenos críticos teve papel importante no avanço dos estudos de sistemas complexos nas décadas de 1980 e 1990 [31–33]. Apesar de não ser recente o envolvimento de físicos em outras áreas do conhecimento [34], conceitos e métodos da física estatística têm sido ex-

tensivamente aplicadas no estudo de fenômenos correlacionados e não correlacionados, fenômenos críticos e difusão em ciências naturais, biologia e economia, na tentativa de compreender e modelar tais sistemas. Particularmente, nesta tese, empregamos alguns desses conceitos para o estudo de sistemas urbanos e biológicos [13–17].

O primeiro deles é o conceito de invariância de escala, que é uma característica de objetos ou leis que não mudam se a escala de comprimento, energia, tamanho ou outras variáveis podem ser reescaladas por meio de uma operação de multiplicação por um fator em comum. A invariância de escala é uma forma exata de autossimilaridade em que em qualquer ampliação (escala) há uma parte menor do objeto que é semelhante ao todo. Outro conceito é a autossimilaridade, que é uma propriedade típica de geometrias fractais, em que partes do objeto mostram propriedades semelhantes em muitas escalas [35]. Em matemática, um objeto é autossimilar se ele é exatamente ou aproximadamente semelhante a uma parte de si mesmo (isto é, o todo tem a mesma forma de uma ou mais partes). Por exemplo, podemos considerar as propriedades de escala de uma função ou curva $f(x)$ sob operações de escala na variável x , isto é, desejamos uma função na forma $f(\lambda x)$ para alguma escala λ , em que λ pode ser o comprimento, tamanho ou energia. Usualmente, para que $f(x)$ seja invariante de escala e autossimilar, a seguinte propriedade é necessária:

$$f(\lambda x) = \lambda^\delta f(x)$$

para algum valor de δ e dilatação (escala) λ , de forma que essa relação estabelece uma equação homogênea de primeira ordem [25].

Leis de escalas autossimilares (ou leis alométricas) foram encontradas em diversos contextos, de sistemas biológicos [36–40] a sistemas urbanos [41–44]. Uma das mais famosas relações alométricas foi encontrada por Kleiber em 1932 [36] e é conhecida como lei dos 3/4 de Kleiber. Essa lei alométrica estabelece que a relação entre a taxa metabólica e a massa de mamíferos aumenta de uma forma lei de potência com expoente 3/4 (Figura 1). Em outras palavras, animais maiores são mais eficientes em relação ao consumo de energia por massa corporal [36–40].

Um dos resultados mais recentes a respeito de sistemas urbanos relata sobre leis de escalas na relação entre indicadores urbanos e o tamanho da cidade. Alguns exemplos de indicadores urbanos que exibem leis de escalas com a população são patentes, postos de gasolina, produto interno

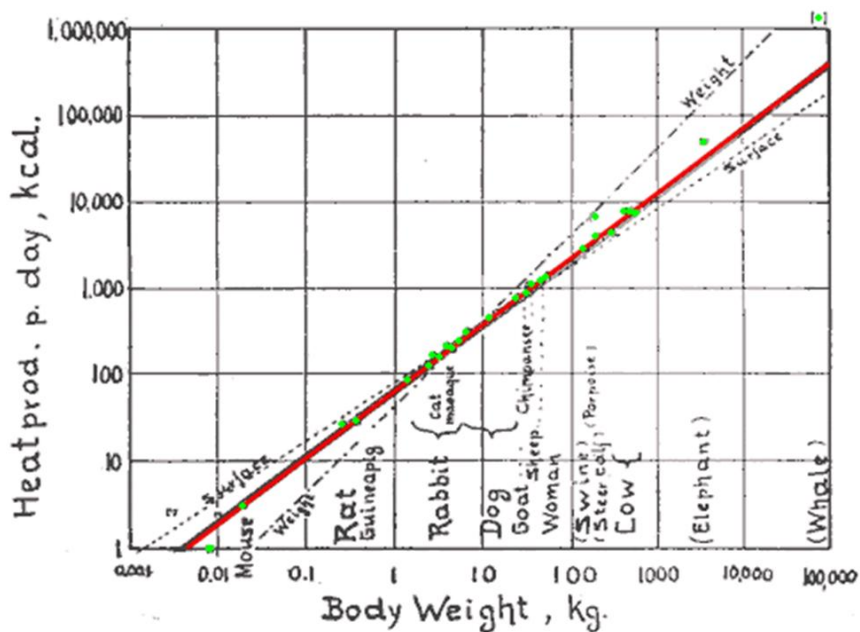


Figura 1: Lei dos 3/4 de Kleiber. Gráfico log-log da produção de calor por dia (kcal) em função do peso corporal (kg) para vários animais. Gráfico reproduzido do artigo de Kleiber M., *Physiological Reviews*, 27, 511–541, 1947 [37].

bruto [41, 42, 45, 46], crime [43, 44, 46–49], indicadores de educação [50], número de candidatos às eleições [51, 52], redes de transportes [53, 54], funcionários de diversos setores [55], medidas de interação social [56]. Analogamente à lei de Kleiber, em que mamíferos maiores são mais eficientes em relação ao consumo de energia por massa corporal, cidades maiores são mais eficientes (em termos *per capita*) com relação ao consumo de recursos como número de postos de gasolinas, comprimento de cabos de energia e área de malha viária (relações sublineares) [41]. Por outro lado, cidades maiores produzem mais riqueza, patentes, problemas sociais e ambientais (também em termos *per capita*) do que as pequenas cidades (relações superlineares) [41]. É nesse contexto de leis de escalas autossimilares que são desenvolvidos os resultados do primeiro capítulo dessa tese. Nesses capítulos, mostramos que devido às não linearidades presentes nas relações entre indicadores urbanos e a população, as métricas *per capita* não são boas para comparar e classificar de cidades. Dessa forma, empregamos uma abordagem que leva em conta essas não linearidades para a classificação/comparação de cidades de uma forma mais justa.

A invariância de escala também está presente no contexto de transições de fase. Quando um sistema sofre transição de fase, as flutuações do sistema podem ser aproximadamente descritas pela

teoria campo médio, em que os expoentes que caracterizam o sistema são chamados de expoentes críticos [29]. O termo transição de fase é mais comumente usado para descrever transições entre estados sólidos, líquidos e gasosos da matéria [23–27]. Os estados físicos e fase da matéria usualmente apresentam propriedades físicas uniformes que durante a transição de fase são alterados de maneira, muitas vezes, abrupta e descontínua como resultado da mudança de alguma condição externa, como temperatura ou pressão [27].

Outro exemplo de transição de fase é a a transição do estado ferromagnético para o paramagnético em temperatura crítica. Os *spins* magnéticos de um material magnético (por exemplo, níquel) interagem uns com os outros. Acima da chamada temperatura de Curie, em média, metade dos *spins* apontam em uma direção e a outra metade na direção oposta. Assim, seus momentos magnéticos se cancelam e o níquel não tem nenhum momento magnético líquido: o material é paramagnético [23]. Abaixo da temperatura de Curie, a maioria dos *spins* no níquel são paralelos e seus momentos magnéticos somam construtivamente e o pedaço de níquel tem um momento magnético resultante não nulo: o material é ferromagnético. Na rede de átomos que compõem o material, os *spins* formam aglomerados/*clusters* que apontam na mesma direção. O tamanho desse aglomerado/*cluster* muda com a temperatura até que, acima da temperatura crítica, o material sofre uma transição de fase tipo percolação em que o maior *cluster* é quebrado em vários fragmentos menores. Baseado nessas ideias de percolação e transição de fase, no segundo capítulo, apresentamos um estudo sobre homicídios *per capita* nas cidades do Brasil. Além disso, vamos discutir as correlações espaciais no número de homicídios *per capita* e como essa dinâmica poderia ser modelada no contexto da cinética estocástica de superfícies na presença de difusão e aleatoriedade ambiental.

O conceito de difusão também é um tema central em mecânica estatística. A difusão é o movimento de moléculas ou átomos de uma região de alta concentração para uma região de baixa concentração. Desde os primeiros relatos feitos pelo botânico Robert Brown, por volta de 1828, a respeito do movimento altamente irregular de partículas suspensas em um fluido [57], e a explicação proposta por Einstein para este fenômeno [58] e também por Smoluchowski [59] de forma independente, os processos difusivos têm sido amplamente investigados em vários contextos. Enquanto em física os estudos de difusão se concentram em partículas e moléculas [60] em um material ou meio, em biologia, economia e sociologia, o estudo da difusão está focado em microrganismos ou células [16, 17, 61, 62], valores de preços, ideias ou rumores [6], respectivamente.

Muitas vezes, em fenômenos naturais, a difusão ocorre de maneira não linear no tempo, em contraste com o processo típico de difusão em que a variância das posições da partícula (dependendo do contexto, microrganismos, ideia ou preço, etc) cresce linearmente no tempo. Nesse caso, a difusão é dita ser anômala [60]. A difusão também pode ser anômala devido a barreiras geométricas que dificultam (ou facilitam) o movimento [13]. É no contexto de difusão anômala que serão apresentados os resultados dos outros três capítulos dessa tese, sendo que o terceiro e o quarto capítulo estão relacionados à dinâmica de organismos biológicos (protozoários e vermes) e o quinto capítulo é de uma vertente teórica que pode ter aplicações diretas para descrever a difusão em sistemas complexos, tais como as células vivas.

Em resumo, essa tese aplica métodos de física estatística no estudo dos problemas que aparecem na organização deste trabalho da seguinte forma: capítulo 1: quantificação e predição do desempenho de cidades [14]; capítulo 2: correlações espaciais e transições de fase tipo percolação no crime de homicídios *per capita* [15]; capítulo 3: análise de correlação e difusão na motilidade de protozoários [16]; capítulo 4: análise de correlação e dinâmica fractal em *C. elegans* [17]; capítulo 5: investigação de mecanismos de difusão anômala [13]. Além dos artigos descritos nessa tese, os quais constituem cada um dos capítulos aqui apresentados, cabe mencionar, por completeza, que o autor também participou de outros projetos que resultaram em publicações nas áreas de ruídos crepitantes [63], mecânica quântica [64] e vazão de rios [65].

CAPÍTULO 1

Métricas ajustadas à escala para prever a evolução de indicadores urbanos e quantificar o desempenho de cidades

Neste capítulo, estudamos a relação entre oito indicadores urbanos e o tamanho da população nas cidades brasileiras. Mostramos que as relações entre métricas urbanas e a população são, em geral, não lineares e podem ser descritas por leis alométricas. Por isso, argumentamos que o uso de métricas *per capita* para comparar o desempenho de cidades favorecem cidades pequenas ou grandes, dependendo se as alometrias são do tipo superlinear ou sublinear. Assim, levamos em conta esse viés, avaliando a diferença entre o valor real de um indicador urbano e o valor esperado pela alometria com o tamanho da população. Proposta inicialmente por Bettencourt *et al.* [46], essa métrica ajustada à escala fornece uma forma mais adequada de avaliar/comparar cidades com diferentes tamanhos de população. Mostramos que essa métrica ajustada à escala é linearmente correlacionada com seus valores passados, assim como a métrica ajustada à escala de cada indicador possui correlações lineares com as demais métricas urbanas ajustadas à escala. Devido a essa característica, usamos modelos lineares para prever os valores dos indicadores urbanos. Ao empregar esses modelos para visualizar espacialmente suas previsões para o ano de 2020, verificamos o surgimento de aglomerados/*clusters* espaciais caracterizados por regiões do território brasileiro onde se espera um aumento ou uma diminuição nos valores desses indicadores urbanos.

1.1 Introdução e apresentação do problema

Nas últimas seis décadas, o mundo passou por um período de rápida e notável urbanização. De acordo com as Nações Unidas [66] no ano de 2007, pela primeira vez, a população urbana mundial superou a rural e, se esse processo continuar, são esperados que dois terços da população mundial viva em áreas urbanas até o ano de 2050. Por um lado, cidades são normalmente associadas a níveis mais elevados de alfabetização, cuidados de saúde e melhores oportunidades. Por outro lado, a urbanização não planejada e as más decisões políticas geram poluição, degradação ambiental, aumento da criminalidade, desigualdade de oportunidades e aumento do número de pessoas que vivem em condições precárias. Nesse sentido, há uma grande necessidade de encontrar padrões, quantificar e prever a evolução de indicadores urbanos, uma vez que essas investigações podem fornecer orientações para melhores decisões políticas e alocação de recursos.

Fomentados por essa necessidade e devido à disponibilidade de uma quantidade sem precedentes de dados a nível de cidades, recentemente, vários pesquisadores têm promovido um progresso impressionante no campo de pesquisa que tem sido chamado de “*Science of Cities*” [67]. Esses novos dados permitiram aos pesquisadores investigarem padrões relacionados às cidades em um grau nunca antes possível. Nesse contexto, um dos resultados mais surpreendentes e universais foi a descoberta de leis de escala alométricas robustas entre vários indicadores urbanos com o tamanho da população. Patentes, postos de gasolina, produto interno bruto [41, 42, 45, 46], crime [43, 44, 46–49, 68], suicídios [50], número de candidatos às eleições [51, 52], redes de transportes [53, 54], funcionários de diversos setores [55], medidas de interação social [56], emissão de CO₂ [69–71] são apenas alguns exemplos de indicadores em que as leis de escala foram encontradas. Similarmente ao que acontece em biologia, em que animais maiores são mais eficientes em relação ao consumo de energia [36, 37], cidades maiores podem ser mais eficientes (em termos *per capita*) dependendo do indicador urbano em questão [41], mas, diferentemente do caso biológico, cidades maiores também podem produzir mais recursos e problemas em termos *per capita*. Apesar dessas relações não lineares intrínsecas, é comum encontrar trabalhos que tentam descobrir relações entre o tamanho da população e indicadores urbanos empregando regressão linear nos dados brutos. Além disso, indicadores *per capita* (ou seja, indicadores divididos pelo tamanho da população) são onipresentes nos relatórios de agências governamentais e, muitas vezes, são utilizados como um

guia para políticas públicas na análise da evolução temporal de uma determinada cidade ou para comparar/classificar o desempenho de cidades com diferentes tamanhos de população. No entanto, essas regressões lineares podem resultar em relações controversas [44, 72]. Indicadores *per capita* são indiferentes às leis de escala alométricas que fazem da cidade um aglomerado complexo que não pode ser modelado como uma combinação linear de seus componentes individuais [46, 73, 74]. Nesse sentido, há uma escassez de métricas alternativas para indicadores urbanos que podem superar os problemas citados anteriormente e que forneçam uma comparação mais justa entre cidades de diferentes tamanhos, bem como uma melhor compreensão da evolução de cidades. Recentemente, Bettencourt *et al.* [46] propuseram uma alternativa simples para superar essas não linearidades por meio do cálculo da diferença entre o valor real dos indicadores urbanos e o valor esperado pelas alometrias com o tamanho da população (isto é, os resíduos nas relações alométricas). Essa métrica ajustada à escala considera, explicitamente, as relações alométricas e já provou ser útil no contexto econômico [75, 76] e para revelar as relações que não são devidamente observadas por análise de regressão quando se investiga correlações entre crimes e métricas urbanas [44].

Nesse capítulo, observamos a evolução dessa métrica relativa em oito indicadores urbanos de cidades brasileiras em três anos (1991, 2000 e 2010), nos quais o censo nacional foi realizado. Ao agrupar as cidades em acima e abaixo das leis alométricas, argumentamos que a média dessa métrica ajustada à escala fornece uma forma mais adequada de resumir a evolução dos indicadores urbanos, quando comparada aos valores *per capita*. Essa métrica também revela padrões que não aparecem quando se analisa somente a evolução dos valores *per capita*. Por exemplo, enquanto os valores de homicídios *per capita* têm aumentado sistematicamente ao longo das últimas três décadas, as médias da métrica ajustada à escala para cidades acima e abaixo da lei alométrica estão se aproximando de zero, isto é, cidades onde o número de homicídios está acima do esperado pela alometria conseguiram reduzir esse crime, enquanto que esse crime aumentou nas cidades onde o número de homicídios está abaixo da alometria. Argumentamos que as não linearidades podem afetar os indicadores *per capita*, criando um viés para cidades grandes em alometrias superlineares e um viés para as cidades pequenas em alometrias sublineares. Mostramos ainda que essas métricas ajustadas à escala estão fortemente correlacionadas aos seus valores em tempos anteriores por uma correspondência linear, tornando-as particularmente boas para prever valores futuros de indicadores urbanos via regressões lineares. Testamos essa hipótese por meio de um modelo linear em

que a métrica ajustada à escala para um indicador em um determinado censo foi prevista por uma combinação linear de todas as oito métricas no censo anterior. Esses modelos simples explicam de 31% a 97% da variância observada nos dados e reproduzem o valor médio da métrica ajustada à escala ao agrupar as cidades em acima e abaixo das leis alométricas. Motivados por esses bons resultados, apresentamos uma previsão para os valores dos indicadores urbanos do ano de 2020, assumindo que os coeficientes lineares ao longo do tempo sejam constantes. Ao visualizar as mudanças previstas, verificamos o surgimento de aglomerados espaciais caracterizados por regiões do território brasileiro onde o modelo prevê um aumento ou uma diminuição nos valores dos indicadores urbanos. Em Ref. [14] reportamos uma lista contendo todas as métricas ajustadas à escala, bem como as previsões para cada cidade, na expectativa de que as agências governamentais possam empregar essas informações.

1.2 Apresentação dos dados

Os dados analisados consistem no tamanho da população, $N(t)$, e oito indicadores urbanos, $Y_i(t)$, para cada cidade brasileira nos anos $t = 1991, 2000$ e 2010 nos quais ocorreram o censo nacional. Desses dados, selecionamos as 1605 cidades para as quais todos os oito indicadores urbanos estão disponíveis, o que corresponde a 28,8% do número total de cidades brasileiras, mas que representam 76,5% do total da população do Brasil. Esses dados são mantidos e disponibilizados gratuitamente pelo Departamento de Informática do Sistema Único de Saúde — DATASUS [77]. As oito métricas urbanas são definidas da seguinte forma:

- *trabalho infantil*: a proporção da população com idade entre 10 e 15 anos que está trabalhando ou procurando trabalho na semana de referência, em uma determinada área geográfica, no ano de referência;
- *população idosa*: o número de habitantes de uma dada cidade com 60 anos ou mais;
- *população feminina*: o número de habitantes de uma determinada cidade que é do sexo feminino;
- *homicídios*: número de mortes causadas por ferimentos infligidos por outra pessoa com a intenção de ferir ou matar, por qualquer meio;

- *analfabetismo*: o número de habitantes de uma determinada área geográfica, no ano de referência, com 15 anos ou mais, que não sabem ler e escrever pelo menos um bilhete simples no idioma que eles falam;
- *renda familiar*: rendimento médio domiciliar dos residentes de uma determinada área geográfica, no ano de referência. Foi considerado como renda familiar a soma do rendimento mensal do agregado familiar, em Reais, dividido pelo número de habitantes na residência;
- *população masculina*: o número de habitantes de uma determinada cidade que é do sexo masculino;
- *desemprego*: o número de habitante com 16 anos ou mais, que está sem trabalhar ou a procura de trabalho na semana de referência, em uma determinada área geográfica, no ano de referência.

Apesar de haver outras definições [78], os resultados que apresentamos foram obtidos considerando-se que as cidades são as menores unidades administrativas com um governo local (municípios). Uma outra definição amplamente utilizada é a de área metropolitana, que é composta por mais de um município e normalmente está associada à coalescência de vários municípios. Como discutido por Bettencourt *et al.* [79], a escolha da “unidade de análise” é crucial quando se estuda as propriedades de cidades. No que diz respeito à análise de escala, por um lado, a desagregação da definição correta de cidade pode introduzir um desvio no valor do expoente de escala (quer reduzindo ou aumentando o seu valor esperado); por outro lado, a agregação da definição correta de cidade normalmente faz a alometria mais linear [79]. Mudanças nos expoentes de escala foram relatadas na escolha de diferentes definições de cidade [70, 79, 80]. No entanto, não há nenhum procedimento à prova de falhas para a definição correta dos limites de uma cidade [80–82]. Além disso, alguns indicadores urbanos são realmente mais espacialmente restritos do que outros (por exemplo, homicídios *versus* renda familiar). Uma alternativa para identificar os limites urbanos foi proposta por Rozenfeld *et al.* [83]. O método proposto, chamado de *City Clustering Algorithm* (CCA), é baseado em distribuições espaciais da população em uma escala geográfica fina (escala menor que a de nossos dados), definindo uma cidade além do alcance de seus limites administrativos, formando as áreas metropolitanas de estatística [83]. Nesse capítulo, também analisamos

os dados depois de considerar simultaneamente os municípios que não pertencem a qualquer área metropolitana e agregar os municípios das 39 áreas metropolitanas existentes no Brasil de acordo com a definição do IBGE [84]. Apesar de observarmos algumas mudanças pequenas nos expoentes de escala, nossas conclusões permanecem inalteradas nesse cenário.

1.3 Resultados e discussão

Começamos considerando a média dos valores *per capita* dos oito indicadores urbanos descritos na seção anterior. Essa é uma prática comum em agências governamentais para acompanhar a evolução de uma determinada cidade ou para comparar um grupo de cidades com populações diferentes. Observamos na Figura 1.1 que quase todos os indicadores *per capita* mostram uma tendência de evolução temporal clara: população de idosos, população feminina, homicídios e renda familiar têm aumentado ao longo dos anos; enquanto que trabalho infantil, analfabetismo e população masculina diminuíram (as taxas de desemprego têm evoluído de uma forma mais complexa, não exibindo uma tendência clara). Também poderíamos listar as cidades em que esses indicadores foram consideravelmente alterados ou classificar as cidades que fizeram mais progressos na redução, por exemplo, das taxas de homicídios ou de analfabetismo.

Um dos principais problemas com essa análise é que ela ignora completamente a hipótese de que a maioria dos indicadores urbanos $Y_i(t)$ exibe uma relação alométrica (lei de potência) com a população $N(t)$, isto é,

$$Y_i(t) = 10^{\mathcal{A}_i} N(t)^{\beta_i}, \quad (1.1)$$

em que \mathcal{A}_i é uma constante, β_i é o expoente alométrico (ou expoente de escala) e t é o ano. Essa relação simples resume os efeitos (médios) nos indicadores urbanos quando aumentamos o tamanho da população. Ela afirma que as cidades são autossemelhantes (*self-similar*) em termos de sua população, no sentido de que as propriedades médias de uma determinada cidade podem ser inferidas apenas pelo conhecimento de sua população. Portanto, são esperados que os indicadores urbanos exibam uma componente linear, emergindo de propriedades gerais das redes urbanas relacionadas aos aspectos sociais e de infraestrutura das cidades [74]. Ao analisar os valores *per capita*, assumimos, implicitamente, que o valor de um indicador urbano é proporcional ao tamanho da população ($\beta_i = 1$), ou, em outras palavras, que as cidades são sistemas extensivos. Essa ideia

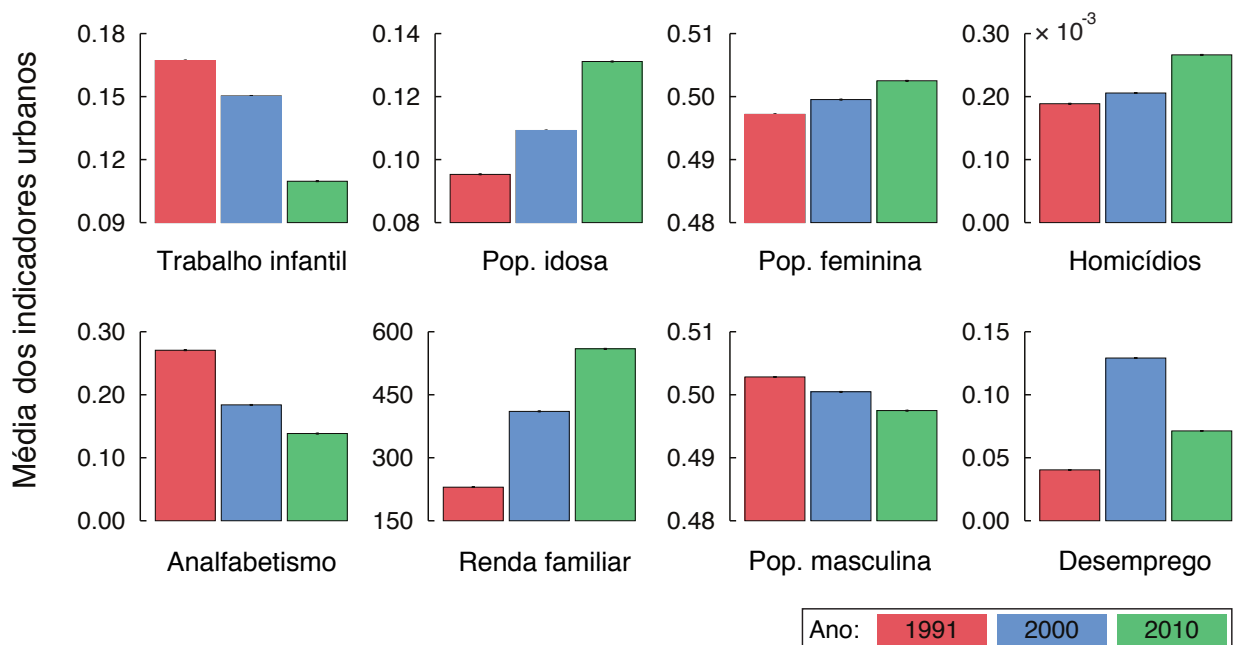


Figura 1.1: Tendências na evolução dos valores *per capita* de indicadores urbanos brasileiros. Os gráficos de barra mostram a evolução temporal dos oito indicadores urbanos nos anos de 1991, 2000 e 2010, nos quais o censo nacional ocorreu. As alturas das barras correspondem aos valores médios obtidos para as cidades brasileiras, em que cada indicador urbano para cada cidade está dividido pela população da cidade correspondente, isto é, de forma *per capita*. A definição de cada indicador é fornecida na Seção 1.2. As pequenas barras de erro são intervalos de confiança de 95% para os valores médios, via *bootstrapping*. Notamos que população de idosos, população feminina, homicídios e renda familiar exibem uma tendência crescente; enquanto trabalho infantil, analfabetismo e população masculina têm diminuído ao longo do tempo. Apesar de terem aumentado, as taxas de desemprego não mostram uma tendência muito clara.

é oposta à abordagem de cidades como sistemas complexos: sistemas complexos são não extensivos ($\beta_i \neq 1$), o que significa que suas partes isoladas não se comportam da mesma maneira como quando estão interagindo entre si. As cidades têm propriedades semelhantes e só fazem sentido como um “organismo” inteiro e, de fato, há evidências robustas favorecendo a natureza não extensiva e universal de cidades (ou seja, a hipótese de escala urbana) em diferentes culturas e períodos históricos [74, 85, 86]. Assim, várias propriedades de uma cidade de um determinado tamanho não podem ser escaladas de forma linear para uma outra cidade com tamanho de população maior ou menor. Os processos dinâmicos mediados pelas redes urbanas tornam a operação de escala uma transformação não linear para vários indicadores urbanos, muitas vezes, resultando em uma economia *per capita* de infraestrutura material e no ganho de produtividade socioeconômica [41]. De

um ponto de vista mais técnico, sempre que o expoente alométrico β_i é diferente de um, há uma componente remanescente relacionada ao tamanho da população ao avaliar os valores *per capita* desses indicadores urbanos, ou seja,

$$y_i = Y_i/N \sim N^{(\beta_i-1)}, \quad (1.2)$$

o que cria um viés para grandes cidades quando $\beta_i > 1$ e para as pequenas cidades quando $\beta_i < 1$. Medidas *per capita* só são eficientes na remoção correta do efeito do tamanho da população em um indicador urbano se $\beta_i = 1$.

Vários trabalhos relatam a existência de relações alométricas robustas entre vários indicadores urbanos e o tamanho da população. Em relação às cidades brasileiras, as leis de escala já foram identificadas para vários indicadores urbanos [42–44, 47–52], principalmente por causa da existência de dados confiáveis disponibilizados por agências brasileiras, tais como o DATASUS e o IBGE. Queremos confirmar que a hipótese de escala urbana é válida para todos os nossos indicadores urbanos e se essas alometrias mudaram ao longo do tempo. Especificamente, testamos a hipótese de que um indicador urbano pode ser descrito por uma função lei de potência do tamanho da população, ou seja,

$$Y_i(t) = 10^{\mathcal{A}_i} N(t)^{\beta_i}, \quad (1.3)$$

em que β_i é o expoente alométrico (ou escala) e \mathcal{A}_i é uma constante. Para isso, analisamos a relação entre o logaritmo de cada indicador urbano e o logaritmo da população e via um modelo de regressão linear de distância ortogonal (implementado no pacote `scipy.odr` da biblioteca `SciPy` da linguagem *Python* [87]). Embora as relações empíricas apresentem diferentes graus de dispersão, todas elas exibem relações lineares de boa qualidade (correlação de Pearson variando de 0,44 a 1,00 — Tabela 1.1) que são bem descritas por modelos lineares em escala logarítmica (Figuras 1.2, 1.3 e 1.4). A partir da Tabela 1.1, notamos que os valores de β_i para analfabetismo, renda familiar e desemprego exibem uma fraca tendência decrescente ao longo dos anos, enquanto os outros indicadores mostram apenas pequenas flutuações (sem uma clara tendência evolutiva). A tendência decrescente fraca para o desemprego também aparece no trabalho de Ignazzi [49] para os mesmos dados nos anos de 2000 e 2010. Assim, os valores de β_i classificam nossos indicadores em dois grupos: população feminina, homicídios e desemprego têm relações superlineares com a

população ($\beta_i > 0$); enquanto trabalho infantil, população idosa, analfabetismo, renda familiar e população masculina têm relações sublineares ($\beta_i < 0$). É interessante notar que, apesar dos expoentes alométricos β_i estarem perto de um para as populações idosas, femininas e masculinas, as alometrias entre esses indicadores e a população total são quase perfeitamente correlacionadas, produzindo valores de β_i muito próximos, mas estatisticamente diferentes de um. Observamos, que os valores para os expoentes alométricos β_i relatados aqui podem diferir ligeiramente dos resultados relatados anteriormente devido às definições urbanas diferentes e a outros procedimentos de ajuste. No entanto, essas discrepâncias são frequentemente muito pequenas (por exemplo, considerando mínimos quadrados generalizados via procedimento de Cochrane-Orcutt e outra definição de cidade, Ignazzi [49] encontrou $\beta = 1,23$ e $\beta = 1,19$ para o desemprego, respectivamente, nos anos de 2000 e 2010).

O problema com indicadores *per capita* nos conduz, portanto, a questão sobre como podemos levar em conta essas alometrias e corretamente remover o efeito do tamanho da população sobre os indicadores urbanos. Bettencourt *et al.* [46] propuseram um procedimento simples e eficiente para superar esse problema, definindo o chamado indicador urbano ajustado à escala — ou, em inglês, *scale-adjusted metric*. A abordagem consiste em avaliar a diferença logarítmica entre o valor real de um indicador urbano $Y_i(t)$ e o valor esperado pela relação alométrica com a população $N(t)$ (isto é, os resíduos nas relações alométricas) em um determinado ano t , ou seja,

$$D_{Y_i}(t) = \log Y_i(t) - [\mathcal{A}_i + \beta_i \log N(t)]. \quad (1.4)$$

A quantidade anterior considera explicitamente a alometria entre um indicador urbano e o tamanho da população, criando uma medida relativa que não é influenciada pela tamanho da população (independente do tamanho) para qualquer valor de β_i . A métrica ajustada à escala $D_{Y_i}(t)$ assimila a excepcionalidade (boa ou má) de uma cidade, que o resultado do processo de aglomeração não linear relacionado às escolhas socioeconômicas e ao percurso histórico de uma cidade [46]. Além disso, D_{Y_i} estabelece uma escala mais “natural” para a comparação entre cidades, identificando se um indicador urbano de uma determinada cidade está acima ($D_{Y_i} > 0$) ou abaixo ($D_{Y_i} < 0$) do valor esperado para cidades de tamanhos semelhantes. Essa abordagem já provou ser útil no contexto econômico [75, 76] e também foi empregada para revelar relações que não são devidamente

Tabela 1.1: Relações alométricas entre indicadores urbanos e o tamanho da população. Os valores dos parâmetros \mathcal{A}_i e β_i obtidos via regressão de distância ortogonal sobre a relação entre $\log Y_i(t)$ e $\log N(t)$ para cada indicador urbano no ano t . Os valores dentro dos parênteses são os erros padrão (SE) na última casa decimal dos valores estimados dos parâmetros. A última coluna mostra os valores do coeficiente de correlação linear de Pearson (ρ) para cada alometria.

Indicador $Y_i(t)$	Ano t	\mathcal{A}_i (SE)	β_i (SE)	ρ
Trabalho infantil	1991	-0,64 (5)	0,96 (1)	0,909
	2000	-0,55 (5)	0,93 (1)	0,906
	2010	-0,80 (5)	0,95 (1)	0,884
Pop. idosa	1991	-0,99 (5)	0,992 (6)	0,976
	2000	-0,83 (2)	0,969 (5)	0,980
	2010	-0,72 (2)	0,963 (5)	0,982
Pop. feminina	1991	-0,367 (3)	1,014 (1)	1,000
	2000	-0,361 (2)	1,013 (1)	1,000
	2010	-0,355 (2)	1,012 (1)	1,000
Homicídios	1991	-5,4 (1)	1,35 (3)	0,769
	2000	-5,7 (1)	1,41 (2)	0,800
	2010	-5,04 (9)	1,29 (2)	0,827
Analfabetismo	1991	-0,29 (7)	0,92 (2)	0,789
	2000	-0,26 (7)	0,87 (2)	0,774
	2010	-0,28 (7)	0,85 (2)	0,749
Renda familiar	1991	0,82 (6)	0,33 (1)	0,428
	2000	1,14 (6)	0,31 (1)	0,440
	2010	1,63 (5)	0,23 (1)	0,440
Pop. masculina	1991	-0,239 (3)	0,987 (1)	1,000
	2000	-0,243 (2)	0,987 (1)	1,000
	2010	-0,249 (2)	0,988 (1)	1,000
Desemprego	1991	-3,50 (8)	1,45 (2)	0,880
	2000	-2,07 (5)	1,25 (1)	0,940
	2010	-2,09 (5)	1,20 (1)	0,931

compreendidas por análises de regressão linear quando investigamos as correlações entre crime e métricas urbanas [44]. A Figura 1.4 ilustra a definição de $D_{Y_i}(t)$ e mostra um exemplo de alometria entre o número de homicídios e tamanho da população (o mesmo mostrado na Figura 1.2).

A fim de mostrar as informações que essa métrica ajustada à escala fornece, estudamos a evolução da média de $D_{Y_i}(t)$ para dois grupos de cidades: aquelas cujo indicador urbano $Y_i(t)$ estava acima ($D_{Y_i}(t) > 0$) e aquelas cujo indicador $Y_i(t)$ estava abaixo ($D_{Y_i}(t) < 0$) da lei alométrica no ano de 1991. A Figura 1.5 mostra essas médias para os oito indicadores urbanos ao longo dos três anos dos nossos dados. Para a maioria dos indicadores urbanos, a média de $D_{Y_i}(t)$ exibe uma tendência decrescente e estatisticamente significativa para as cidades inicialmente acima da lei alo-

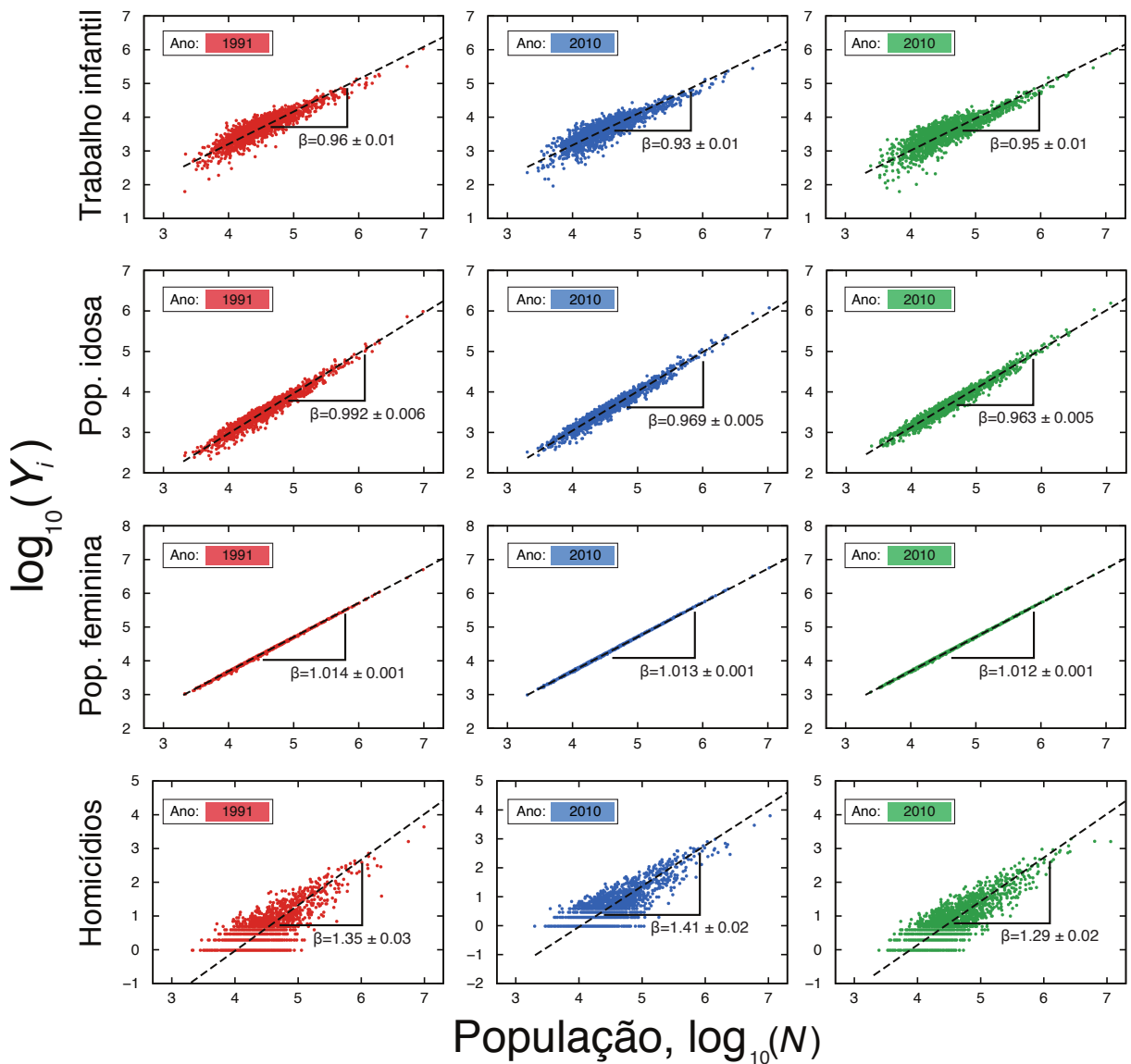


Figura 1.2: Leis alométricas com o tamanho da população. Os gráficos de dispersão mostram as relações alométricas entre os indicadores urbanos (de cima para baixo: trabalho infantil, população de idosos, população feminina e homicídios) e o tamanho da população para os anos $t = 1991$ (pontos vermelhos), 2000 (pontos azuis) e 2010 (pontos verdes) em escala log-log. Os expoentes alométricos β_i são mostrados na figura. Veja Figura 1.3 para os outros indicadores.

métrica; por outro lado, observamos uma tendência crescente para as médias de $D_{Y_i}(t)$ para as cidades que estavam inicialmente abaixo da lei alométrica. Por exemplo, cidades onde o número de crianças trabalhando, homicídios e desemprego estava acima do valor esperado pela lei alométrica foram bem sucedidas (em média) em reduzi-los; em contraste, cidades onde esses indicadores

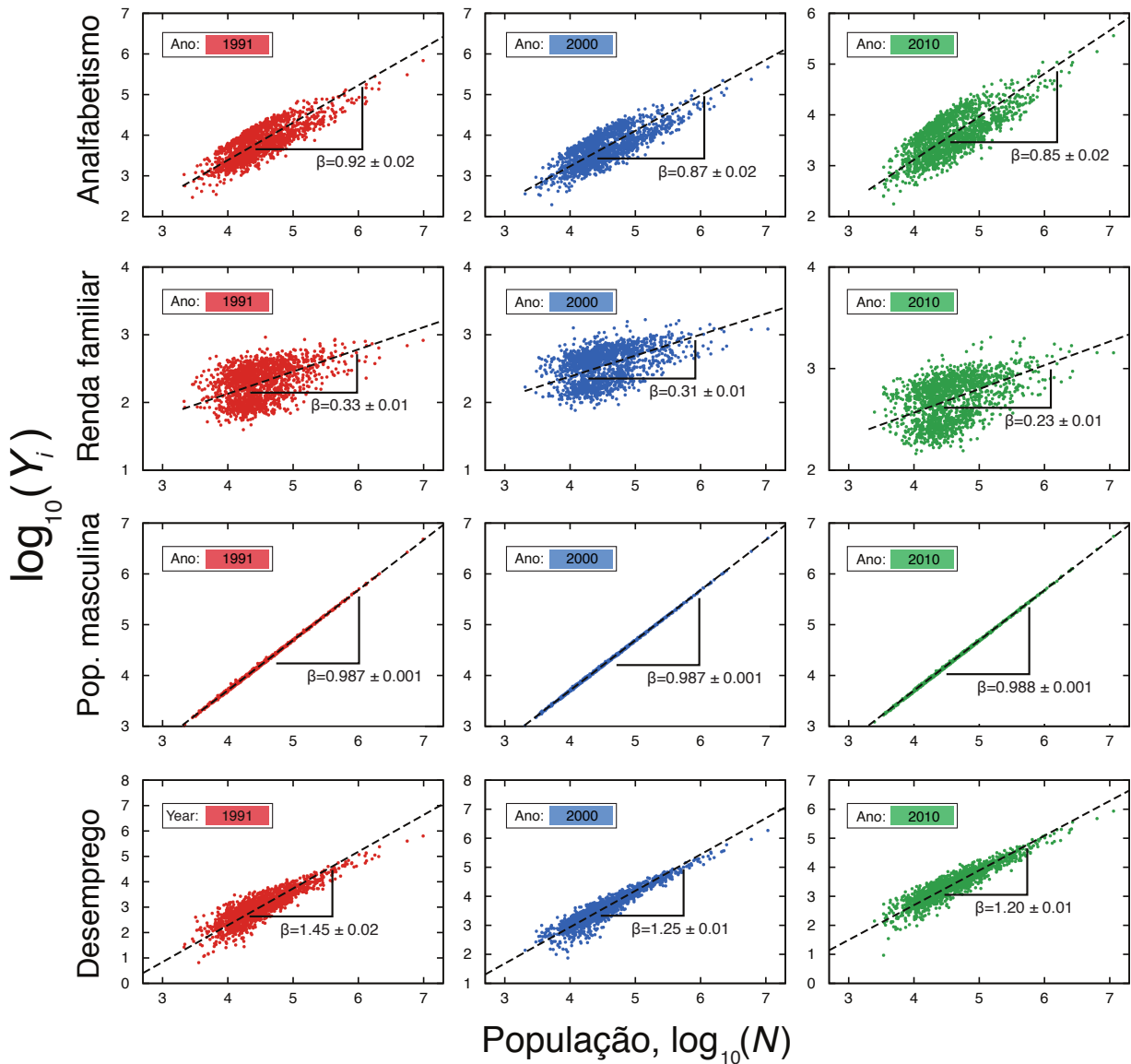


Figura 1.3: Leis alométricas com o tamanho da população. O mesmo que a Figura 1.2 para os indicadores de analfabetismo, renda familiar, população masculina e desemprego.

estavam abaixo das leis alométricas se mostraram incapazes (em média) de melhorar ou manter essa situação. O caso do analfabetismo é uma exceção ao padrão anterior, uma vez que a média de $D_{Y_i}(t)$ aumentou para cidades inicialmente acima da lei alométrica e é quase uma constante para as cidades inicialmente abaixo. Assim, cidades nas quais o analfabetismo estava inicialmente acima da lei alométrica falharam (em média) em aumentar o número de alfabetizados; por outro lado, as cidades inicialmente abaixo da lei alométrica não apenas mantiveram essa situação (em

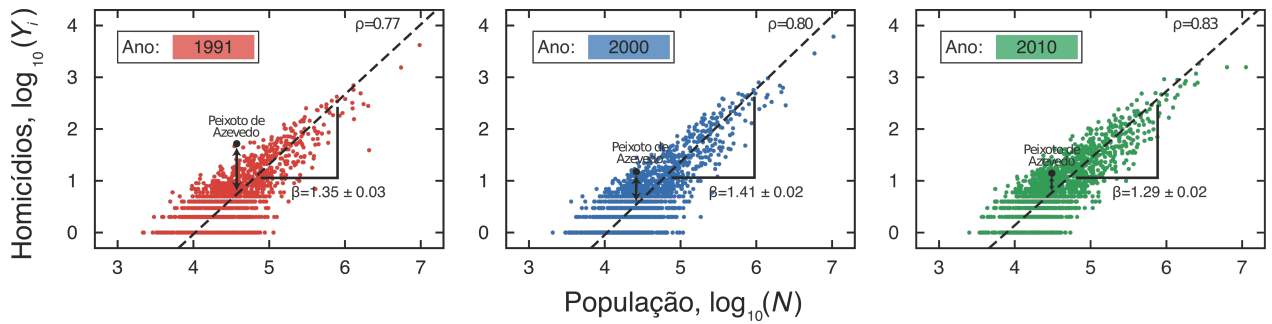


Figura 1.4: Leis alométricas e a definição da métrica ajustada à escala $D_{Y_i}(t)$. Os gráficos mostram as relações alométricas entre o número de homicídios e o tamanho da população para os anos de 1991, 2000 e 2010 em escala log-log. Os expoentes alométricos β_i e o coeficiente de correlação de Pearson ρ são mostrados nas figuras. Destacamos uma cidade em particular (Peixoto de Azevedo) nos três anos para ilustrar a evolução de $D_{Y_i}(t)$. Para essa cidade, o número de homicídios estava bem acima da lei alométrica no ano de 1991; no entanto, tornou-se mais próximo do valor esperado pela lei alométrica ao longo dos anos.

média), mas também conseguiram melhorar ainda mais os níveis de alfabetização. Um fenômeno parecido foi observado por Christakis *et al.* [88] para o caso de obesidade. As chances de uma pessoa se tornar obesa aumentam se ela tem algum um amigo ou parente próximo (irmão, marido, esposa, etc) que se tornou obeso em um dado intervalo de tempo [88]. Nesse contexto, podemos imaginar que os aumentos nos níveis de analfabetismo estão relacionados às redes de interação pessoa-pessoa nas cidades, de forma que a falta de conhecimento pode se espalhar de forma similar às epidemias. No caso de métricas populacionais, especialmente sobre as populações feminina e masculina, a aproximação às leis alométricas (juntamente com a diminuição da proporção de homens na população) podem representar um bom aspecto com relação à redução da violência, uma vez que um contingente excessivo de homens podem produzir um aumento no comportamento antissocial [89]. Além disso, tanto a população masculina acima e população feminina abaixo da lei alométrica correlacionam com o número de homicídios acima da lei alométrica [44]. Da mesma forma, a renda familiar acima da lei alométrica correlaciona-se com homicídios acima, enquanto a renda familiar abaixo da lei alométrica correlaciona-se com homicídios abaixo [44]. Notavelmente, essas informações permanecem ocultas ou distorcidas quando olhamos apenas para os valores *per capita* desses indicadores urbanos.

As tendências nos valores médios das métricas ajustadas à escala levam a questão de como os valores dessa métrica relativa são afetados por seus valores anteriores, ou seja, os valores de

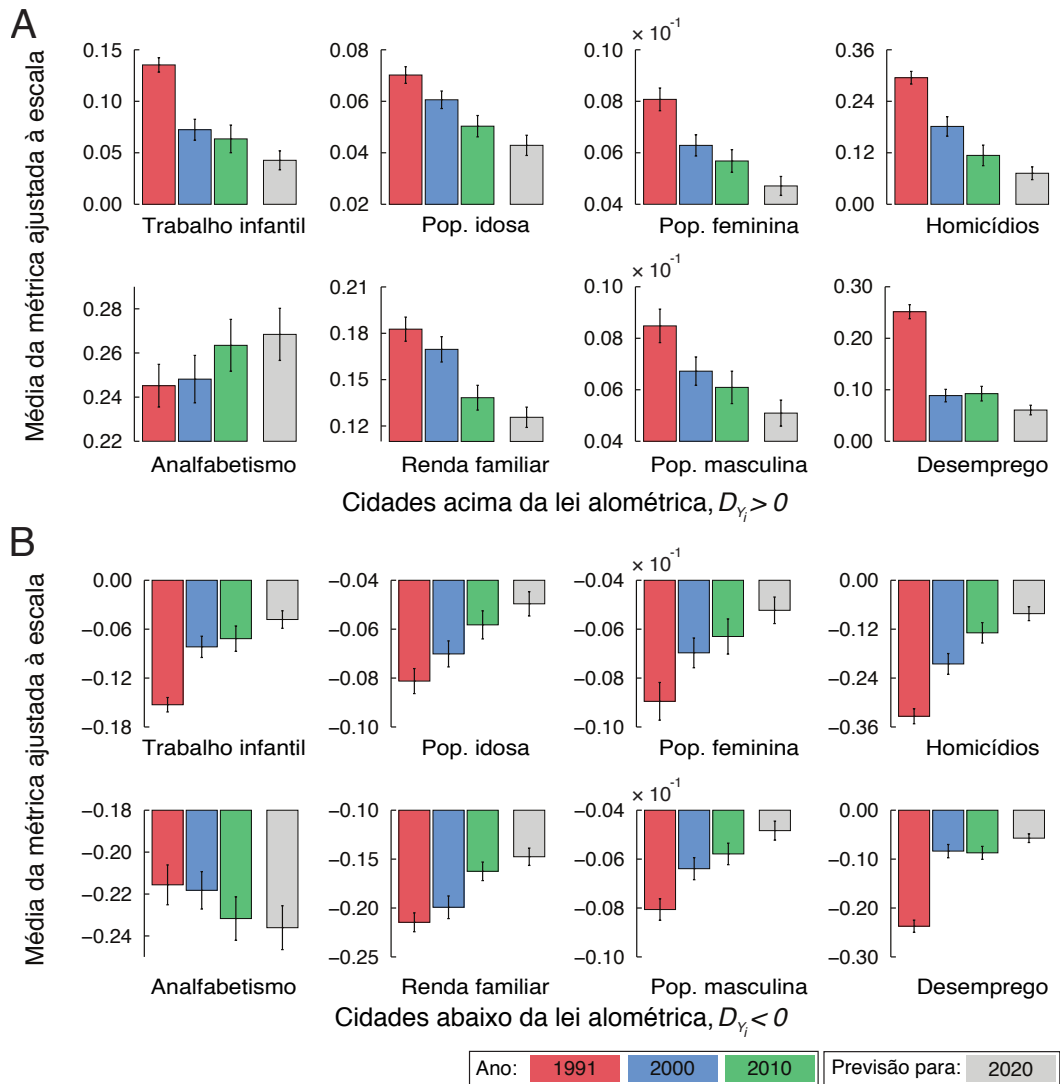


Figura 1.5: Tendências na evolução dos valores médios das métricas ajustadas à escala. Os gráficos de barra mostram a evolução da média de $D_{Y_i}(t)$ ao agrupar as cidades cujos indicadores urbanos estavam A) acima e B) abaixo da lei alométrica no ano de $t = 1991$. As barras coloridas são os dados empíricos (vermelho para $t = 1991$, azul para $t = 2000$ e verde para $t = 2010$) e as barras cinzas representam as previsões obtidas a partir do modelo linear da Equação 1.8 (ver discussão no texto principal). As barras de erro são intervalos de confiança de 95% via *bootstrapping* para os valores médios. Notamos que a média de $D_{Y_i}(t)$ exibe uma tendência decrescente e estatisticamente significativa para as cidades cujo indicador urbano estava inicialmente acima da lei alométrica, e é observada uma tendência crescente para as cidades cujo indicador urbano estava inicialmente abaixo da lei alométrica. A única exceção é o caso do analfabetismo, para o qual o valor médio de $D_{Y_i}(t)$ aumentou para cidades inicialmente acima da lei alométrica e é aproximadamente constante para as cidades inicialmente abaixo. Observamos, ainda, que os valores previstos (barras cinza) mantêm a tendência principal dos dados empíricos. O modelo para prever D_{Y_i} será discutido mais adiante.

$D_{Y_i}(t + \Delta t)$ e $D_{Y_i}(t)$ são correlacionados de alguma forma particular? Para responder essa pergunta, analisamos os gráficos de dispersão da métrica ajustada à escala em um determinado ano *versus* os seus valores em tempos anteriores para cada indicador urbano para todas as cidades. A Figura 1.6 mostra esses gráficos de dispersão, considerando os valores de $D_{Y_i}(2010)$ *versus* $D_{Y_i}(2000)$ e a Figura 1.7 mostra os gráficos para $D_{Y_i}(2000)$ *versus* $D_{Y_i}(1991)$. Observamos que, apesar dos diferentes graus de dispersão (correlação de Pearson variando de 0,47 a 0,98), funções lineares são boas aproximações para a tendência média dessas relações. Portanto, ajustamos o seguinte modelo linear

$$D_{Y_i}(t + \Delta t) = A_i + \alpha_i D_{Y_i}(t) \quad (1.5)$$

para cada indicador urbano usando o método dos mínimos quadrados e os melhores parâmetros α_i (assim como os coeficientes de correlação de Pearson) são mostrados na Tabela 1.2 para as duas combinações de anos. Omitimos os valores de A_i , porque eles são muito pequenos ($\sim 10^{-6}$). Apesar da tendência crescente observada para os valores de α_i ao longo dos anos (para homicídios e renda familiar), observamos que $\alpha_i < 1$ para quase todos os indicadores urbanos, exceto para o analfabetismo ($\alpha_i = 1,01 \pm 0,01$ para 2000 *versus* 1991 e $\alpha_i = 1,05 \pm 0,01$ para 2010 *versus* 2000). Esses resultados estão de acordo com a evolução da média de $D_{Y_i}(t)$, ou seja, os indicadores caracterizados por $\alpha_i < 1$ apresentam uma tendência de se aproximar das leis alométricas médias, enquanto para $\alpha_i \geq 1$ há uma tendência de se afastar das leis alométricas médias.

A fim de compreender melhor o papel do parâmetro α_i , consideramos o limite em que Δt é muito pequeno para reescrever a Equação 1.5 como a seguinte equação diferencial

$$\frac{d}{dt} D_{Y_i}(t) = A_i + (\alpha_i - 1) D_{Y_i}(t), \quad (1.6)$$

cuja solução é

$$D_{Y_i}(t) = \begin{cases} A_i/(1 - \alpha_i) + [k - A_i/(1 - \alpha_i)] \exp[-(1 - \alpha_i)t] & (\text{se } \alpha \neq 1), \\ A_i t + k & (\text{se } \alpha = 1), \end{cases} \quad (1.7)$$

onde k é uma constante de integração. Assim, a Equação 1.7 prevê que $D_{Y_i}(t)$ aproxime exponencialmente do valor $A_i/(1 - \alpha_i)$ para $\alpha_i < 1$ e que $D_{Y_i}(t)$ aumente exponencialmente ao longo do

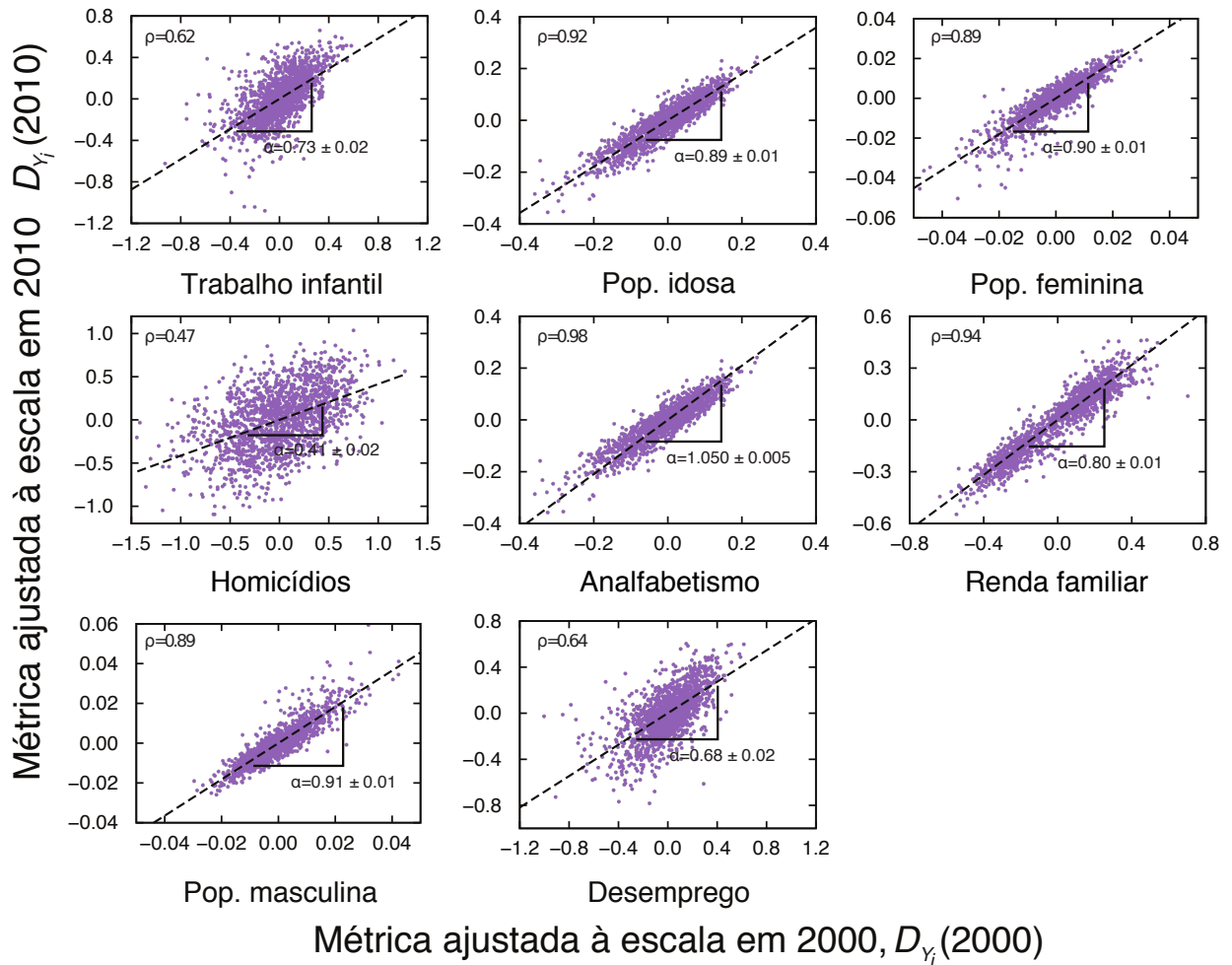


Figura 1.6: Efeitos de memória na evolução da métrica ajustada à escala $D_{Y_i}(t)$. Os pontos roxos mostram os valores de $D_{Y_i}(2010)$ versus $D_{Y_i}(2000)$ para cada cidade [veja também a Figura 1.7 para $D_{Y_i}(2000)$ versus $D_{Y_i}(1991)$]. Apesar dos diferentes graus de dispersão (ver o valor do coeficiente de correlação de Pearson ρ), observamos uma correspondência linear entre $D_{Y_i}(2010)$ e $D_{Y_i}(2000)$. As linhas tracejadas são os ajustes do modelo linear $D_{Y_i}(2010) = A_i + \alpha_i D_{Y_i}(2000)$ (Equação 1.5) obtidas via regressão por mínimos quadrados. Os valores de α_i e seus erros padrão estão apresentados nos gráficos e também resumidos na Tabela 1.2. Notamos que $\alpha_i < 1$ para quase todos os indicadores, fato que está de acordo com a aproximação às leis alométricas observadas para os valores médios de $D_{Y_i}(t)$ (Figura 1.5). Analfabetismo é a única exceção, no qual $\alpha_i \gtrsim 1$. Esse resultado também concorda com o afastamento da lei alométrica da média de $D_{Y_i}(t)$ no caso do analfabetismo.

tempo para $\alpha_i > 1$. Para $\alpha_i = 1$, temos um comportamento linear para a evolução de $D_{Y_i}(t)$. Vale a pena lembrar que A_i é um número muito pequeno ($\sim 10^{-6}$) para todos os indicadores urbanos e, portanto, os valores de $D_{Y_i}(t)$ estão na verdade se aproximando de zero para $\alpha_i < 1$ (isto é, os

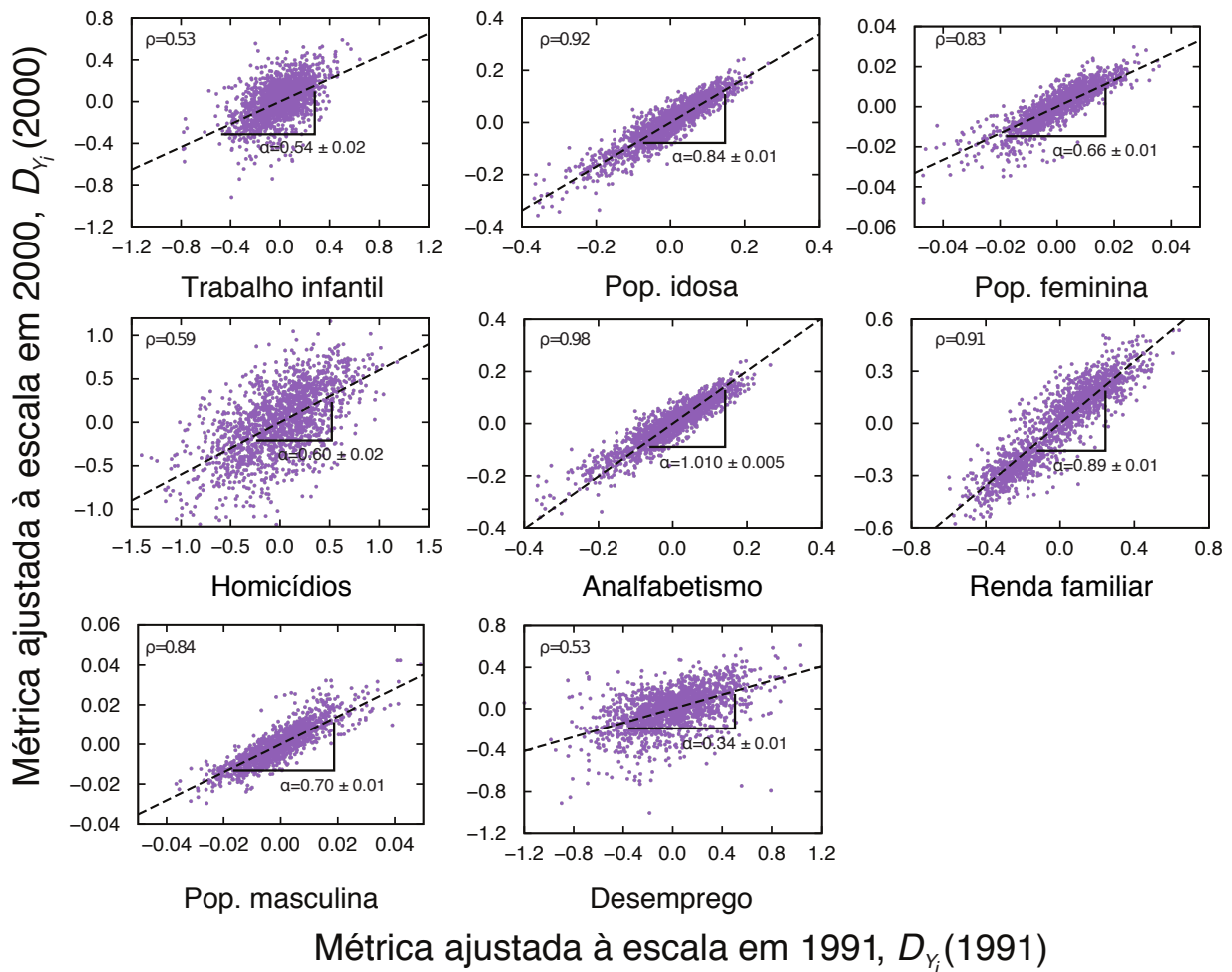


Figura 1.7: Efeitos de memória na evolução da métrica ajustada à escala $D_{Y_i}(t)$. Os pontos roxos mostram os valores de $D_{Y_i}(2000)$ versus $D_{Y_i}(1991)$ para cada cidade. As linhas tracejadas são os ajustes do modelo linear $D_{Y_i}(2000) = A_i + \alpha_i D_{Y_i}(1991)$ (Equação 1.5) obtidas via regressão de mínimos quadrados. Os valores de α_i e seus erros padrão são mostrados nas figuras e resumidos na Tabela 1.2.

valores dos indicadores urbanos estão ficando mais perto do valor esperado pela lei alométrica). Observamos ainda que $1/(1 - \alpha_i)$ desempenha o papel de um tempo característico e assumindo que $\alpha_i < 1$, quanto menor o valor de α_i , mais rapidamente esperamos que sejam as mudanças no indicador urbano. Para os indicadores relacionados à população, observamos uma tendência crescente nos valores de α_i e também que esses valores estão entre os maiores, o que concorda com a redução do crescimento da população brasileira nas últimas décadas. Para os indicadores socioeconômicos, temos, por um lado, que os valores de α_i para o trabalho infantil e o desemprego também aumentaram (apontando uma moderação de sua dinâmica); por outro lado, homicídios e renda familiar

Tabela 1.2: Coeficientes lineares α_i das relações de memória entre as métricas ajustadas à escala em anos diferentes. Valores dos parâmetros α_i obtidos por meio do ajuste de mínimos quadrados para modelo da Equação 1.5 para as relações entre as métricas ajustadas à escala: $D_{Y_i}(2000)$ versus $D_{Y_i}(1991)$ e $D_{Y_i}(2010)$ versus $D_{Y_i}(2000)$. Omitimos os valores dos parâmetros A_i , porque eles são muito pequenos ($\sim 10^{-6}$). Os valores dentro dos parênteses são os erros padrão (SE) na última casa decimal do parâmetro estimado. A última coluna mostra os valores dos coeficientes de correlação linear de Pearson ρ para cada relação.

Indicadores	Anos	α_i (SE)	ρ
Trabalho infantil	2000 – 1991	0,54 (2)	0,53
	2010 – 2000	0,73 (2)	0,62
Pop. idosa	2000 – 1991	0,84 (1)	0,92
	2010 – 2000	0,89 (1)	0,92
Pop. feminina	2000 – 1991	0,66 (1)	0,83
	2010 – 2000	0,90 (1)	0,89
Homicídios	2000 – 1991	0,60 (2)	0,59
	2010 – 2000	0,41 (2)	0,47
Analfabetismo	2000 – 1991	1,010 (5)	0,98
	2010 – 2000	1,050 (5)	0,98
Renda familiar	2000 – 1991	0,89 (1)	0,91
	2010 – 2000	0,80 (1)	0,94
Pop. masculina	2000 – 1991	0,70 (1)	0,84
	2010 – 2000	0,91 (1)	0,89
Desemprego	2000 – 1991	0,34 (1)	0,53
	2010 – 2000	0,68 (2)	0,64

tiveram seus valores de α_i reduzidos, sugerindo que suas taxas de variação aumentaram. Além da evolução nos valores de α_i maiores, indicadores relacionados à população apresentam (em geral) valores de α_i do que os observados para os indicadores socioeconômicos, apontando que esses últimos são mais suscetíveis às mudanças naturais ou de políticas públicas. O analfabetismo tem um comportamento único porque o seu valor de α_i , que estava muito perto de um [$D_{Y_i}(2000)$ versus $D_{Y_i}(1991)$], aumentou, resultado que concorda com o distanciamento da alometria relatado na Figura 1.5 e sugere uma aceleração nesse processo.

Além dos $D_{Y_i}(t + \Delta t)$ serem autocorrelacionados com seus valores em tempos anteriores, eles também exibem correlações cruzadas significativas uns com os outros (Figura 1.8). Esses efeitos de memória e o fato de que os resíduos em torno das relações $D_{Y_i}(t + \Delta t)$ versus $D_{Y_i}(t)$ estão muito perto de distribuições gaussianas (Figura 1.9) com desvios padrão praticamente constantes, (Figura 1.10) fazem essas métricas ajustadas à escala particularmente boas para serem usadas em modelos de regressão linear visando previsões. Assim, ajustamos o seguinte modelo linear genera-

lizado (via método dos mínimos quadrados)

$$D_{Y_i}(t + \Delta t) = C_0 + \sum_{k=1}^8 C_k D_{Y_k}(t) + \eta_i(t) \quad (1.8)$$

considerando as relações $(t + \Delta t) = 2000$ versus $t = 1991$ e $(t + \Delta t) = 2010$ versus $t = 2000$. Na Equação 1.8, C_k é o coeficiente linear que quantifica o poder preditivo de $D_{Y_k}(t)$ em $D_{Y_i}(t + \Delta t)$ (C_0 é o coeficiente de interceptação) e $\eta_i(t)$ é o termo de ruído que representa o efeito de fatores não mensuráveis. Notamos que esses modelos simples descreverem de 31% a 97% da variância observada em $D_{Y_i}(t + \Delta t)$ e reproduzem corretamente os valores médios da métrica ajustada à escala acima e abaixo das leis alométricas para os anos de 2000 e 2010, usando apenas os dados dos anos de 1991 e 2000, respectivamente (Figura 1.11). Comparamos ainda as distribuições dos valores empíricos de D_{Y_i} com as previsões desses modelos lineares e observamos que o acordo é bom para os indicadores de população idosa, feminina e masculina, bem como para o analfabetismo e renda (Figuras 1.12 e 1.13). Motivados por esses bons acordos, propusemos prever os valores de $D_{Y_i}(t + \Delta t)$ no ano de 2020 (próximo censo nacional brasileiro). A fim de fazer isso, consideramos que os coeficientes lineares C_k são constantes ao longo do tempo e empregamos o valor médio de C_k nas duas combinações de anos utilizados na Equação 1.8 para prever os valores de $D_{Y_i}(t + \Delta t)$. Vale a pena notar que, assumindo C_k constante, estamos ignorando a evolução dos fatores socioeconômicos e políticos. Numa situação ideal, poderíamos investigar a evolução dos valores de C_k para conseguir previsões mais confiáveis. No entanto, nossos dados (isto é, os dois valores para C_k) não nos permitem investigar possíveis comportamentos evolutivos nos valores de C_k . Mesmo assim, como apontado por Bettencourt [47], a dinâmica dos indicadores urbanos parece ser dominada por prazos longos (~ 30 anos) e, portanto, a abordagem de coeficientes constantes deve ser vista como uma primeira aproximação. As barras cinzas na Figura 1.5 mostram as médias $D_{Y_i}(2020)$ após agrupar as cidades em que $D_{Y_i}(1991) > 0$ e $D_{Y_i}(1991) < 0$. Observamos que as previsões para os valores médios, basicamente, mantêm as tendências apresentadas nos anos anteriores. Para o desemprego, em que a tendência não foi muito clara, as previsões colocam esse indicador junto com a maioria dos indicadores, em que a média de $D_{Y_i}(t)$ tem diminuído para cidades inicialmente acima da lei alométrica e aumentando para aqueles inicialmente abaixo.

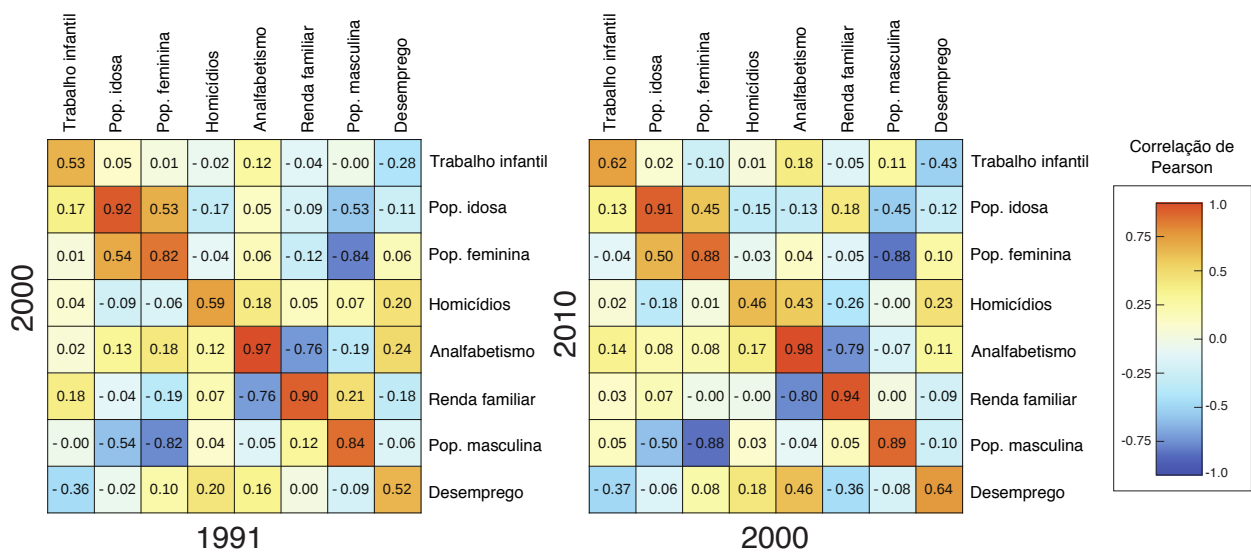


Figura 1.8: Correlação cruzada entre os indicadores urbanos. A matriz na esquerda mostra os valores do coeficiente de correlação de Pearson entre a métrica ajustada à escala $D_{Y_i}(t)$ para um dado indicador (um indicador por linha) no ano $t = 2000$ e todos os outros indicadores no ano de $t = 1991$ (um indicador por coluna). A matriz da direita representa o mesmo para os anos $t = 2010$ e $t = 2000$. O valor dentro de cada célula é a correlação de Pearson e cada uma é colorida de acordo com esse valor. Notamos que todos os indicadores estão fortemente correlacionados aos seus próprios valores no passado. Além disso, todos os indicadores apresentam correlações relevantes com pelo menos um outro indicador.

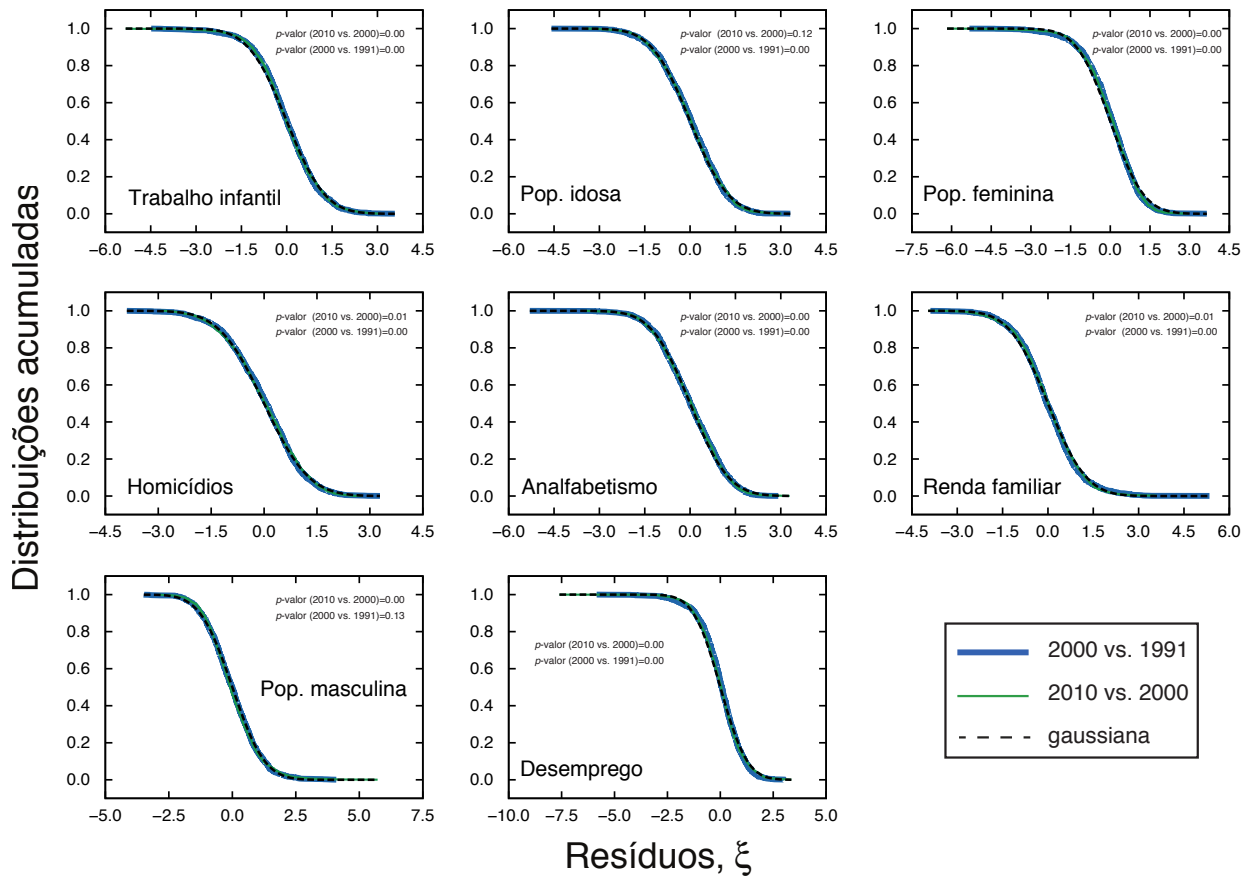


Figura 1.9: Distribuições acumuladas das flutuações normalizadas em torno das relações entre as métricas ajustadas à escala $D_{Y_i}(t + \Delta t)$ e $D_{Y_i}(t)$. Os gráficos mostram as distribuições acumuladas dos resíduos normalizados ξ da regressão linear entre $D_{Y_i}(t + \Delta t)$ e $D_{Y_i}(t)$ (Figuras 1.6 e 1.7) para os anos 2000-1991 (linhas azuis) e 2010-2000 (linhas verdes) em comparação com a gaussiana de média nula e variância unitária (linhas tracejadas). Também mostramos os p -valores do método Cramér von Mises que testa a hipótese nula de que os resíduos ξ são distribuídos normalmente. Observamos que a normalidade dos dados é rejeitada na maioria dos casos (provavelmente, devido à heterocedasticidade presente nessas relações — Figura 1.10). No entanto, não são observadas grandes diferenças entre a curva gaussiana acumulada e as distribuições empíricas acumuladas, o que sugere que ξ pode ser aproximadamente descrito por uma distribuição gaussiana.

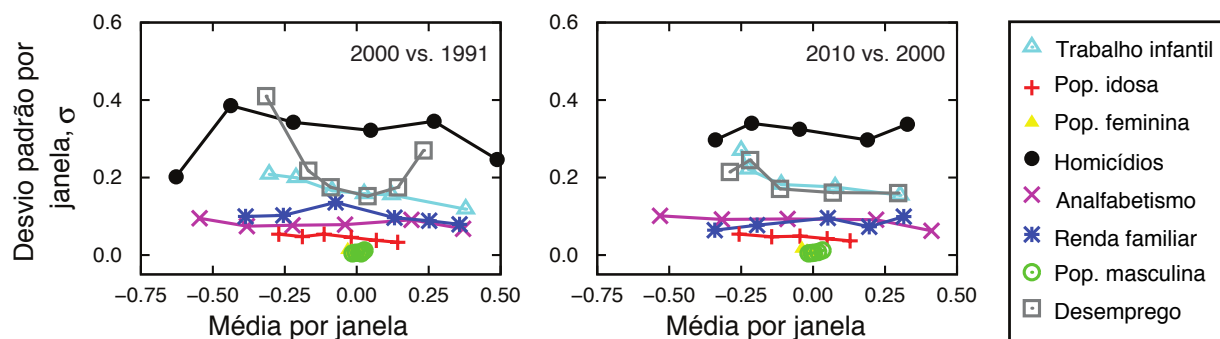


Figura 1.10: Desvio padrão σ calculado em janelas para a relação entre as métricas ajustadas à escala $D_{Y_i}(t + \Delta t)$ e $D_{Y_i}(t)$. Esses gráficos mostram o desvio padrão σ da métrica ajustada à escala $D_{Y_i}(t + \Delta t)$ versus a média de $D_{Y_i}(t)$ calculada em cinco janelas igualmente espaçadas a partir da relação entre $D_{Y_i}(t + \Delta t)$ e $D_{Y_i}(t)$ (Figuras 1.6 e 1.7) para os anos 2000-1991 (painel a esquerda) e 2010-2000 (painel a direita). Notamos que o desvio padrão pode ser aproximado por uma constante para a maioria dos indicadores em ambas as combinações de anos. Observamos ainda que as pequenas flutuações em σ são, provavelmente, o motivo do teste de Cramér von Mises rejeitar a normalidade das flutuações ξ mostradas na Figura 1.9. Ao ajustar os modelos lineares da Equação 1.8, também levamos em conta essa pequena heterocedasticidade (modelo implementado no Stata 13 — <http://www.stata.com> — via a opção `robust` da função `regress`), mas os coeficientes lineares permanecem praticamente os mesmos.

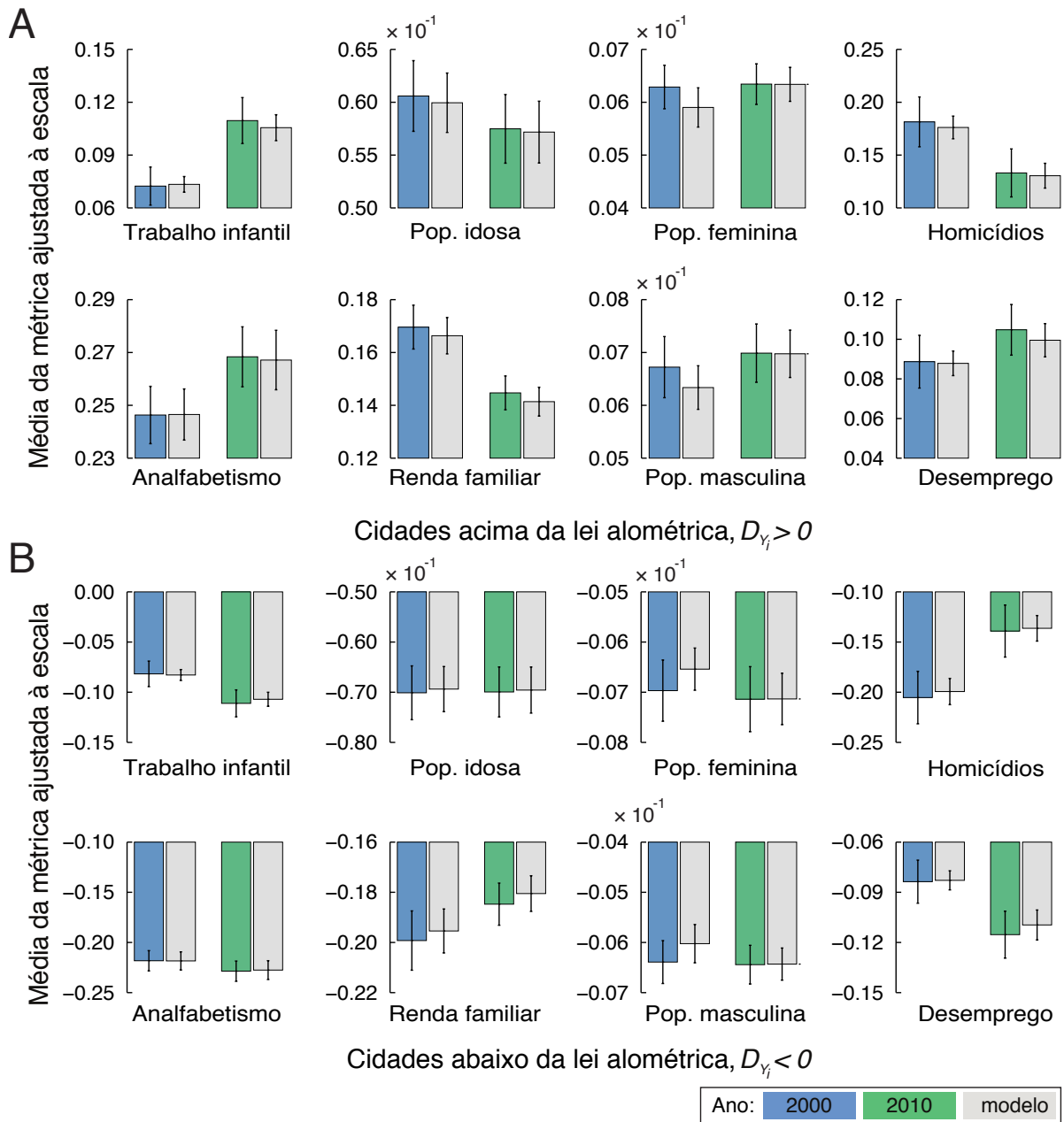


Figura 1.11: Comparações entre os valores médios das métricas ajustadas à escala obtidas a partir dos modelos lineares e os empíricos. Aplicamos o modelo linear da Equação 1.8 para prever os valores de $D_{Y_i}(t)$ no ano de 2000 usando apenas os dados do ano de 1991 e também para prever $D_{Y_i}(t)$ no ano 2010 apenas usando os dados de 2000. Em ambos os casos, calculamos a média de $D_{Y_i}(t)$ para as previsões (barras em cinza) após agrupar as cidades em acima A) e abaixo B) das leis alométricas (nesse ano) e comparamos esses resultados com as mesmas médias utilizando os dados empíricos (barras azuis para o ano de 2000 e barras verdes para o ano de 2010). As barras de erros são intervalos de confiança de 95% calculados via *bootstrapping* para os valores médios. Observamos que os valores médios previstos estão em boa concordância com os valores empíricos para todos os indicadores urbanos em ambos os anos.

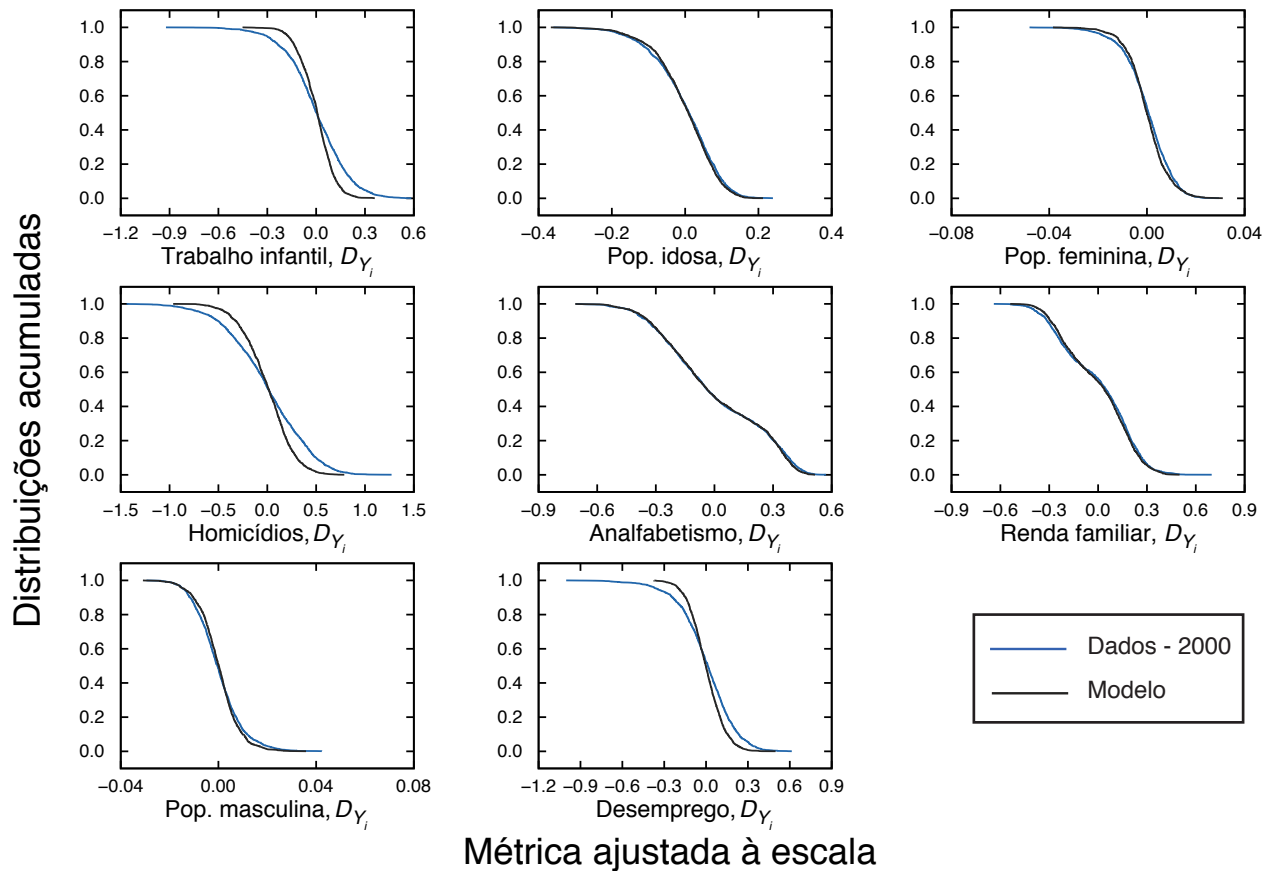


Figura 1.12: Comparações entre as distribuições das métricas ajustadas à escala obtidas a partir dos modelos lineares e os empíricos. Obtivemos os valores de $D_{Y_i}(t)$ no ano de 2000 usando o modelo linear da Equação 1.8 e empregando os dados do ano de 1991. Calculamos as distribuições acumuladas de $D_{Y_i}(t)$ para os valores previstos (linhas pretas) e os empíricos (linhas azuis). Observamos que o acordo é muito bom para os indicadores de população, analfabetismo e renda familiar. Para o restante dos indicadores, observamos que o modelo falha em reproduzir as caudas das distribuições.

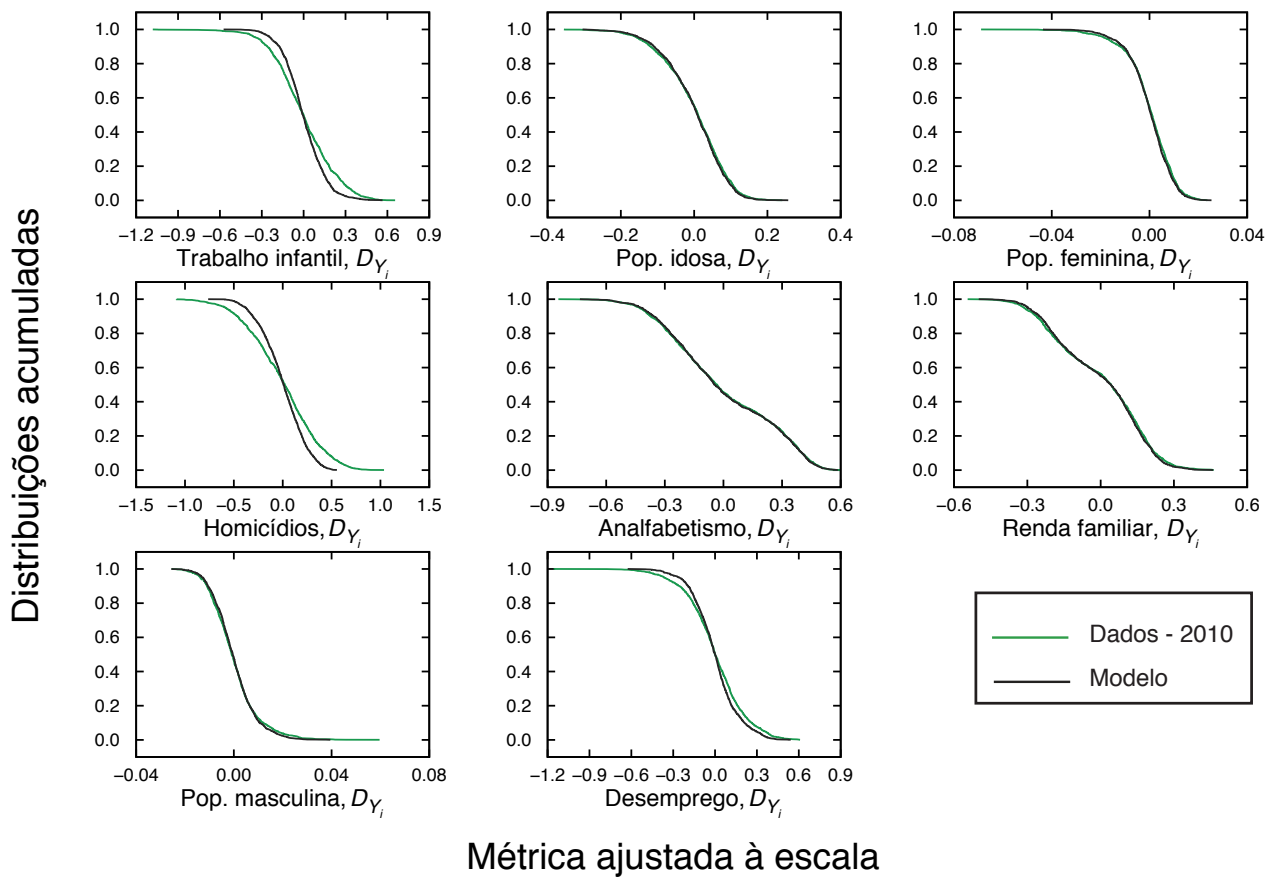


Figura 1.13: Comparações entre as distribuições das métricas ajustadas à escala obtidas a partir dos modelos lineares e os empíricos. O mesmo que a Figura 1.12 considerando o ano 2010.

A fim de obter mais informações sobre as previsões, construímos uma visualização geográfica das mudanças esperadas em $D_{Y_i}(t)$ entre os anos de $t = 2010$ e $t = 2020$. Os círculos sobre os mapas na Figura 1.14 mostram a localização geográfica de cidades brasileiras; os raios desses círculos são proporcionais à $|D_{Y_i}(2020) - D_{Y_i}(2010)|$ e são coloridos com tons de azul para as cidades onde $[D_{Y_i}(2020) - D_{Y_i}(2010)] < 0$ (o indicador deverá diminuir) e com tons de vermelho para as cidades onde $[D_{Y_i}(2020) - D_{Y_i}(2010)] > 0$ (o indicador deverá aumentar). Em ambos os casos, quanto mais escuro o tom da cor, maior é o valor absoluto da diferença $[D_{Y_i}(2020) - D_{Y_i}(2010)]$. Talvez, a característica mais marcante dessas visualizações seja o fato de que as mudanças previstas aparecem espacialmente agrupadas para quase todos os indicadores, o que de alguma forma reflete as desigualdades geográficas existentes no Brasil; no entanto, alguns padrões intrigantes são dependentes do indicador.

Para o trabalho infantil, $D_{Y_i}(t)$ deverá aumentar em torno das três das regiões mais povoadas que contêm as áreas metropolitanas de São Paulo, Rio de Janeiro e as áreas metropolitanas de quase todas as capitais do nordeste; observamos que a diminuição dos casos de trabalho infantil é esperada, em sua maioria, nas cidades do interior e do sul. Para os indicadores de população idosa, feminina e masculina a aglomeração das mudanças em $D_{Y_i}(t)$ é bastante evidente: população idosa e feminina devem diminuir na maior parte das cidades do nordeste e exibem uma tendência crescente em grande parte das outras regiões (população masculina comporta-se antissimetricamente comparada a população feminina). No caso dos homicídios, o nosso modelo prevê uma diminuição de $D_{Y_i}(t)$ para a grande maioria das cidades do sul e observamos uma faixa perto da costa leste, onde $D_{Y_i}(t)$ deverá diminuir na maioria das cidades (ambas as áreas densamente povoadas). Por outro lado, o modelo prevê que cidades do interior (especialmente cidades do interior do estado de São Paulo e da região nordeste) aumentem $D_{Y_i}(t)$, sugerindo que esse crime violento pode estar se deslocando para áreas menos povoadas do interior do Brasil.

Para o analfabetismo, mais uma vez, o agrupamento das mudanças é facilmente perceptível: esperamos um aumento de $D_{Y_i}(t)$ para a maior parte do nordeste e de cidades mais ao norte. Já, para a maioria das cidades de outras regiões (excluindo várias cidades do interior da região mais ao sul), o valor de $D_{Y_i}(t)$ deverá diminuir. Para renda familiar, observamos também um agrupamento nas mudanças, no qual esperamos que a maioria das cidades do norte aumentem o valor de $D_{Y_i}(t)$ (especialmente as cidades do interior dessas regiões), enquanto que para a maioria

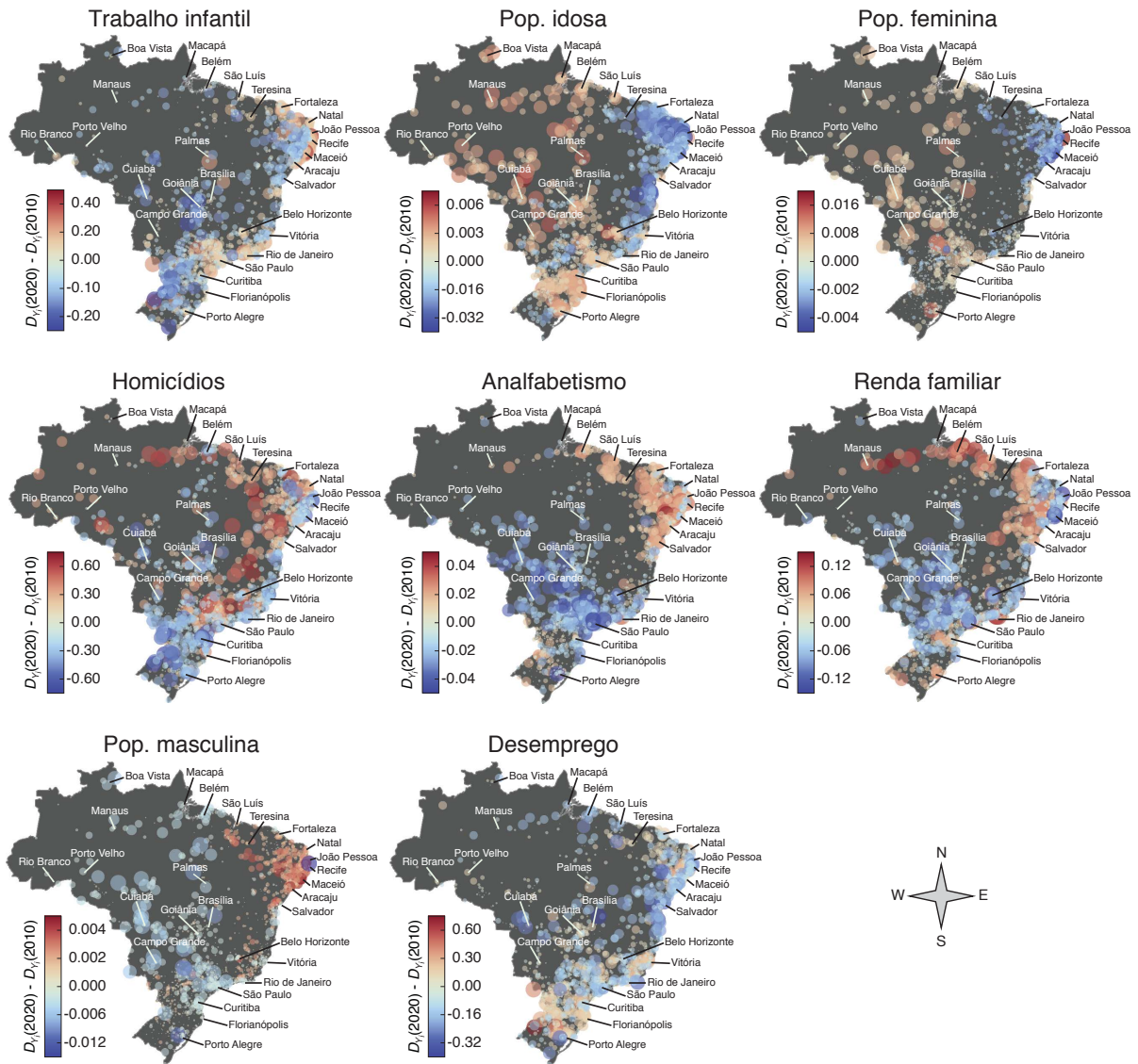


Figura 1.14: Visualização geográfica das mudanças previstas nas métricas ajustadas à escala $D_{Y_i}(t)$ entre os anos de $t = 2010$ e $t = 2020$. Cada círculo representa uma cidade e o raio do círculo é proporcional à $|D_{Y_i}(2020) - D_{Y_i}(2010)|$. Os círculos são coloridos de acordo com a diferença entre $D_{Y_i}(2020)$ e $D_{Y_i}(2010)$: os tons de azul indicam que esperamos uma diminuição nos valores de $D_{Y_i}(2020)$, enquanto os tons de vermelho mostram as cidades onde esperamos um aumento nos valores de $D_{Y_i}(2020)$. As cidades marcadas são as capitais das vinte e sete unidades federativas do Brasil (os estados brasileiros e o Distrito Federal). As previsões para os valores de $D_{Y_i}(2020)$ foram obtidas usando o modelo linear da Equação 1.8, em que os coeficientes lineares C_k são os valores médios das duas combinações de anos 2000–1991 e 2010–2000. Notamos que as mudanças em $D_{Y_i}(t)$ aparecem especialmente agrupadas para a maioria dos indicadores, com regiões onde se espera que a maioria das cidades aumentem ou diminuam o valor da métrica ajustada à escala.

das cidades na parte central do Brasil o valor de $D_{Y_i}(t)$ deverá diminuir. Essas mudanças esperadas podem ser (pelo menos em parte) relacionadas ao programa “bolsa família” — um programa de bem-estar social em grande escala do governo brasileiro (mais de 14 milhões de famílias foram beneficiadas em 2013) para fornecer ajuda financeira às famílias pobres por meio de transferência direta de renda — visto que grande parte das famílias nesse programa são das regiões norte e nordeste. Vale a pena notar que, para a participação no programa, as famílias devem assegurar que os seus filhos frequentem a escola e, portanto, seria natural de se esperar uma redução $D_{Y_i}(t)$ para o analfabetismo nas mesmas regiões em que se concentram os beneficiários do programa, o que não foi previsto pelo modelo. Esse resultado sugere que simplesmente reforçar a frequência escolar pode não ser eficiente para reduzir o analfabetismo, o que também pode ser explicado por más condições observadas na rede pública de ensino dessas regiões. Finalmente, para o desemprego não há um agrupamento claro das alterações em $D_{Y_i}(t)$. Em vez disso, esperamos uma diminuição de $D_{Y_i}(t)$ difundida em todo o território brasileiro (exceto no extremo sul do país, onde podemos constatar a prevalência de tons leves de vermelho).

1.4 Conclusões e perspectivas

Nesse capítulo, argumentamos que apesar de serem amplamente utilizados em relatórios do governo e também em trabalhos acadêmicos, indicadores *per capita* ignoraram completamente as leis alométricas universais que emergem na dinâmica de crescimento de cidades. Discutimos que os indicadores *per capita* podem apresentar tendências para pequenas ou grandes cidades, dependendo se a alometria é sublinear ou superlinear. Assim, empregamos uma métrica ajustada à escala avaliando a diferença entre os valores reais dos indicadores urbanos e os que são esperados pelas relações alométricas. Ao investigar a evolução das métricas ajustadas à escala, relatamos padrões que não aparecem nos indicadores *per capita*. As métricas ajustadas à escala também exibem uma correspondência linear com seus valores em tempos anteriores, uma característica que facilita o uso de regressão linear para modelar os indicadores urbanos. Com o emprego de modelos lineares simples para descrever as métricas ajustadas à escala, baseados apenas em seus valores anteriores, verificamos que esses modelos representam de 31% a 97% da variância observada e que eles reproduzem corretamente os valores médios das métricas ajustadas à escala. Assumindo os coeficientes lineares constantes ao longo do tempo, apresentamos previsões para os valores das métricas ajustadas à escala no ano de 2020, quando o próximo censo brasileiro deverá acontecer. Observamos que as mudanças previstas para os indicadores urbanos aparecem (na maioria dos casos) espacialmente agrupadas, ou seja, formando regiões onde se espera que a maioria das cidades aumentem ou diminuam o valor da métrica ajustada à escala. Acreditamos que a nossa análise pode encontrar aplicações potenciais no desenvolvimento e atribuição de novas políticas e recursos no contexto do planejamento urbano. Finalmente, vale mencionar que os métodos empregados aqui também podem ser diretamente aplicados em outros contextos relacionados às alometrias, como para índices econômicos e quantidades biológicas.

Correlações espaciais, aglomeração e transições tipo percolação em crimes de homicídio

A dinâmica espacial de atividades criminosas tem sido estudada via métodos de física estatística. No entanto, modelos e resultados costumam ser restritos a escalas locais (escala de cidades) e muito menos é conhecido sobre esses padrões em escalas maiores, como a nível de um país. Nesse capítulo, relatamos uma caracterização da dinâmica espacial dos crimes de homicídio ao longo do território brasileiro usando dados de todas as cidades (~ 5000) em um período de mais de trinta anos. Nossos resultados mostram que a função de correlação espacial nos homicídios *per capita* decai exponencialmente com a distância entre as cidades e que o comprimento de correlação característica apresenta uma tendência crescente notável nos últimos anos. Também investigamos a formação de agrupamentos/*clusters* espaciais de cidades por meio de uma análise de percolação, na qual observamos agrupamentos de cidades e comportamentos similares à transição de fase ao descrever o tamanho do maior *cluster* como função do limiar de homicídios. Esse comportamento semelhante a uma transição de fase apresenta características evolutivas caracterizadas por um aumento do limiar de homicídios (onde as transições ocorrem) e por uma diminuição nas magnitudes de transição (comprimento dos saltos no tamanho do *cluster*).

2.1 Introdução e apresentação do problema

O estudo dos fenômenos sociais é atualmente onipresente na agenda de pesquisa do físico. Em particular, métodos baseados na Física Estatística mostram-se muito poderosos quando aplicados a sistemas sociais [6, 9, 90, 91]. Um exemplo importante está relacionado às investigações e modelagens empíricas de atividades criminosas [14, 41–44, 46–48], que foram recentemente revistos por D’Orsogna e Perc [92]. Em uma escala local (nível municipal), um padrão bem conhecido está relacionado à chamada “teoria das janelas quebradas”. Proposta por Wilson e Kelling [93] em 1982 e amplamente aceita entre criminologistas, essa teoria pode ser resumida pela ideia de que ambientes urbanos degradados fomentam atividades criminosas em suas vizinhanças. De fato, resultados empíricos sugerem que criminosos tendem a voltar para locais anteriormente visitados [94] e a violação de normas/regras sociais faz com que a desordem se propague [95]. Esses comportamentos foram modelados por equações de reação–difusão [96–98] e por processos pontuais auto-excitante — em inglês, *self-exciting point processes* [99–101]. Notavelmente, métodos estatísticos baseados nesses modelos foram efetivamente utilizados para a identificação de possíveis áreas de criminalidade em vários departamentos de administração de segurança dos Estados Unidos da América [92, 95, 102].

Em escalas maiores, como por exemplo, ao pensarmos em países, conhecemos muito menos sobre a dinâmica espacial de atividades criminosas. Perguntas sobre se as atividades criminosas em uma determinada cidade têm influência nas cidades vizinhas ou se as atividades criminosas também formam aglomerados/*clusters* espaciais de cidades ainda não foram abordadas. Nesse contexto, o Brasil é (infelizmente) um lugar ideal para abordar essas questões: é um dos países mais violentos do mundo e é também um país com dimensões continentais. Assim, investigamos as correlações espaciais e os padrões de agrupamento de homicídios *per capita* em todas as cidades brasileiras; quantificamos ainda como esses dois aspectos da dinâmica espacial de homicídios têm evoluído ao longo do período de 1980 a 2011.

Nossos resultados mostram que a função de correlação espacial dos homicídios *per capita* decai exponencialmente com a distância entre as cidades. Observamos, ainda, que o comprimento de correlação característico apresenta uma tendência crescente notável nos últimos anos (≈ 14 km por ano). Esses comportamentos não estão diretamente relacionados às correlações espaciais do

tamanho da população, que apresentam decaimento tipo lei de potência com um expoente bastante estável. Os resultados revelam a emergência de agrupamentos/*clusters* espaciais e que o tamanho do maior aglomerado depende do limiar empregado de forma não trivial. De modo semelhante ao que acontece em transições de fase, o tamanho do maior grupo muda abruptamente em torno de um valor específico do limiar de homicídios. Esse comportamento de transição também evoluiu: o limiar de homicídios em que as transições ocorrem aumentou, enquanto a magnitude das transições (comprimento dos saltos no tamanho do *cluster*) tem diminuído ao longo dos anos. Assim, os resultados apóiam a ideia de que o número de homicídios de uma cidade afeta o valor dessa métrica urbana em cidades vizinhas (a dezenas de quilômetros de distância) e que esse comportamento correlacionado se intensificou nos últimos anos.

2.2 Apresentação dos dados

Acessamos dados de todas as cidades brasileiras no período de 1980 a 2011 (32 anos), disponibilizados gratuitamente pelo Departamento de Informática do Sistema Único de Saúde [77]. Para cada cidade i e em um determinado ano t , temos o número de homicídios $H_i(t)$, o tamanho da população $N_i(t)$ e a localização geográfica da cidade (latitude e longitude). A Figura 2.1 ilustra o conjunto de dados, mostrando a evolução do homicídio *per capita* no Brasil $h(t) = [\sum H_i(t)]/[\sum N_i(t)]$ e a distribuição espacial do homicídio *per capita* nas cidades $h_i(t) = H_i(t)/N_i(t)$, sobre o território brasileiro. Apesar de mostrarmos anteriormente que indicadores urbanos *per capita* possuem uma componente extra relacionada à população (o que poderia distorcer o nossos resultados), os resultados sobre homicídios *per capita* nesse contexto não são afetados pela população. Quando escalamos o número de homicídios por N_i^β , os resultados sobre percolação e formação de *clusters* permanecem inalterados. Por simplicidade, mostramos os resultados apenas para métrica *per capita*. Notamos que o homicídio *per capita*, no Brasil, mais que duplicou durante o período de 32 anos cobertos por nossos dados. Os mapas dessa figura também fornecem indícios de correlações espaciais e agrupamentos na distribuição espacial de $h_i(t)$; especificamente, observamos o surgimento de áreas (grupo de cidades) com valores pequenos e grandes de $h_i(t)$. Um resultado que se assemelha aos resultados do capítulo anterior para o agrupamento espacial da evolução da métrica ajustada à escala (Figura 1.14).

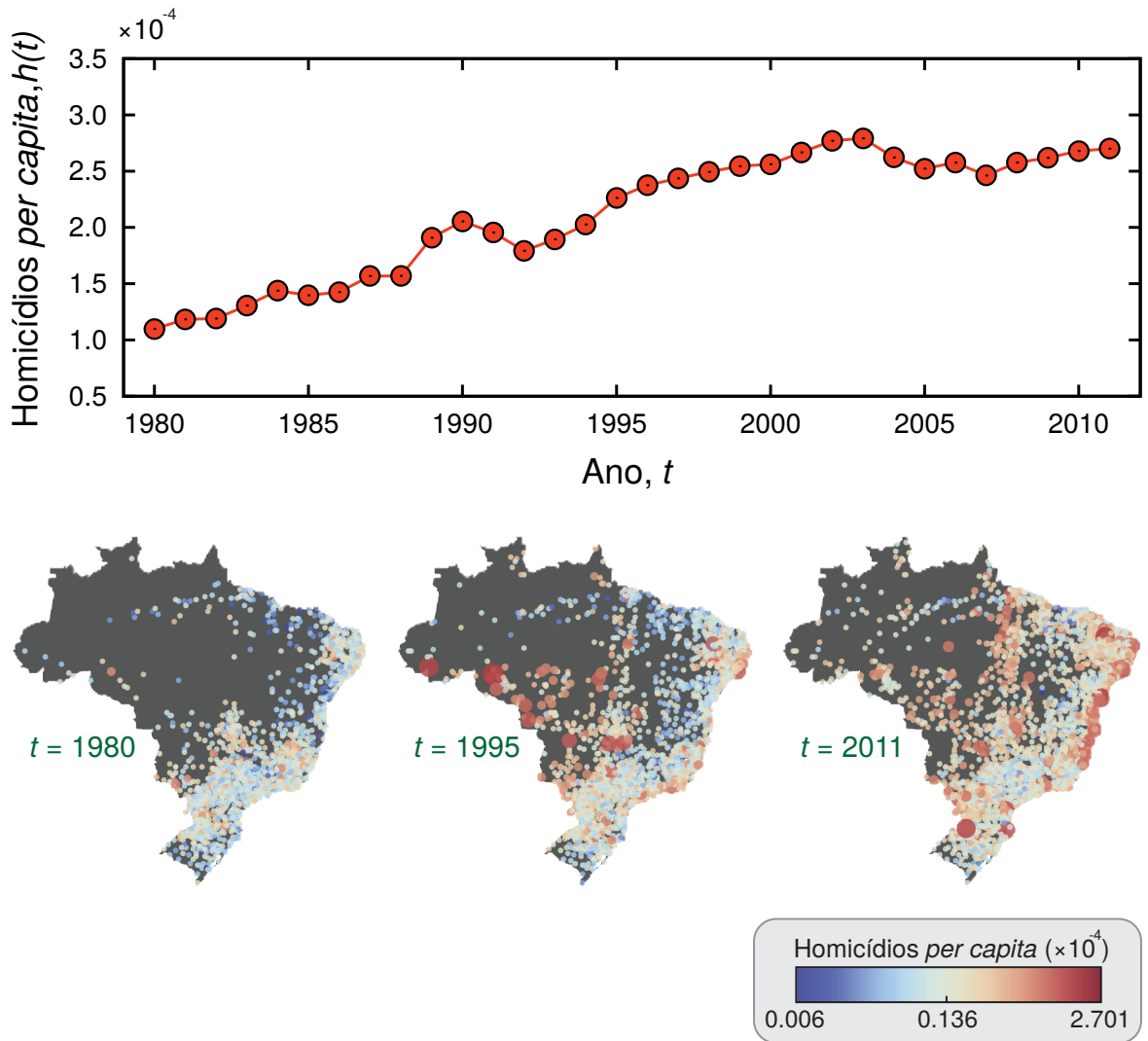


Figura 2.1: Evolução temporal do homicídio *per capita* no Brasil e sua distribuição espacial entre as cidades brasileiras. O gráfico mostra a evolução do total *per capita* de homicídios $h(t)$ no Brasil durante o período de 1980 a 2011 e os três mapas ilustram a distribuição espacial de $h_i(t)$ sobre o território brasileiro nos anos de 1980, 1995 e 2011. Nos mapas, cada ponto representa uma cidade e o código de cores representa o valor de $h_i(t)$. O tamanho do ponto também é proporcional $h_i(t)$.

2.3 Resultados e discussão

Para começar a quantificar a dinâmica espacial em $h_i(t)$, calculamos a função de correlação espacial

$$C_h(r,t) = \frac{\langle [h_i(t) - \mu(r,t)][h_j(t) - \mu(r,t)] \rangle_{r_{ij}=r}}{\sigma^2(r,t)}, \quad (2.1)$$

em que $h_i(t)$ e $h_j(t)$ são os homicídios *per capita* nas cidades i e j , $\mu(r, t)$ é a média, $\sigma^2(r, t)$ é a variância dos homicídios *per capita* de cidades separadas por r quilômetros e $\langle \dots \rangle_{r_{ij}=r}$ representa a média sobre todos os pares de cidades separadas por r quilômetros. Consideramos efetivamente intervalos de r para calcular $C_h(r, t)$ devido à discretização natural de nossos dados. A Figura 2.2A mostra a dependência espacial da função de correlação $C_h(r, t)$ para o ano de $t = 2005$ em escala mono-log. Observamos que a função $C_h(r, t)$ é bem aproximada por um decaimento exponencial abrangendo centenas de quilômetros, ou seja,

$$C_h(r, t) \sim \exp[-r/r_c(t)] \quad (2.2)$$

em que $r_c(t)$ é o comprimento de correlação característico no ano t . Similarmente, observamos decaimentos exponenciais para todos os anos em nosso conjunto de dados (Figura 2.3). No entanto, o valor de $r_c(t)$ varia de ano para ano, como representado na Figura 2.2B. Entre 1980 e 2003, $r_c(t)$ mostra uma tendência decrescente aproximadamente linear de 1,8 km por ano, seguida de uma tendência linear crescente de 14 km por ano no período de 2003 a 2011. Portanto, os resultados indicam que o valor do homicídio *per capita* em uma cidade tem influência nos valores de suas cidades próximas, por meio de correlações de curto alcance. No entanto, é interessante notar que os valores de $r_c(t)$ são comparáveis com as distâncias típicas entre as cidades brasileiras (valor médio de ~ 1000 km). Além disso, a tendência notável de aumento de $r_c(t)$ ao longo dos últimos anos, sugere que a influência/correlação que cidades exercem umas sobre as outras tem sido intensificada. É muito difícil apontar origens concretas para explicar essa variação acentuada nos valores de $r_c(t)$ e, talvez, a investigação de outros indicadores urbanos poderia ajudar a esclarecer esse comportamento. Casos de homicídios são correlacionados com vários outros indicadores urbanos [14, 44]. Além disso, como vimos no capítulo anterior (Figura 1.8), o número de homicídios também é positivamente correlacionado no tempo com indicadores de analfabetismo e desemprego. Visto que analfabetismo e desemprego também passaram por um período de variação rápida ao longo dos últimos anos, esses indicadores também podem ter contribuído para as mudanças na dinâmica espacial de $h_i(t)$.

Além de serem correlacionados com outros indicadores urbanos, as características evolutivas de $C_h(r, t)$ também podem estar relacionadas à dinâmica espacial do tamanho da população $N_i(t)$,

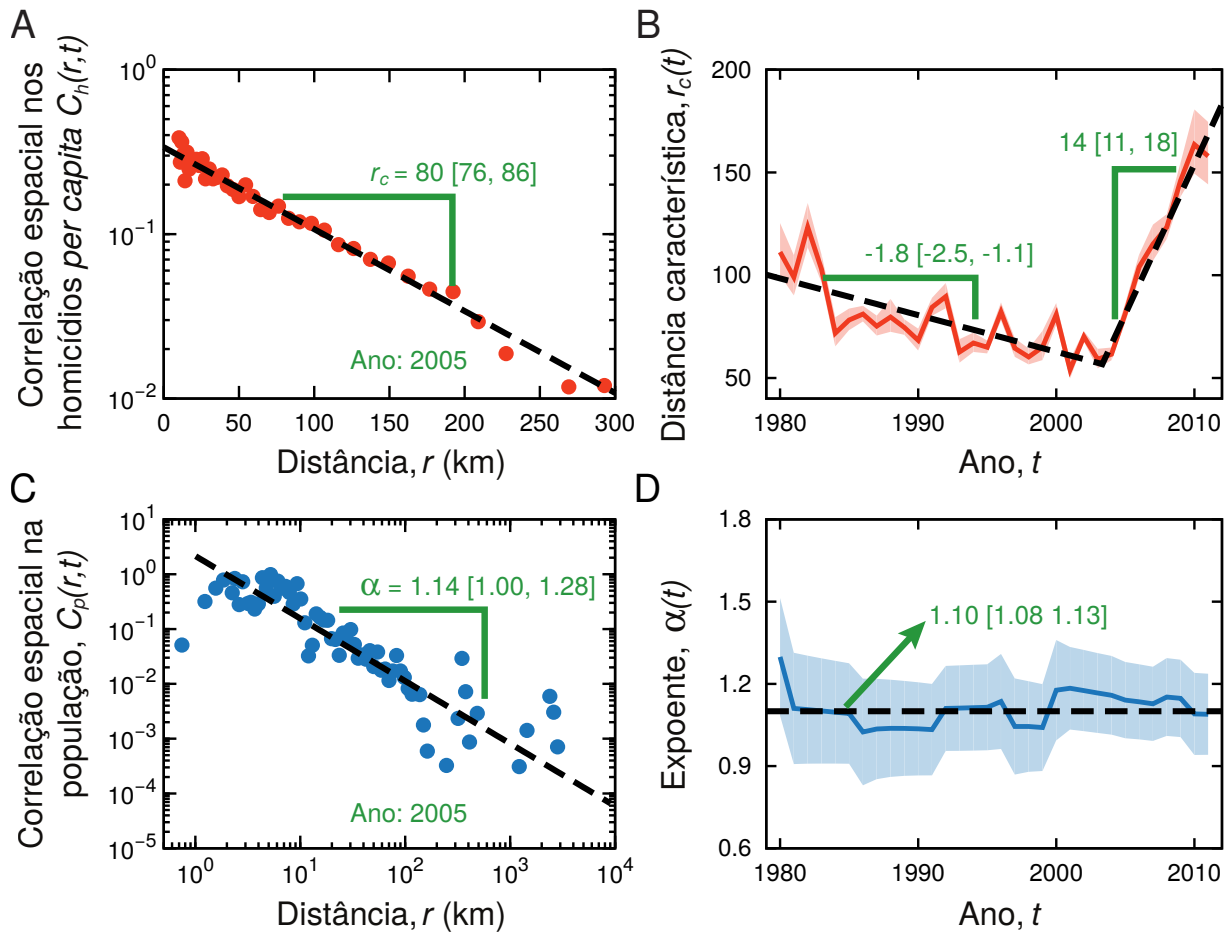


Figura 2.2: Correlações espaciais nos homicídios *per capita* e no tamanho da população. A) Dependência espacial da função de correlação $C_h(r, t)$ calculada para os homicídios *per capita* $h_i(t)$ no ano de $t = 2005$. Os pontos vermelhos são os valores empíricos de $C_h(r, t)$ e a linha tracejada é um ajuste exponencial (via método dos mínimos quadrados) aos dados em escala mono-log. O valor do comprimento característico de correlação $r_c(t)$ e seus limites de confiança de 95% são mostrados no gráfico. B) A linha vermelha mostra a evolução de $r_c(t)$ ao longo dos anos em nosso conjunto de dados. As áreas sombreadas são os intervalos de confiança de 95% para $r_c(t)$ e as duas linhas tracejadas são ajustes lineares para os valores $r_c(t)$ *versus* t , nos períodos de 1980-2003 e 2003-2011 (os coeficientes lineares e seus limites de confiança são mostrados no gráfico). C) Dependência espacial da função de correlação $C_p(r, t)$ calculada para o tamanho da população $N_i(t)$ no ano de $t = 2005$. Os pontos azuis são os valores empíricos e a linha tracejada é um ajuste lei de potência para os dados em escala log-log (o expoente da lei de potência α é mostrado no gráfico). D) A linha azul ilustra o comportamento aproximadamente constante de $\alpha(t)$ ao longo dos anos (áreas sombreadas representam intervalos de confiança de 95%), caracterizado por um valor médio de 1,10.

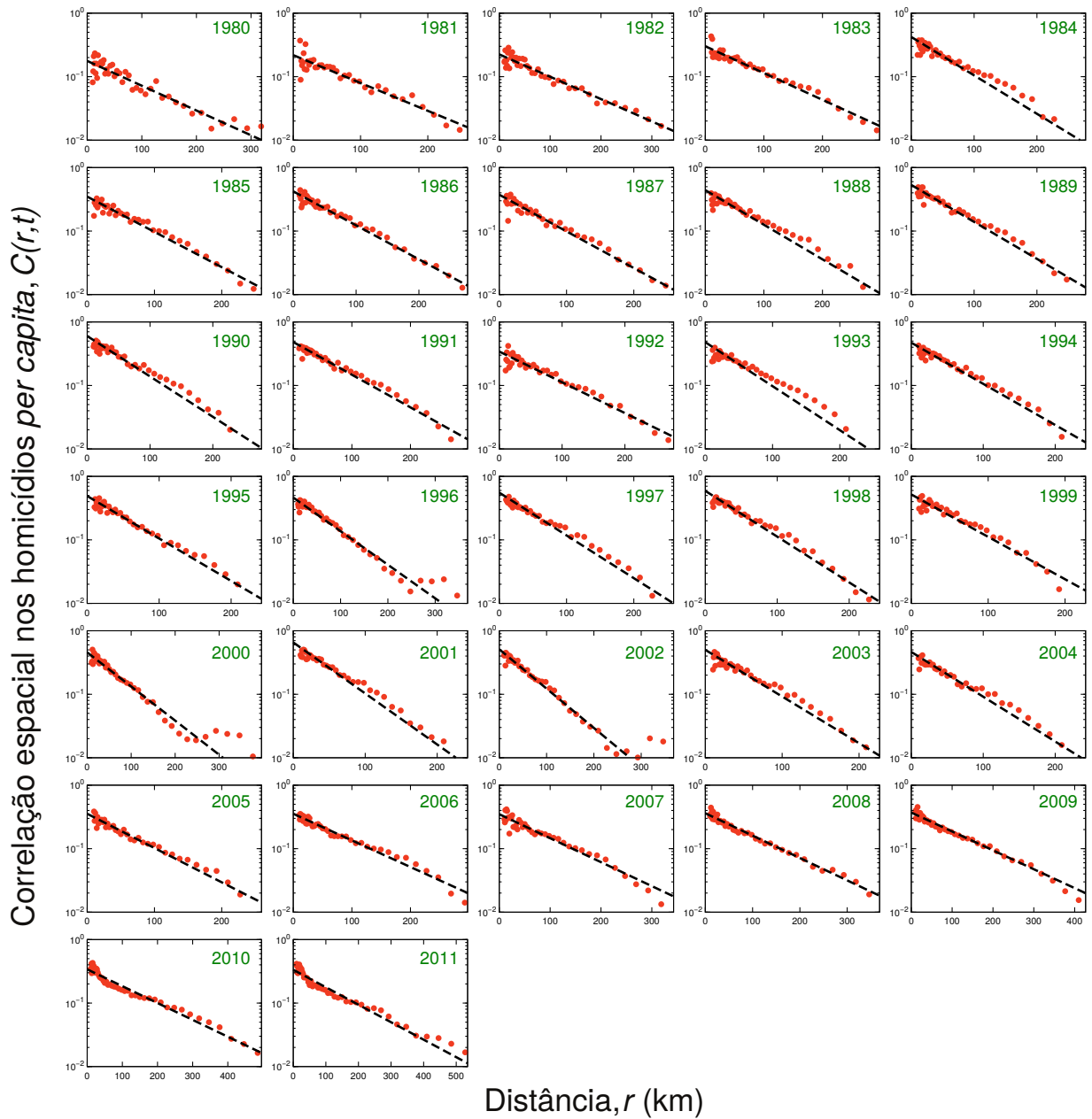


Figura 2.3: Correlação espacial nos homicídios *per capita*. Cada gráfico mostra a dependência espacial da função de correlação $C_h(r, t)$ calculada para o homicídio *per capita* $h_i(t)$ em um dado ano t (como indicado nos gráficos). Os pontos vermelhos são os valores empíricos de $C_h(r, t)$ e as linhas tracejadas são ajustes exponenciais (via métodos de mínimos quadrados) para esses dados na escala mono-log.

devido à definição de $h_i(t)$ envolver essa quantidade. Para investigar essa hipótese, calculamos a função de correlação $C_p(r, t)$ considerando o tamanho da população $N_i(t)$ no lugar de $h_i(t)$ na Equação 2.1. A Figura 2.2C mostra que a correlação espacial para o tamanho da população é bem

descrita por uma decaimento lei de potência,

$$C_p(r, t) \sim r^{-\alpha(t)}, \quad (2.3)$$

com um expoente $\alpha(t) = 1,14$ no ano de $t = 2005$. Comportamentos análogos são obtidos para todos os outros anos (Figura 2.4) com praticamente os mesmos expoentes. De fato, como mostrado na Figura 2.2D, o comportamento de $\alpha(t)$ é bem aproximado por seu valor médio global ($\alpha(t) \approx 1,10$). Resultados semelhantes foram relatados para as correlações no tamanho da população de cidades dos Estados Unidos da América [103]. Dessa forma, a evolução de $C_h(r, t)$ para homicídios *per capita* não pode ser diretamente explicada pela dinâmica espacial do tamanho da população.

Outro aspecto interessante da correlação exponencial $h_i(t)$ é que ela pode ser considerada não trivial no contexto da equação Edwards-Wilkinson [104]. Essa equação é geralmente empregada para descrever a cinética estocástica de crescimento de superfícies na presença de difusão e aleatoriedade ambiental. No limite do contínuo, podemos escrever a equação de Edwards-Wilkinson para os homicídios *per capita* como

$$\frac{\partial}{\partial t} h(\mathbf{r}, t) = D \nabla^2 h(\mathbf{r}, t) + \eta(\mathbf{r}, t), \quad (2.4)$$

em que $h(\mathbf{r}, t)$ representa homicídios *per capita* em um ponto localizado pelo vetor \mathbf{r} , D é o coeficiente de difusão, ∇^2 é o Laplaciano e $\eta(\mathbf{r}, t)$ é um ruído não correlacionado (no espaço e no tempo): $\langle \eta(\mathbf{r}, t) \eta(\mathbf{r}', t') \rangle \propto \delta(\mathbf{r} - \mathbf{r}') \delta(t - t')$. Para esse modelo, a correlação é geralmente obtida pelas transformadas de Fourier no espaço e no tempo e retornando para o espaço usual via transformada inversa de Fourier (veja Refs. [104–106] para mais detalhes). Seguindo esses procedimentos, podemos escrever a correlação espacial em d dimensões e para grandes valores de t como

$$C_h(r) \sim \begin{cases} r^{2-d} & \text{se } d \neq 2 \\ \ln(1/r) & \text{se } d = 2. \end{cases} \quad (2.5)$$

O resultado anterior deve ser entendido como um modelo nulo para nossos dados e, portanto, seria de esperar um decaimento muito lento (um decaimento logaritmo, uma vez que o nosso sistema é essencialmente bidimensional) para a correlação espacial em $h_i(t)$, assumindo que a dinâmica

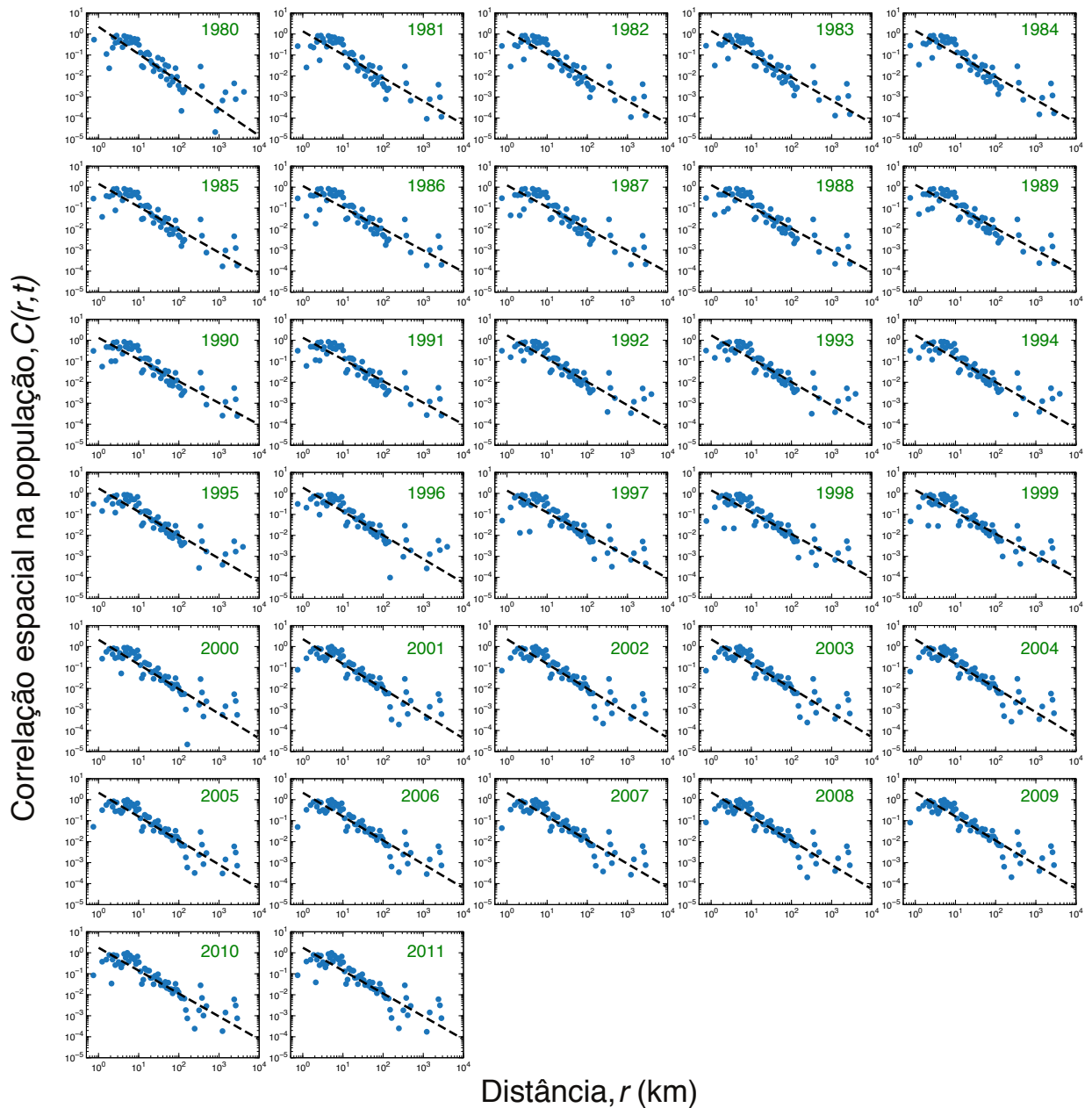


Figura 2.4: Correlação espacial na população. Cada gráfico mostra a dependência espacial da função de correlação $C_p(r, t)$ calculado para a população $N_i(t)$ em um dado ano t (como indicado nos gráficos). Os pontos vermelhos são os valores empíricos de $C_p(r, t)$ e as linhas tracejadas são ajustes lei de potência (via métodos de mínimos quadrados) para esses dados na escala log-log.

espacial de homicídios é exclusivamente governada por aspectos difusivos e pela aleatoriedade ambiental. De fato, uma aproximação de decaimento logaritmo para a correlação espacial foi observada em alguns sistemas sociais (taxas de participação nas eleições [107, 108] e obesidade [103])

e também para as taxas de ocorrências de raios [109]. Naturalmente, o modelo da Equação 2.4 é bastante grosseiro diante da complexidade envolvida em sistemas sociais. Talvez, um primeiro ingrediente a ser adicionado a esse modelo seria um termo de reação (levando-nos a uma equação de reação-difusão). Outro aspecto importante do nosso sistema é a distribuição espacial discreta e não homogênea das cidades, que também pode desempenhar um papel importante nas correlações espaciais.

As correlações espaciais em $h_i(t)$ também fornecem um indicativo da existência de aglomerados/*clusters* espaciais de cidades com valores similares de homicídios *per capita*. Para analisar essa questão, estudamos a formação de aglomerados/*clusters* espaciais de cidades por meio de uma análise de percolação [103, 110, 111]. Especificamente, definimos um limiar de homicídios T e selecionamos todas as cidades para as quais $h_i(t) \geq T$. Para o conjunto de cidades que satisfaz essa condição, aplicamos o algoritmo baseado na densidade espacial de aglomeração para aplicações com ruídos (DBSCAN — implementado na biblioteca `scikit-learn` da linguagem *Python* [112]) para a identificação de possíveis agrupamentos espaciais. Esse algoritmo encontra grupos estatisticamente significativos e, assim, podemos investigar como o tamanho do maior aglomerado/*cluster* S (isto é, o número de cidades no maior *cluster*) muda em função do limiar T , como mostrado na Figura 2.5 (linha vermelha) para o ano de $t = 1980$. Analogamente a uma transição de fase, S sofre uma mudança abrupta em torno de um valor específico de $T = T^*(t)$, a partir do qual o maior *cluster* começa a ser quebrado em aglomerados menores. Um comportamento semelhante foi observado para as taxas de obesidade nos Estados Unidos da América [103]. No nosso caso, o valor de $T^*(t)$ é definido como o ponto no qual o salto em S é máximo. Definimos ainda a magnitude de transição $M(t)$ como sendo o tamanho desse salto máximo. Nesse exemplo em particular, temos que em $T^*(t = 1980) = 5,67 \times 10^{-5}$ o número de cidades no maior *cluster* diminui em $M(t = 1980) = 314$. Observamos ainda a existência de mudanças subsequentes e abruptas de magnitudes menores em S , produzindo um comportamento tipo cascata. A fim de confirmar que esse comportamento semelhante ao de transição de fase não está relacionado à localização espacial das cidades, embaralhamos os valores de $h_i(t)$ entre as cidades e investigamos novamente a relação entre S e T . Para os dados aleatorizados, a Figura 2.5 mostra que o tamanho do maior aglomerado vai continuamente para zero conforme aumentamos T (linha azul). Assim, o comportamento de S versus T observado para os dados originais não pode ser diretamente relacionada à distribuição

espacial das cidades brasileiras.

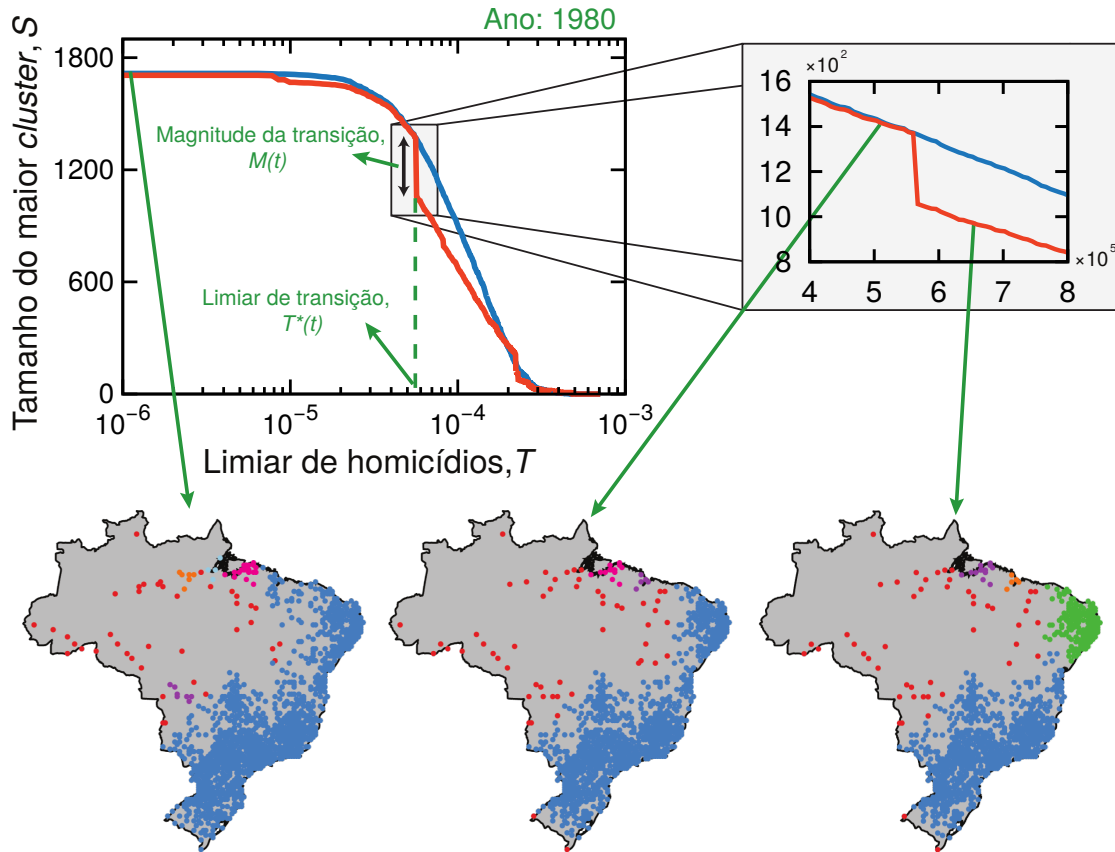


Figura 2.5: Formação de aglomerados/clusters espaciais de cidades e análise de percolação. O gráfico mostra o tamanho do maior *cluster* S (definido como o número de cidades que pertencem ao *cluster*) em função do limiar T do número de homicídios *per capita*. A linha vermelha mostra os resultados considerando os dados do ano $t = 1980$ e a linha azul mostra os resultados após o embaralhamento dos valores dos homicídios *per capita* $h_i(t)$ entre as cidades. Para os dados originais (linha vermelha), observamos que S sofre uma mudança abrupta em torno de um valor específico de $T = T^*(t)$, seguido de mudanças abruptas menores. A inserção mostra uma ampliação da região onde esse comportamento semelhante a uma transição de fase ocorre. O valor de $T^*(t)$ é definido como o ponto em que o salto em S é máximo e definimos ainda a magnitude de transição $M(t)$ como o tamanho desse salto máximo ($T^*(t)$ e $M(t)$ são ilustrados na figura). Os três mapas ilustram os aglomerados/clusters identificadas pelo algoritmo DBSCAN para os valores de T indicados pelas setas verdes.

Aplicamos a análise de *cluster* para todos os anos em nosso conjunto de dados e encontramos um comportamento de semelhante ao de uma transição de fase em todos eles (Figura 2.6), mas com algumas características evolutivas. O valor crítico $T^*(t)$ tem aumentado ao longo dos anos com uma taxa linear de aproximadamente $(0,10 \pm 0,01) \times 10^{-4}$ homicídios *per capita* por ano, como

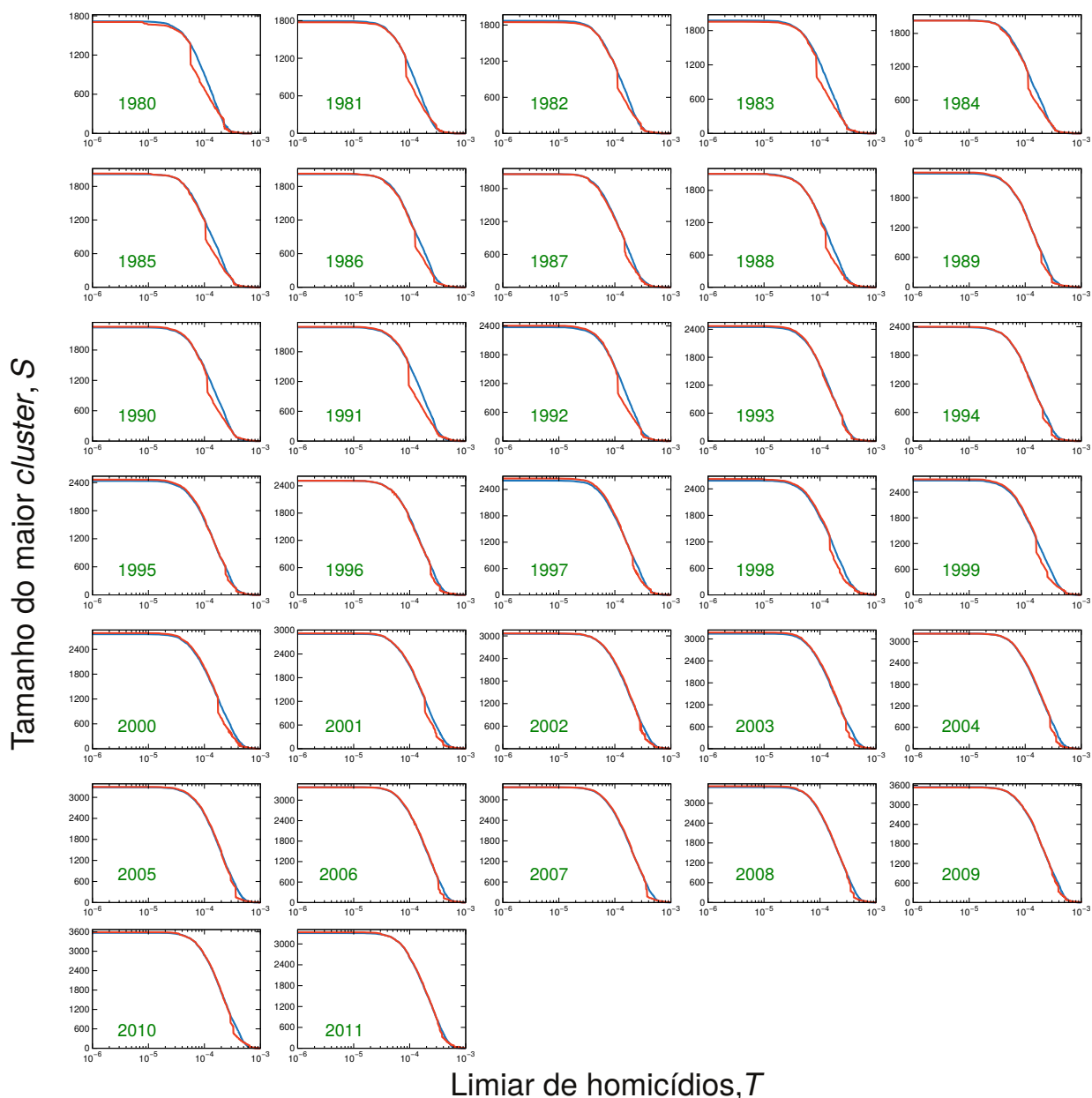


Figura 2.6: Formação de *clusters* espaciais das cidades e análise de percolação. Os gráficos mostram o tamanho do maior *cluster* S (definido como o número de cidades que pertencem ao *cluster*) como função do limiar T nos homicídios *per capita* para um dado ano t (tal como indicado nos gráficos). As linhas vermelhas mostram os resultados considerando os dados originais e as linhas azuis mostram os resultados após o embaralhamento dos valores dos homicídios *per capita* $h_i(t)$ entre as cidades.

mostrado na Figura 2.7A. Naturalmente, o crescimento de $T^*(t)$ deve estar relacionado à tendência crescente no homicídio *per capita* total do Brasil (Figura 2.1). A fim de explicar esse efeito,

dividimos $T^*(t)$ pelo homicídio global *per capita* no Brasil $h(t)$. A Figura 2.7B mostra que parte do comportamento evolutivo de $T^*(t)$ é realmente associado à $h(t)$. No entanto, observamos uma tendência estatisticamente significativa para o limiar escalado $T^*(t)/h(t)$. O limiar de homicídios escalado quase triplicou durante os 32 anos cobertos por nossos dados. Esse resultado concorda com o aumento no comprimento característico de correlação $r_c(t)$, no sentido que a intensificação das correlações nos últimos anos requer maiores limiares $T^*(t)$ para dividir os *clusters* espaciais. Para quantificar ainda mais os aspectos evolutivos dessas transições, investigamos se a magnitude de transição $M(t)$ evoluiu ao longo dos anos. Apesar de ser uma relação mais ruidosa, a Figura 2.7C mostra que $M(t)$ tem uma tendência decrescente significativa que pode ser aproximada por uma diminuição linear de ≈ 5 cidades por ano. Suspeitamos que parte desse comportamento também pode estar relacionado ao aumento de $r_c(t)$.

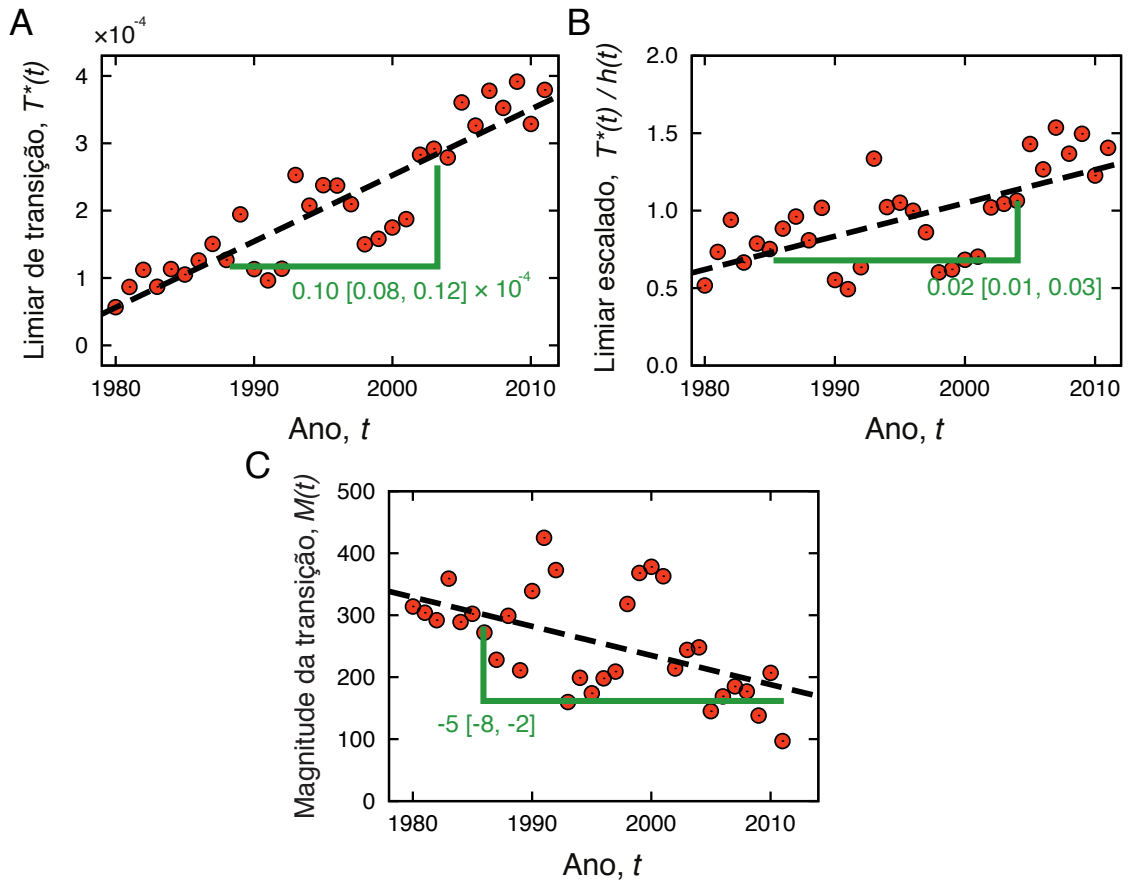


Figura 2.7: Aspectos evolutivos do comportamento de transição. A) Evolução no tempo do valor crítico $T^*(t)$ em que o tamanho do maior *cluster* S sofre uma mudança abrupta (pontos vermelhos). A linha a tracejada é um ajuste linear aos dados e o coeficiente linear bem como o seu intervalo de confiança são mostrados no gráfico. B) Evolução do limiar de homicídios escalado $T^*(t)/h(t)$ (pontos vermelhos). Mais uma vez, observamos uma tendência crescente que é aproximada por uma função linear (linha a tracejada) cujo coeficiente linear e seu intervalo de confiança são mostrados no gráfico. C) Alterações na magnitude de transição $M(t)$ ao longo dos anos (pontos vermelhos). Apesar da relação ruidosa, uma tendência de diminuição estatisticamente significativa é observada e pode ser aproximada por um decaimento linear (linha tracejada) cujo coeficiente linear e seu intervalo de confiança são mostrados no gráfico.

2.4 Conclusões e perspectivas

Nesse capítulo, caracterizamos a dinâmica espacial dos crimes de homicídio em todas as cidades brasileiras. Nossos resultados mostram que o homicídio *per capita* em uma cidade tem influência nas cidades próximas por meio de correlações de curto alcance. Ao investigar a evolução do comprimento de correlação característico, verificamos que a influência/correlação entre as cidades se intensificou consideravelmente ao longo dos últimos anos. Devido à existência dessas correlações, investigamos a formação de aglomerados/*clusters* espaciais de cidades por meio de uma análise de percolação. Observamos aglomerados estatisticamente significativos e um comportamento similar ao de transição de fase para o tamanho do maior *cluster* em função do limiar de homicídios. Esse comportamento apresenta características evolutivas caracterizadas por um aumento do limiar de homicídios em que as transições ocorrem e por uma diminuição nas magnitudes de transição (comprimento dos saltos no tamanho do *cluster*). Essas duas características podem ser relacionadas ao aumento do comprimento de correlação típico, uma vez que correlações mais intensas requerem limiares de homicídios maiores para dividir os aglomerados/*clusters* espaciais e podem contribuir também para a diminuição dos saltos nos tamanhos dos aglomerados. Assim, nossos resultados contribuem para indentificar padrões espaciais de atividades criminosas em grandes escalas e indicam que padrões empíricos observados no contexto da “teoria das janelas quebradas” no nível de cidade, isto é, o agrupamento de atividades criminosas, parecem emergir também em uma escala espacial muito maior. Acreditamos, ainda, que os nossos resultados empíricos abrem novas possibilidades para a modelagem de atividades criminosas. Por causa da descrição bem-sucedida (no nível de cidade), modelos baseados em equações de reação–difusão e processos pontuais — em inglês, *pointing process* — em conjunto com extensões da equação Edwards-Wilkinson são candidatos naturais a serem testados para reproduzir os padrões observados em escalas maiores.

Transientes de superdifusão e correlações de longo alcance nos padrões de motilidade dos protozoários tripanosomatídeos flagelados

Nesse capítulo, relatamos uma análise de difusão no movimento de quatro espécies de protozoários flagelados. Esses parasitas são os agentes etiológicos¹ de doenças tropicais negligenciadas: leishmaniose causada pela *Leishmania amazonensis* e *Leishmania braziliensis*, a doença do sono africano causada pelo *Trypanosoma brucei* e a doença de Chagas causada pelo *Trypanosoma cruzi*. Ao rastrear as posições desses parasitas e avaliar a variância relacionada às posições radiais, descobrimos que seus movimentos são caracterizados por um comportamento superdifusivo transiente. Além disso, a distribuição de probabilidade das posições radiais são autossemelhantes e podem ser aproximadas por uma distribuição gaussiana alongada. Investigamos também as distribuições de probabilidade das velocidades radiais das trajetórias individuais. Entre as várias distribuições candidatas, descobrimos que a distribuição gama generalizada mostra uma boa concordância com as distribuições empíricas. As séries temporais das velocidades têm correlações de longo alcance, exibindo um comportamento persistente (expoentes de Hurst próximos a um). A prevalência de

¹Os agentes etiológicos são aqueles microrganismos e toxinas microbianas que causam doenças em seres humanos e incluem bactérias, toxinas bacterianas, vírus, fungos, rickettsias, protozoários e parasitas [113].

padrões aproximadamente “universais” em todas as espécies analisadas indica que mecanismos semelhantes podem estar governando o movimento desses parasitas, apesar de suas diferentes características morfológicas. Além disso, uma análise mais aprofundada desses padrões poderia ser uma ferramenta útil para investigar a eficácia de novas drogas contra esses e outros parasitas relacionados à doenças tropicais negligenciadas.

3.1 Introdução e apresentação do problema

Protistas cinetoplastídeos² são responsáveis por inúmeras doenças em humanos e animais. Muitos desses protozoários são os agentes etiológicos de doenças tropicais negligenciadas [114]. Essas doenças afetam a vida de cerca de um bilhão de pessoas ao redor do mundo [115] e são consideradas um problema grave de saúde pública em vários países. As principais regiões afetadas são países localizados em áreas tropicais, onde os parasitas têm condições naturais adequadas para o ciclo de vida e onde insetos são normalmente abundantes [116]. Leishmaniose, doença de Chagas e a doença do sono africano são exemplos de doenças tropicais negligenciadas causadas por parasitas cinetoplastídeos — *Leishmania spp*, *Trypanosoma cruzi* e *Trypanosoma brucei*, respectivamente [117]. Esses protozoários têm um ciclo de vida complexo, alternando entre invertebrados (vetor) e hospedeiros vertebrados. Esses parasitas têm um flagelo, pelo menos durante uma das formas evolutivas de seu ciclo de vida [118–120]. O flagelo é uma organela multifuncional responsável pela propulsão da célula e está associado ao controle da morfogênese celular (processo de origem e desenvolvimento das formas e estruturas de um organismo) quimiotaxia (nome dado ao processo de atração de células em direção à um gradiente químico [121]), e ao processo de citocinese³ durante a última etapa do ciclo de divisão celular [122]. A motilidade⁴ dos parasitas é uma questão chave para a invasão hospedeiro–célula na fixação e colonização de tecidos do

²Reino protista agrupa organismos eucariontes, unicelulares, autótrofos e heterótrofos. Cinetoplastídeos é um grupo de protistas heterotróficos flagelados que inclui vários parasitas [113].

³A citocinese é o processo durante a divisão celular em que o citoplasma de uma única célula eucariótica é dividido para formar duas células filhas [113].

⁴Motilidade é a capacidade que os seres vivos possuem de se mover de forma espontânea. Um sinônimo aproximado é mobilidade; no entanto, no caso de células vivas que possuem a habilidade de se mover devido a mecanismos próprios (como o flagelo, no caso de parasitas flagelados) é mais comum o emprego da palavra motilidade, do inglês *motility*, visto que mobilidade não necessariamente significa que o objeto ou indivíduo tem a capacidade de se mover. Por exemplo, no caso do movimento do pólen em um líquido, o correto seria o emprego da palavra mobilidade para expressar a ideia de movimento, visto que o movimento do pólen é o resultado de forças externas que o impulsiona.

hospedeiro [122–127].

Movimentos difusivos estão onipresentes na natureza e desempenham um papel fundamental na motilidade de nado de microrganismos em meios líquidos [60, 128–133]. Pesquisadores de mecânica estatística concentraram-se em fenômenos em que a difusão anômala emerge [60, 61, 134, 135]. Por exemplo, a difusão anômala e distribuições de velocidade não gaussianas foram observadas no contexto de células *Hydra* [61] em que as densidades de entropia máxima associadas às medidas entrópicas não usuais foram usadas para descrever o movimento dessas células. Motilidade e padrões de difusão também foram investigadas em protozoários [123, 124, 126, 136–140]. Por exemplo, há estudos sobre *T. brucei* investigando o movimento de propulsão [126], flexibilidade e direcionalidade [138] e adaptações ambientais do corpo [139]. As interações célula-hospedeiro [123] para *Leishmania spp* e os batimentos do flagelo [140] para *T. cruzi* também foram estudadas. A motilidade está fortemente relacionada à viabilidade celular em todas as espécies flageladas cinetoplastídeas e é amplamente utilizada como uma medida para viabilidade [124, 136, 137].

Cada espécie de protozoário tem adaptações únicas dependendo das diferentes condições de vida. A investigação da dinâmica e difusão desses microrganismos pode contribuir para um avanço na compreensão dos mecanismos de patogênese microbiana e também para o campo de padrões difusivos. No entanto, ainda falta uma compreensão mais completa e geral de padrões de motilidade desses parasitas. Para superar essa lacuna, estudamos a dinâmica de difusão de agentes causadores de doenças tropicais negligenciadas. Ao rastrear as posições desses parasitas, apresentamos uma caracterização completa de seus padrões de motilidade. Especificamente, mostramos que a propagação das trajetórias é superdifusiva para tempos curtos e que as distribuições de probabilidades relacionadas às posições radiais diferem das previsões da equação de difusão usual. Verificamos ainda que a série temporal da velocidade das trajetórias individuais tem correlações de longo alcance e são bem aproximadas por uma distribuição gama generalizada. Nossos resultados revelam alguns padrões “universais” de motilidade dos parasitas que poderiam facilitar a identificação de novos alvos para intervenção terapêutica.

3.2 Métodos

3.2.1 Manutenção e cultura dos parasitas

Os experimentos descritos a seguir foram realizados em parceria com o Departamento de Farmácia da Universidade Estadual de Maringá, no Laboratório de Inovação Tecnológica no Desenvolvimento de Fármacos e Cosméticos.

Leishmania amazonensis (cepa MHOM/BR/75/Josefa) e *Leishmania braziliensis* em forma de promastigotas foram cultivadas dentro de frascos de cultura celular contendo meio de Warren (infusão de cérebro e coração mais hemina bovina e ácido fólico, pH 7,2) suplementadas com 10% de soro fetal bovino (FBS). Os parasitas foram incubados a 25°C por 48 h. *Trypanosoma cruzi* em formas epimastigotas (cepa Y) foram mantidas em meio LIT (*Liver Infusion Tryptose*, pH 7,2), suplementado com 10% de FBS e incubados a 28°C durante 96 h. Tripomastigotas de *Trypanosoma brucei brucei* (cepa EATRO-427) foram cultivadas em meio HMI-9, suplementados com 10% de FBS e incubados a 37°C e 5% de pressão de CO₂ durante 24 h. Esses períodos de incubação são essenciais para colher os protozoários em fase de crescimento exponencial. A Figura 3.1 ilustra as formas amastigota, promastigota, epimastigota e tripomastigota dos protozoários.

3.2.2 Configuração experimental, aquisição de imagem e rastreamento

Após os períodos de incubação, preparamos uma suspensão contendo cerca de 6×10^6 parasitas para as espécies de *Leishmania*, *T. cruzi* e *T. brucei* em Warren, LIT, e DMEM (*Dulbecco Modified Eagle Medium*, suplementado com 2 mM de L-glutamina, pH 7,4), respectivamente. Os meios não foram suplementados com FBS. Em seguida, 10 μ L das suspensões de protozoários foram colocadas entre uma lâmina e lamínula de vidro para iniciar a aquisição de imagens (Figura 3.2A). A espessura entre a lâmina e lamínula de vidro é comparável com o tamanho do protozoários ($\sim 5 \mu\text{m}$) [142–144], confinando o movimento em duas dimensões. Usamos o microscópio BA410E da Motic equipado com uma câmera CMOS de 5.0 *Megapixels* com uma resolução de $800 \times 600 \text{ pixels}$ (Figura 3.2B), taxa de aquisição de 10 *frames* por segundo e ampliação de 20 vezes. A área coberta pelo microscópio nessa configuração é de $285,12 \times 213,84 \mu\text{m}^2$. O comprimento dos vídeos adquiridos foi de 10 minutos para cada amostra. Repetimos esse procedimento

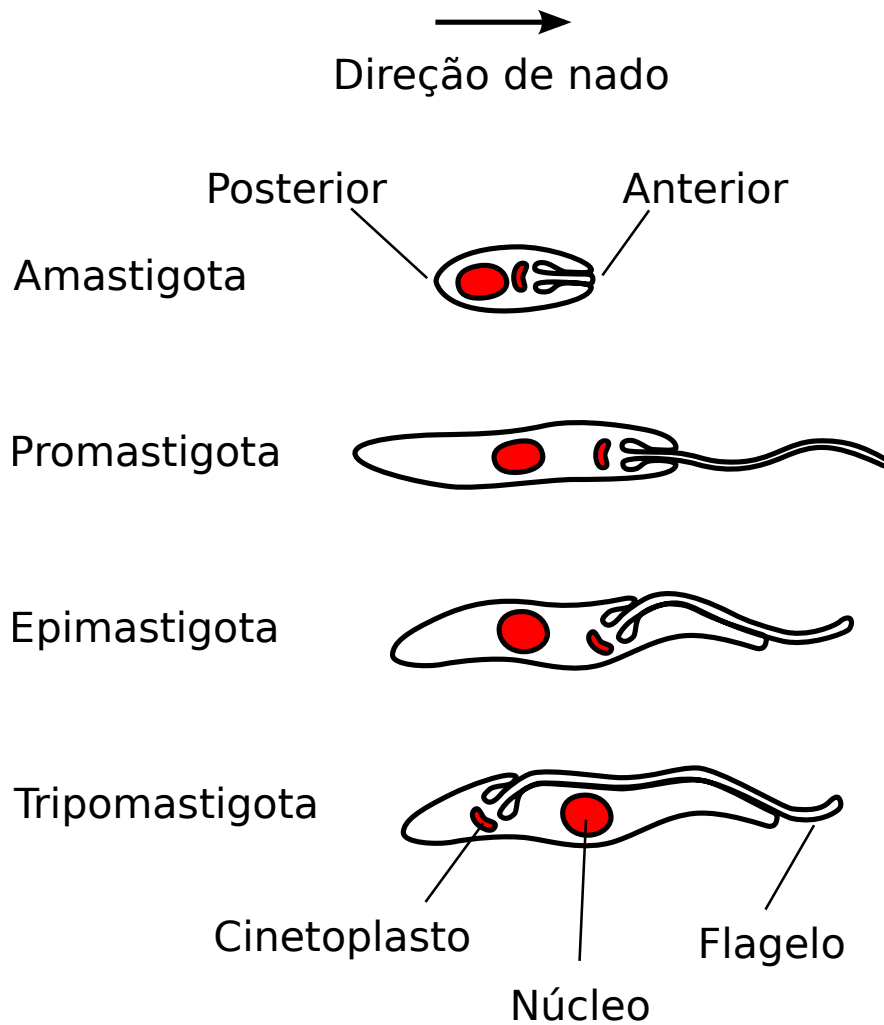


Figura 3.1: Ilustração das formas dos protozoários. A estrutura maior em vermelho representa o núcleo da célula, enquanto que, a estrutura menor (também em vermelho) representa o cinetoplasto. A estrutura citoplasmática fornece energia para a motilidade do flagelo. A forma amastigota é uma morfologia em que o flagelo é muito curto [141]. A forma promastigota é uma morfologia em que o flagelo está na parte anterior da célula e não está anexado ao corpo celular [141]. A forma epimastigota é uma morfologia em que o flagelo sai da extremidade anterior da célula e está ligado ao longo de parte do corpo celular [141]. A forma tripomastigota é uma forma em que o flagelo está ligado à maior parte do corpo da célula por uma membrana ondulante [141].

três vezes para cada espécie de protozoário. A fim de extrair as trajetórias dos vídeos, utilizamos um algoritmo de rastreamento de movimento em sequências de imagens implementado em Matlab [145]. Excluimos as trajetórias com menos de 500 passos de tempo para garantir trajetórias com tempo suficiente para a análise estatística. Para essas trajetórias, removemos os primeiros e

últimos 50 passos devido à imprecisão do algoritmo em rastrear as posições dos protozoários⁵. Por inspeção visual, removemos também as trajetórias para as quais o algoritmo confundiu dois ou mais microrganismos em algum passo ao longo da trajetória. Todas as trajetórias foram suavizadas mediante a aplicação de um filtro de média móvel de comprimento 10. A viscosidade do meio de cultura foi medida num viscosímetro (*Visco Star Plus*) a 25°C e 50 rpm. Os valores da viscosidade e o número de trajetórias analisadas para cada protozoário são mostradas na Tabela 3.1.

Tabela 3.1: Resumo dos dados

Protozoário	Total de trajetórias	Meio	Viscosidade [mPa·s]
<i>L. amazonensis</i>	105	Warren	0,38
<i>L. braziliensis</i>	135	Warren	0,38
<i>T. brucei</i>	131	DMEM	0,71
<i>T. cruzi</i>	153	LIT	0,39

Na Figura 3.3A mostramos trajetórias típicas da *L. amazonensis*, isto é,

$$\vec{r}_i(t) = [x_i(t), y_i(t)], \quad (3.1)$$

em que $x_i(t)$ e $y_i(t)$ são as componentes horizontal e vertical do vetor posição $\vec{r}_i(t)$ no tempo t para o i -ésima trajetória. Na Figura 3.3B, mostramos a série temporal da velocidade radial correspondente à essas trajetórias, ou seja,

$$v_i(t) = \sqrt{v_{x_i}^2(t) + v_{y_i}^2(t)}, \quad (3.2)$$

em que

$$v_{x_i}(t) = [x_i(t + \Delta t) - x_i(t)]/\Delta t \quad (3.3)$$

e

$$v_{y_i}(t) = [y_i(t + \Delta t) - y_i(t)]/\Delta t \quad (3.4)$$

com $\Delta t = 1/10$ s. Exemplos de trajetórias e séries temporais de velocidades para os outros protozoários são mostrados na Figura 3.4. A partir desses dados, podemos observar padrões circulares rela-

⁵O algoritmo de aquisição de imagens depende da diferença entre *frames* para detectar o movimento e identificar os indivíduos (Figura 3.2). No começo e fim das trajetórias, o algoritmo não é tão eficiente em identificar os objetos, ou porque a sequência de imagens não é clara, ou porque o protozoário saiu do foco do microscópio por algumas imagens.

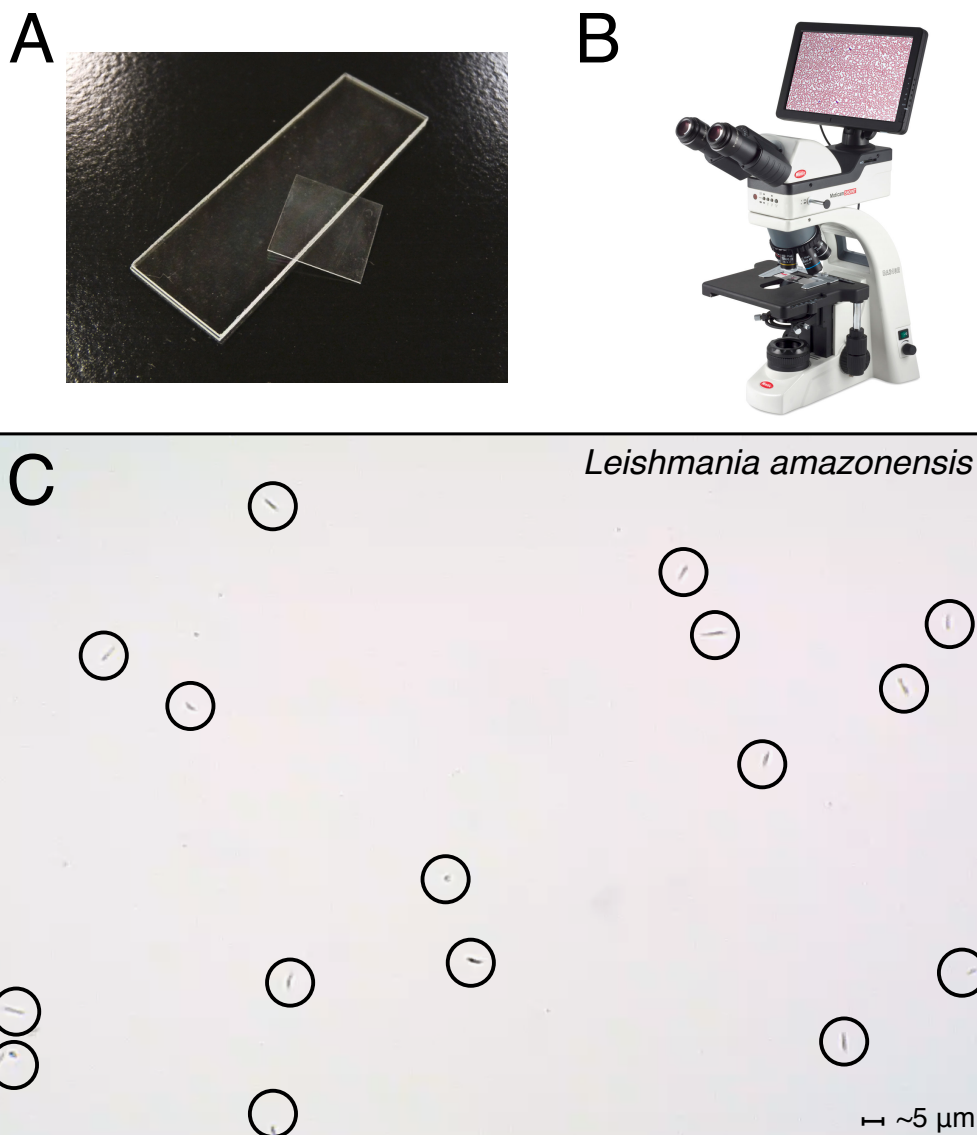


Figura 3.2: Configuração experimental para aquisição das trajetórias. A) Lâmina e lamínula de vidro. B) Microscópio BA410E da Motic equipado com uma câmera CMOS de 5.0 Megapixels com uma resolução de 800×600 pixels. C) Screenshot de um dos vídeos dos protozoários da espécie *L. amazonensis*. Os círculos identificam os protozoários para essa imagem.

cionados a possíveis interações hidrodinâmicas com as paredes, como já observado no movimento de bactérias *Escherichia coli* [146]. Experimentos adicionais considerando diferentes limites e espessura entre a lâmina a lamínula de vidro poderiam clarificar melhor essa questão sobre os efeitos das paredes sobre a motilidade desses protozoários. Porém, nesse estudo, apenas essa configuração foi testada. Calculamos também a média (Figura 3.3C) e desvio padrão (Figura 3.3D) das velocidades para cada espécie de protozoário. É interessante notar que as velocidades dos microrganismos

dependem da viscosidade do meio de cultura [139] e outros valores podem ser encontrados na literatura devido à viscosidades diferentes [126, 138–140]. Por exemplo, foi verificado que na forma selvagem na corrente sanguínea o *T. brucei* pode atingir velocidades muito mais elevadas (cerca de $30 \mu\text{m/s}$) [139]. *Trypanosoma* spp usam a motilidade como uma ferramenta para contornar as células do sistema imunológico e remover a ligação entre sua superfície e as moléculas de anticorpos. Nos vasos sanguíneos, onde os protozoários estão adaptados para sobreviver, as células vermelhas e brancas do sangue comportam-se como suporte para o flagelo que ajuda a impulsionar o corpo do protozoário. O mesmo comportamento é observado em líquidos mais viscosos, justificando o aumento da velocidade do parasita nesses ambientes [139]. Outros fatores que afetam as velocidades incluem sinais químicos, teor de oxigênio, pressão, vazão e confinamento [139].

3.3 Resultados e discussão

Começamos pela caracterização das trajetórias de todos os protozoários. A fim de fazer isso, consideramos a série temporal da magnitude do vetor posição $\vec{r}_i(t)$ depois de subtrair a sua posição inicial $\vec{r}_i(0) = [x_i(0), y_i(0)]$. Medimos a difusão calculando a dependência temporal da variância da posição radial

$$\begin{aligned}\sigma^2(t) &= \langle [\vec{r}_i(t) - \langle \vec{r}_i(t) \rangle]^2 \rangle \\ &= \frac{1}{N_k(t) - 1} \sum_{i=1}^{N_k(t)} (\vec{r}_i(t) - \langle \vec{r}_i(t) \rangle)^2,\end{aligned}\quad (3.5)$$

em que

$$\langle \vec{r}_i(t) \rangle = \frac{1}{N_k(t)} \sum_{i=1}^{N_k(t)} \vec{r}_i(t) \quad (3.6)$$

é a posição média radial e $N_k(t)$ é o número de trajetórias para a k -ésima espécie (Tabela 3.1). Para processos difusivos usuais a variância aumenta linearmente com tempo, isto é,

$$\sigma^2(t) \sim t. \quad (3.7)$$

Esse comportamento usual (movimento browniano) está relacionado à falta de memória ao longo da trajetória das partículas, assim como indica uma escala característica finita para os incrementos

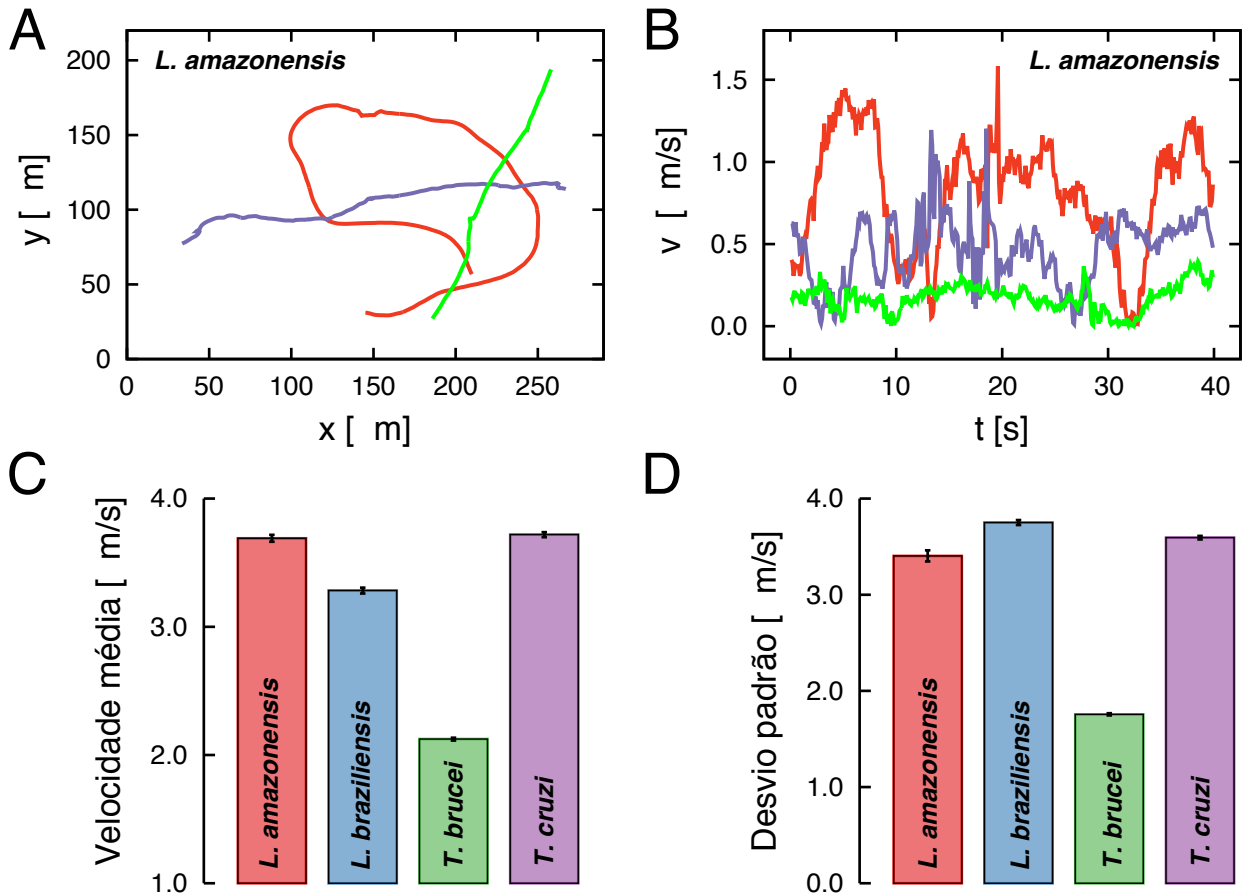


Figura 3.3: Ilustração de trajetórias e velocidades dos protozoários. Mostramos em A) as trajetórias típicas do protozoário *L. amazonensis* e em B) as séries temporais correspondentes às velocidades radiais $v(t)$. A Figura 3.4 mostra essas quantidades para os outros protozoários. Os gráficos de barras representam as C) velocidades médias e D) os desvios padrão da velocidade de todas as trajetórias para cada espécie de protozoário. As barras de erro são intervalos de confiança de 95% calculados via *bootstrapping*.

de posição e para os tempos de espera entre voos. No entanto, processos difusivos mais complexos costumam exibir desvios desse comportamento linear. Quando isso acontece, um comportamento típico para a variância é uma dependência lei de potência no tempo [147],

$$\sigma^2(t) \sim t^\lambda, \quad (3.8)$$

em que $0 < \lambda < 1$ corresponde a subdifusão e $\lambda > 1$ a superdifusão. No nosso caso, a evolução das variâncias é mostrada na Figura 3.5A, em que é evidente que a difusão das trajetórias dos

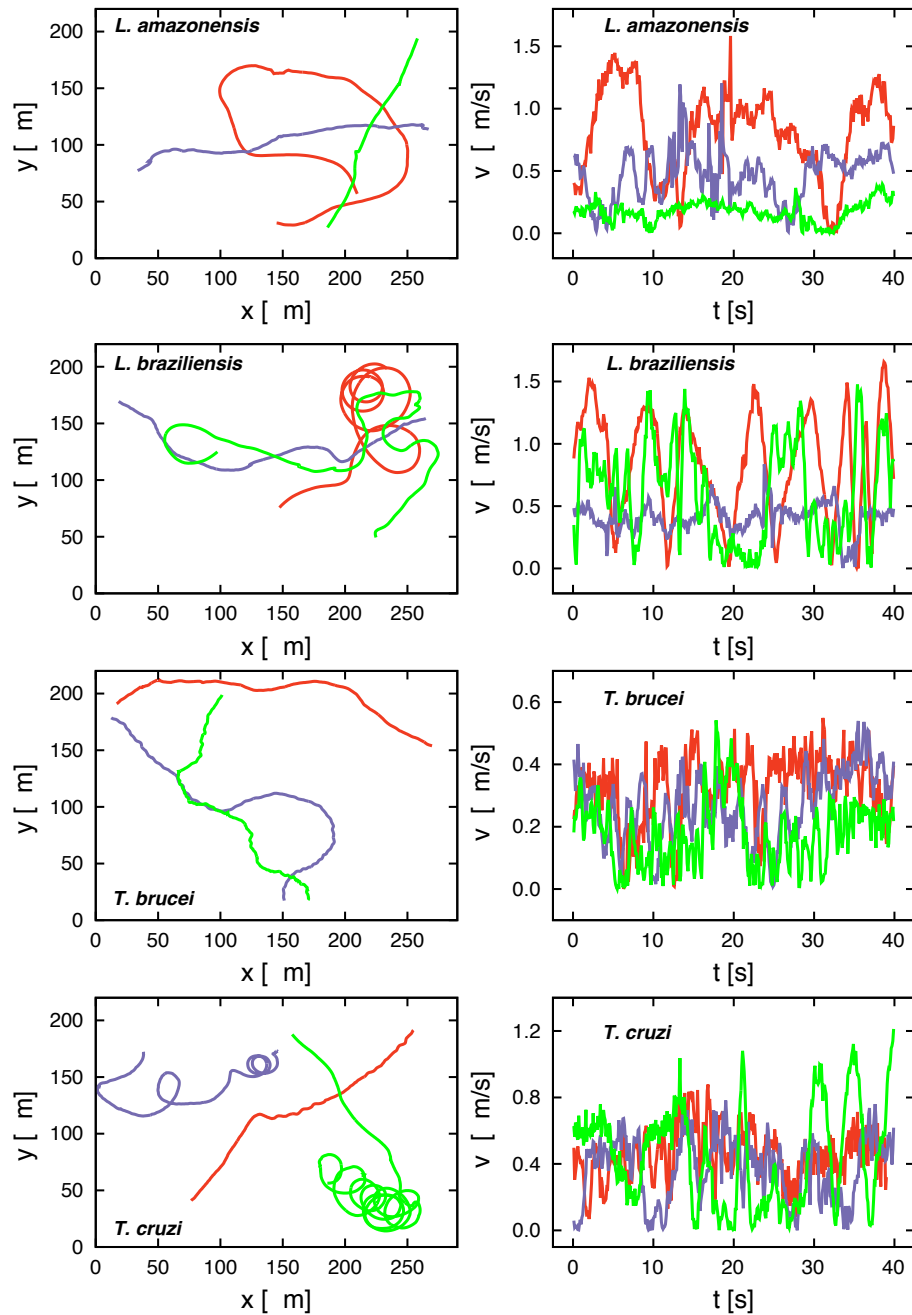


Figura 3.4: Trajetórias e velocidades das quatro espécies de protozoários. Os painéis à esquerda mostram as trajetórias típicas dos quatro protozoários. Os painéis à direita mostram as séries temporais das velocidades $v(t)$.

protozoários ocorre muito mais rapidamente do que o esperado por um movimento browniano. Observamos, ainda, um comportamento aproximadamente lei de potência por mais de duas décadas na escala temporal ($t < 10$ segundos). Ajustamos as relações via métodos dos mínimos quadrados

em log-log ($\log \sigma^2(t)$ versus $\log t$) e descobrimos que os valores de λ são, de fato, muito maiores do que um. Como mostrado na Figura 3.5B, λ varia de 1,69 para o *L. braziliensis* até 1,93 para o *T. brucei*. Portanto, as quatro espécies de protozoários flagelados exibem um comportamento fortemente superdifusivo para tempos curtos. Foi observado um expoente semelhante para o movimento de partículas intracelulares, com $\lambda \approx 1,8$ [62]. Vale a pena notar que microrganismos podem apresentar alguma persistência relacionada ao seu mecanismo de orientação para curtos períodos. Assim, esperamos que a difusão possa, eventualmente, aproximar do regime usual para trajetórias mais longas. Portanto, os valores de λ obtidos aqui, representam um comportamento inicial de tempos curtos e outros regimes poderiam ser observados para trajetórias mais longas. De fato, podemos constatar na Figura 3.5A que as curvas começam a dobrar para baixo para $t > 10$ s. Na Figura 3.5C, mostramos o expoente λ calculado dentro de uma janela de tamanho 30 segundos centralizada em t_w como função de t_w , na qual observamos que os valores de λ diminuem com t_w e se aproximam do valor esperado pela difusão usual $\lambda = 1$.

Embora todas as espécies estudadas apresentem superdifusão para períodos curtos, percebemos um aspecto intrigante do movimento relacionado às velocidades e à análise de difusão para o *T. brucei*. Esse protozoário mostra a menor velocidade e o maior expoente de difusão em meio de cultura, sugerindo que a motilidade do *T. brucei* é mais direcional do que as outras espécies de protozoários considerada nesse trabalho. A motilidade direcional, provavelmente, ocorre porque o *T. brucei* é um parasita extracelular livre e se espalha por vários tecidos. Formas tripomastigotas de *T. brucei* são altamente adaptadas para viver em espaços intercelulares [148–150]. Durante o espalhamento no hospedeiro, parasitas *T. brucei* penetram entre as células do tecido onde existem fibras de colágeno que podem facilitar ou dificultar o movimento. Para os tripanossomas atingirem a velocidade máxima para a frente, uma densidade específica de células é necessária. Se a densidade de obstáculos assemelha-se a das redes de colágeno, os protozoários nadam para trás, a fim de evitar ficarem presos [139]. Esses parasitas têm um antígeno⁶ na superfície inteira chamado glicoproteína de superfície variante (VSG) [151]. A VSG é diferente entre os indivíduos de uma população de *T. brucei* e impede a ligação específica com os anticorpos que podem matar os parasitas [151]. Supomos que a presença de moléculas de superfície, como a VSG, reduz a dificuldade

⁶Antígeno é uma substância orgânica, de natureza geralmente proteica que, inoculada no organismo, provoca a produção de anticorpos específicos como reação.

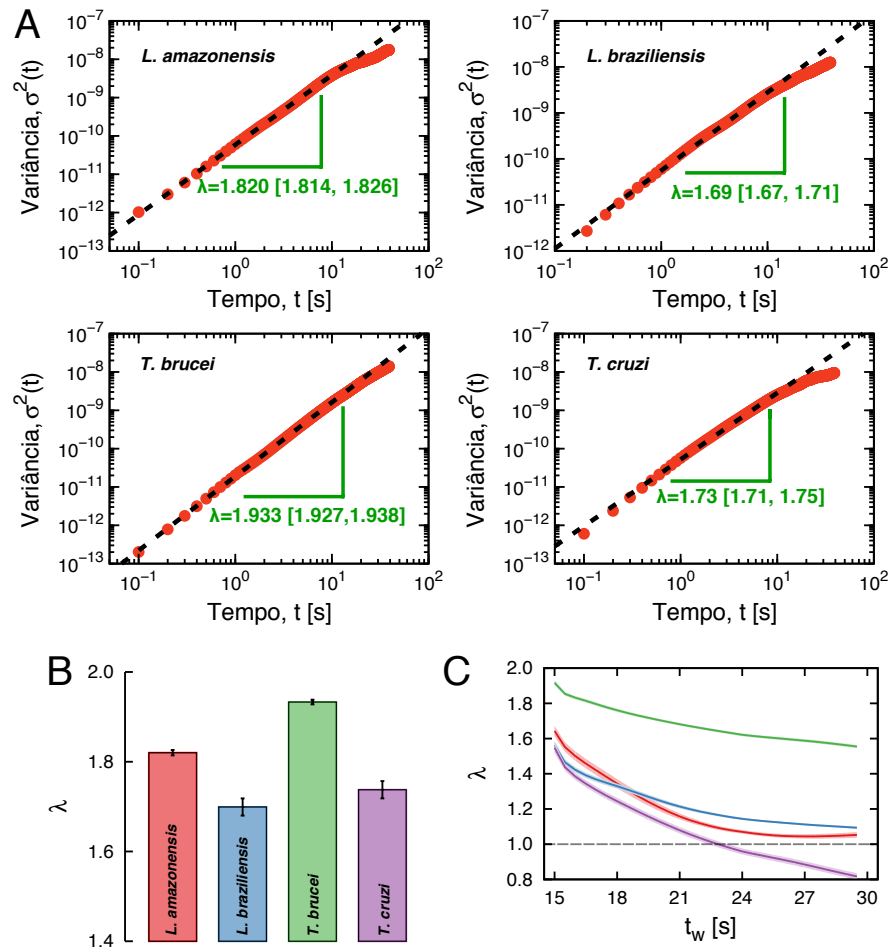


Figura 3.5: Transiente superdifusivo nas trajetórias dos protozoários. A) Mostramos os valores experimentais da variância $\sigma^2(t)$ das posições radiais (pontos vermelhos) para as quatro espécies (como indicado nos gráficos) em escala log-log. As linhas tracejadas representam as relações lei de potência, $\sigma^2(t) \sim t^\lambda$, para as quais os valores de λ foram obtidos (via método dos mínimos quadrados) ajustando o modelo linear para essas relações em escala log-log (considerando $t < 10$ segundos). Mostramos os valores de λ (e seus intervalos de confiança de 95% via *bootstrapping*) nos gráficos e também no painel B) como gráficos de barra, em que as barras de erro correspondem aos intervalos de confiança de 95% via *bootstrapping*. C) Expoentes λ calculados dentro de uma janela de tamanho 30 segundos centralizada em t_w como função de t_w . As diferentes cores (as mesmas usadas em B) representam os quatro protozoários. As pequenas regiões sombreadas representam os intervalos de confiança de 95% via *bootstrapping* e a linha cinza tracejada representa o regime usual $\lambda = 1$.

de movimento de células nos meios e em seus arredores e permite o parasita *T. brucei* se difundir mais rápido.

Formas epimastigotas de *T. cruzi* e promastigotas de *Leishmania spp* exibem velocidades e expoente de difusão semelhantes. Essas formas são mais rápidas do que as formas tripomastigotas

de *T. brucei*, têm menor expoente de difusão e têm diferentes estratégias nesse estágio do ciclo de vida. No ambiente natural, epimastigotas de *T. cruzi* e promastigotas de *Leishmania spp* não precisam viajar/nadar longas distâncias. Especificamente, as formas epimastigotas de *T. cruzi* devem migrar para o intestino do inseto vetor durante um período específico do ciclo de vida para a proliferação [152]. No caso de formas promastigotas de *Leishmania spp*, elas devem ser fagocitadas⁷ pelas células de mamíferos para completar o seu ciclo de vida. Nessa fase, os parasitas *Leishmania* secretam uma substância chamada gel secretado por promastigotas (PSG). O PSG é um gel de tipo mucina⁸ e foi demonstrado ser um fator importante para o crescimento de amastigotas (Figura 3.1) no seu ambiente intracelular [153, 154]. A produção de PSG parece ser responsável pela formação de aglomerados (rosetas) e poderia desempenhar um papel de uma barreira para o movimento difusivo [155, 156]. Em geral, sugerimos que as mudanças nos parâmetros de motilidade e as moléculas de superfície, as quais afetam barreiras mecânicas ou físicas que restringem a capacidade das células de se difundir livremente, poderiam contribuir significativamente para a virulência desses parasitas.

Outra característica marcante dos processos difusivos está relacionada às distribuições de probabilidade das posições. Para a difusão usual com simetria radial, essa distribuição pode ser obtida ao resolver a seguinte equação diferencial

$$\frac{\partial P(\vec{r}, t)}{\partial t} = D \nabla^2 P(\vec{r}, t) \quad (3.9)$$

em que D é o coeficiente de difusão e ∇^2 é o Laplaciano em duas dimensões. Considerando $P(r \rightarrow \infty, t) \rightarrow 0$ e $P(\vec{r}, 0) = \delta^2(\vec{r})$, podemos mostrar que a distribuição de probabilidade das posições radiais é

$$P(r, t) = \frac{r}{2Dt} \exp\left(\frac{-r^2}{4Dt}\right). \quad (3.10)$$

Vale a pena notar que essa solução conduz a um comportamento linear para a variância ao longo do tempo, isto é, $\sigma^2(t) = 4Dt$. Além disso, as distribuições dadas pela Equação 3.10 são autossimilares

⁷Fagocitado, que sofreu fagocitação. A fagocitação é o processo que engloba e ingere (os fagócitos) corpos estranhos, bactérias, células mortas ou quase mortas do organismo.

⁸Mucina é constituída de proteínas fortemente glicosiladas e constituem o principal componente do muco.

no tempo e colapsam em uma única curva,

$$P(\xi) = 2 \xi \exp(-\xi^2), \quad (3.11)$$

para posição reescalada $\xi(t) = r(t)/\sigma(t)$. A equação de difusão usual (Equação 3.9) e sua solução (Equação 3.10) podem ser entendidas como um modelo nulo para as distribuições das posições radiais dos protozoários se assumirmos que eles se comportam como partículas brownianas. Assim, o desvio desse comportamento é outro indicativo de difusão anômala.

Calculamos a evolução temporal das distribuições empíricas das posições radiais (Figura 3.6A). Essas distribuições deslocam-se para valores positivos de r e tornam-se de cauda mais longas com o passar do tempo t . Depois de considerar a posição reescalada $\xi(t) = r(t)/\sigma(t)$, observamos que todas as distribuições colapsam em uma única curva (Figura 3.6B). Esse resultado demonstra a natureza autossimilar das trajetórias dos protozoários, mas também revela desvios notáveis entre as distribuições empíricas e o esperado pela equação de difusão usual (linhas tracejadas cinzentas na Figura 3.6B).

Em uma tentativa de encontrar uma melhor descrição para essas distribuições empíricas, propomos uma substituição da expressão gaussiana na Equação 3.10 por uma gaussiana alongada (caracterizada por um outro parâmetro $\delta > 0$). Especificamente, consideramos a seguinte distribuição de probabilidade para as posições radiais

$$P(r,t) = \frac{\delta}{\Gamma(2/\delta)} \frac{r}{(2Dt)^{2/\delta}} \exp\left(-\frac{|r|^\delta}{2Dt}\right), \quad (3.12)$$

em que $\Gamma(x) = \int_0^\infty y^{x-1} e^{-y} dy$ é a função gama. Observamos que $\delta = 2$ recupera a gaussiana usual. No entanto, para $0 < \delta < 2$ a cauda dessa distribuição vai a zero mais lentamente do que o caso usual, enquanto para $\delta > 2$ ela decai mais rapidamente. Uma generalização semelhante foi proposta por Richardson [157] no contexto de difusão atmosférica, considerando um coeficiente de difusão dependente do espaço. Essa generalização é considerada uma das primeiras equações de difusão anômala. A distribuição dada pela Equação 3.12 também é autossemelhante e para a posição radial reescalada $\xi(t) = r(t)/\sigma(t)$, ela pode ser escrita como

$$P(\xi) = \frac{\delta}{\Gamma(2/\delta)} \xi e^{-|\xi|^\delta}. \quad (3.13)$$

Ajustamos a Equação 3.13 para os valores médios por janela das distribuições empíricas via método dos mínimos quadrados. As linhas contínuas na Figura 3.6B mostram que o ajuste está longe de ser perfeito, mas essa generalização é uma descrição melhor quando comparada à distribuição obtida a partir da equação de difusão usual (Equação 3.11). Portanto, o movimento difusivo dos protozoários exibe simultaneamente um comportamento de escala anômalo para a variância $\sigma^2(t)$ e distribuições de posição radial com caudas muito mais longas do que as esperadas para caminhantes brownianos. Os valores dos melhores parâmetros de ajuste δ são mostrados na Figura 3.6C, em que se observa que o *L. amazonensis* e o *T. brucei* exibem caudas maiores ($\delta \approx 1$) quando comparados ao *L. braziliensis* e ao *T. cruzi* ($\delta \approx 1,25$). Também é interessante notar que os valores de δ estão relacionados aos valores de λ . Na verdade, a dependência do tempo da variância calculada a partir da Equação 3.12 é

$$\sigma^2(t) \propto t^{2/\delta}, \quad (3.14)$$

levando a $\delta = 2/\lambda$, uma relação que é aproximadamente válida para os valores de δ obtidos por meio do procedimento de ajuste por mínimos quadrados (Figura 3.6C).

Para caracterizar ainda mais o movimento difusivo dos protozoários, investigamos a série temporal da velocidade radial $v_i(t)$ relacionada às trajetórias individuais (Figura 3.3B). Primeiro perguntamos se essas séries temporais têm correlações de curto ou de longo alcance. Para responder a essa pergunta, aplicamos a análise de flutuação destendenciada (DFA — *Detrended Fluctuation Analysis*) [7, 158] para essas séries temporais. Essa técnica consiste em quatro passos: *i*) primeiro, definimos o perfil integrado $Y(t) = \sum_{k=1}^t [v_i(k) - \bar{v}_i]$, em que $\bar{v}_i = 1/t_{\max} \sum_t^{t_{\max}} v_i(t)$ com t_{\max} , sendo o comprimento da *i*-ésima série temporal; *ii*) a seguir, dividimos $Y(t)$ em $N_n = t_{\max}/n$ segmentos não sobrepostos de tamanho n ; *iii*) para cada segmento, calculamos uma tendência polinomial local e subtraímos de $Y(t)$ (aqui usamos uma função linear, mas ordens superiores não alteram nossos resultados), definindo $Y_n(t) = Y(t) - p_\nu(t)$, em que $p_\nu(t)$ representa a tendência local no ν -ésimo segmento; *iv*) finalmente, calculamos a função de flutuação

$$F(n) = \left[\frac{1}{N_n} \sum_{\nu=1}^{N_n} \langle Y_n(t)^2 \rangle_\nu \right]^{1/2}, \quad (3.15)$$

em que $\langle Y_n(t)^2 \rangle_\nu$ é o valor quadrático médio de $Y_n(t)$ sobre os dados no ν -ésimo segmento. Se a série temporal de velocidade é autossimilar, a função de flutuação $F(n)$ apresenta uma dependência

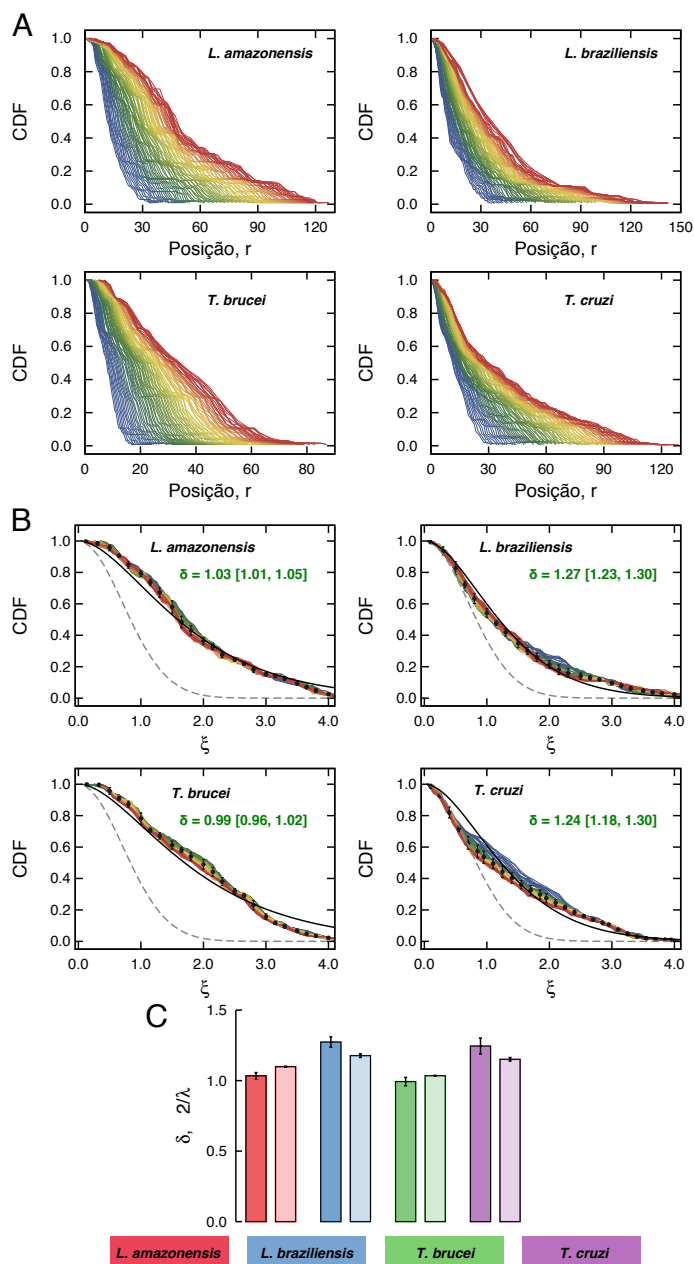


Figura 3.6: Natureza autossimilar das posições dos protozoários. A) Evolução temporal das distribuições acumuladas (CDFs) da posição radial $r(t)$ para cada protozoário. O código de cores indica o tempo t , variando do azul ($t = 2$ s) ao vermelho ($t = 10$ s). B) As mesmas distribuições para as posições reescaladas $\xi(t) = r(t)/\sigma(t)$. Os pontos pretos representam médias por janela das CDFs e as barras de erro representam uma unidade de desvio padrão. As linhas cinzas tracejadas representam a distribuição acumulada da gaussiana radial normalizada (Equação 3.11). As linhas contínuas são a versão acumulada da distribuição da Equação 3.13 e os melhores valores para o parâmetro δ são mostrados nos gráficos. C) Comparação entre os valores de δ obtidos via ajuste das médias por janela das CDFs pelo métodos dos mínimos quadrados e os valores esperados pela dependência temporal dos desvios, isto é, $\delta = 2/\lambda$.

lei de potência na escala de tempo n , ou seja,

$$F(n) \sim n^h, \quad (3.16)$$

em que h é o expoente de Hurst. Se $h = 1/2$, as velocidades não são correlacionadas ou apresentam correlação de curto alcance, enquanto $h \neq 1/2$ indica que a série temporal é correlacionada de longo alcance.

Aplicamos o procedimento acima descrito para todas as séries temporais de velocidade e um comportamento típico para a função de flutuação $F(n)$ está representado na Figura 3.7A. Mostramos outros exemplos de função de flutuação para cada espécie de protozoário na Figura 3.8. Nesse gráfico em log-log, ajustamos um modelo linear para a obtenção do expoente de Hurst h , que é de $h = 1,12$ para a série temporal original e está perto de $1/2$ para versões embaralhadas da série temporal. Calculamos o expoente de Hurst para todas as séries temporais e os valores médios para cada protozoário são mostrados na Figura 3.7B. Essas médias são praticamente indistinguíveis umas das outras e indicam que as velocidades dos protozoários têm correlações de longo alcance. Além disso, essas velocidades apresentam um comportamento persistente forte (visto que $h \approx 1$), isto é, incrementos positivos de velocidade são seguidos por incrementos positivos e incrementos negativos são seguidos por incrementos negativos, muito mais frequentemente do que ao acaso. Também avaliamos as distribuições de h para cada espécie protozoário para todas as trajetórias. Notamos na Figura 3.7C que essas distribuições têm um pico ao redor $h \approx 1$ e que elas são bastante sobrepostas.

Outra questão intrigante é se as distribuições de velocidade dos protozoários exibem uma forma funcional particular (Figura 3.9A). Nesse caso, a distribuição bidimensional de Maxwell-Boltzmann (ou, distribuição de Rayleigh)

$$P(v) = \frac{2v}{T} e^{-v^2/T} \quad (v > 0), \quad (3.17)$$

é um modelo nulo natural para a velocidade dos protozoários (T é um parâmetro da distribuição). Essa distribuição representa a velocidade de partículas de um gás bidimensional em equilíbrio termodinâmico (em uma temperatura T) e também representa a distribuição da magnitude de vetores de velocidades cujas componentes não são correlacionadas e são normalmente distribuídas (com

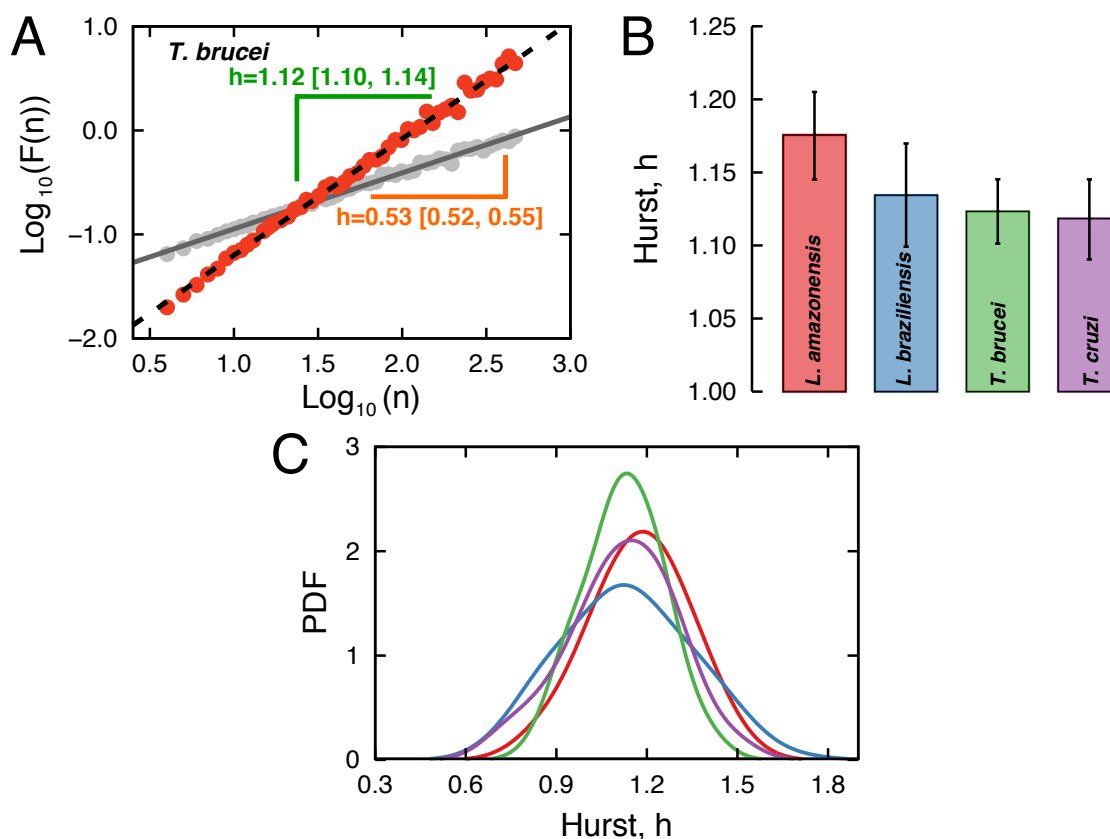


Figura 3.7: Correlações persistentes e de longo alcance nas velocidades dos protozoários. A) Análise de flutuação destendenciada (DFA) da série temporal das velocidades $v(t)$. Um comportamento típico para a função de flutuação $F(n)$ versus a escala n (em um gráfico log-log) de uma série temporal das velocidades do protozoário *T. brucei* (veja outros exemplos na Figura 3.8). Os pontos vermelhos são os valores de $F(n)$ obtidos para a série temporal original, enquanto os pontos cinzas foram obtidos para uma versão embaralhada aleatoriamente. A linha tracejada preta é um ajuste de mínimos quadrados à relação $\log_{10}[F(n)]$ versus $\log_{10}[n]$ para a série temporal original e a linha contínua cinza é o mesmo para a versão embaralhada. O valor do expoente de Hurst é de $h = 1,12$ para os dados originais e $h \approx 0,5$ para os dados embaralhados (um comportamento similar é observado para todas as séries temporais). B) Os valores médios do expoente de Hurst sobre todas as trajetórias para cada protozoário. As barras de erro são intervalos de confiança de 95% calculados via *bootstrapping*. C) Distribuições de probabilidade (PDF) dos expoentes de Hurst h para cada espécie de protozoário obtidas via método de estimativa de densidade de kernel.

média nula) [159]. Essa forma funcional descreve muito bem as velocidades de seres humanos em uma situação muito peculiar (*mosh pit*) [160] e ela também tem sido utilizada na tentativa de modelar as distribuições de velocidade de células *Hydra* em uma configuração bidimensional [61]. Desvios desse modelo também fornecem pistas sobre se as velocidades são correlacionadas ou não. No nosso caso, testamos a hipótese de Maxwell-Boltzmann (bidimensional), ajustando essa

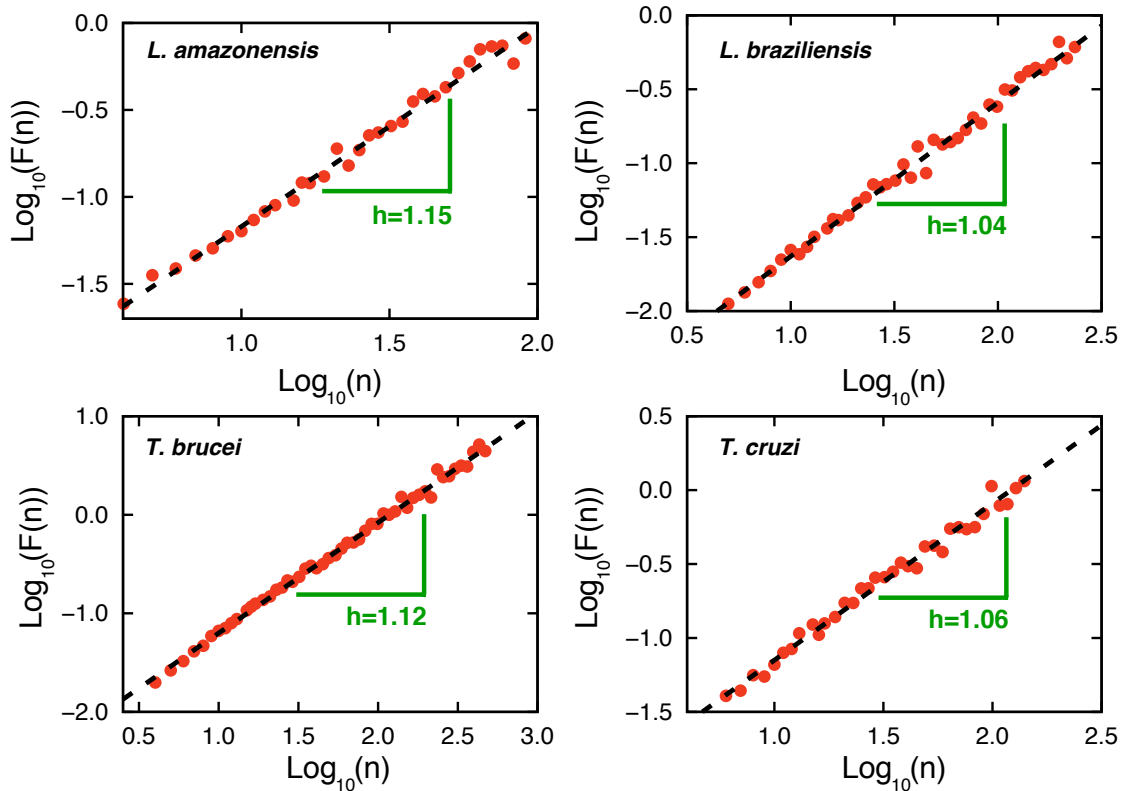


Figura 3.8: Função de flutuação. Outros exemplos típicos de análise de flutuação destendenciada (DFA) para as séries temporais de velocidade.

distribuição para cada série temporal de velocidade e verificamos a qualidade do ajuste via teste de Kolmogorov-Smirnov. Essa hipótese foi rejeitada em quase todas as séries temporais de velocidade ($\approx 99\%$), um resultado que concorda com o comportamento mais complexo observado na análise de correlação dessas séries temporais. Com o objetivo de encontrar uma melhor descrição para essas distribuições empíricas, consideramos a distribuição gama generalizada

$$P(v) = \frac{\gamma}{\beta \Gamma(\alpha)} (v/\beta)^{\alpha\gamma-1} e^{-(v/\beta)^\gamma}, \quad (v > 0) \quad (3.18)$$

em que α e γ são os parâmetros de forma e β é um parâmetro de escala. Apesar de ser uma generalização *ad hoc*, vale a pena notar que essa distribuição recupera a Maxwell-Boltzmann bidimensional (com $\alpha = 1$, $\gamma = 2$ e $\beta^2 = T$) e tem sido empregada por vários pesquisadores como um modelo de velocidades de vento [161–164]. Para os nossos dados, o teste de Kolmogorov-Smirnov não pode rejeitar a hipótese da gama generalizada em cerca de 50% das trajetórias (Figura 3.10),

uma descrição melhor quando comparada com a distribuição de Maxwell-Boltzmann bidimensional. Além disso, testamos as distribuições gama usual (Equação 3.18 com $\gamma = 1$), log-normal, Weibull e a q -exponencial. Verificamos que essas distribuições não superam a descrição da gama generalizada (Tabela 3.2). Na Figura 3.9B, mostramos os valores médios dos melhores parâmetros de ajuste α , γ e β . Esses valores são praticamente indistinguíveis entre os quatro protozoários. Uma exceção ocorre para o *T. brucei*, o qual é caracterizado por um valor significativamente menor de β . Esse fato é uma consequência direta do pequeno desvio padrão observado para as velocidades de *T. brucei* (veja Figura 3.3 D), uma vez que o desvio padrão de v calculado a partir Equação 3.18 é proporcional a β .

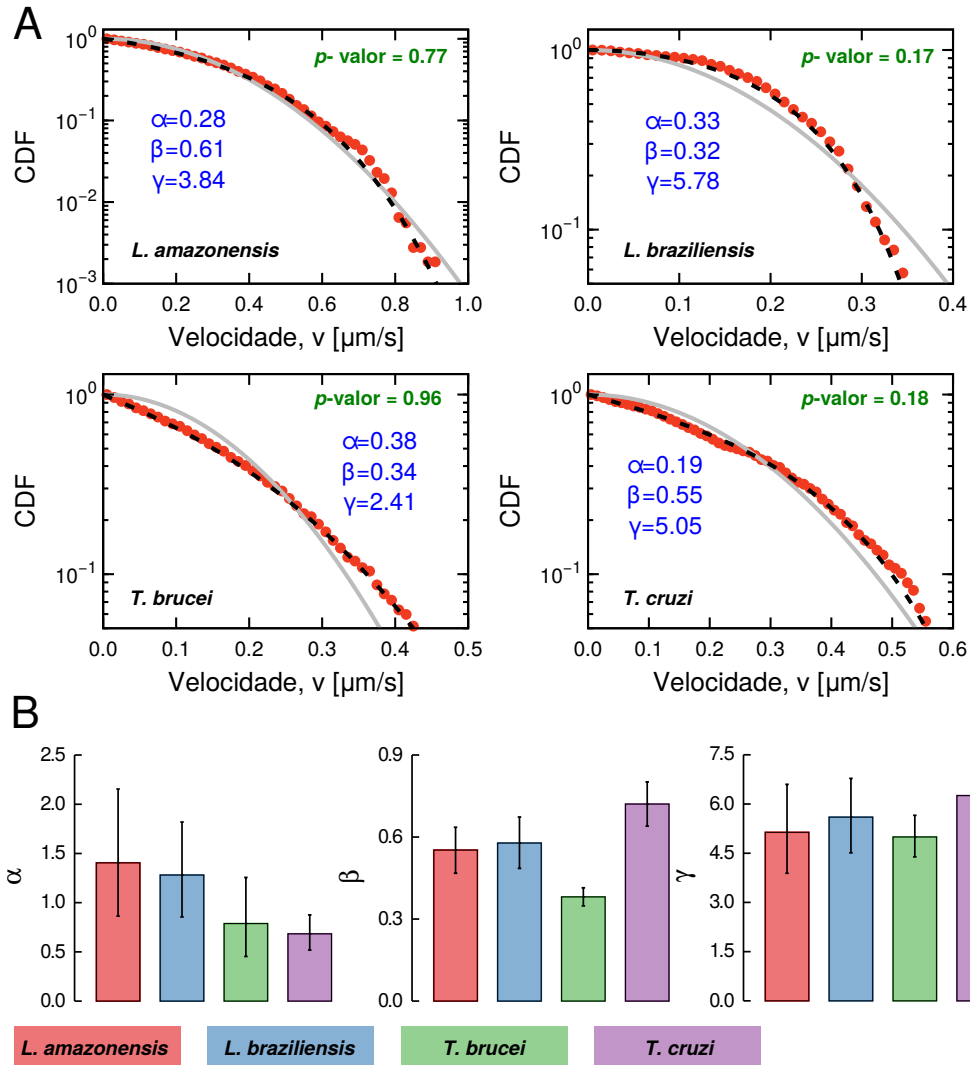


Figura 3.9: As velocidades dos protozoários não são descritas por uma distribuição de Maxwell-Boltzmann bidimensional. A) Exemplos típicos de funções de distribuição acumuladas (CDF) das velocidades $v(t)$ partir de uma única trajetória dos protozoários. As linhas contínuas em cinza são os melhores ajustes obtidos para distribuição de Maxwell-Boltzmann bidimensional (Equação 3.17), enquanto as linhas pretas tracejadas são os ajustes para a distribuição gama generalizada (Equação 3.18). Os valores dos parâmetros α , β e γ foram obtidos via método de máxima verossimilhança e são mostrados nos gráficos. Apresentamos também os p -valores do teste de Kolmogorov-Smirnov (KS) que não pode rejeitar a hipótese da distribuição gama para essas trajetórias particulares. Testamos todas as trajetórias e o teste KS não pode rejeitar essa hipótese para cerca de 50% das trajetórias (todos p -valores são mostrados na Figura 3.10). B) Mostramos os valores médios dos parâmetros de ajuste de todas as trajetórias. As barras de erro são intervalos de confiança de 95% calculados via *bootstrapping*.

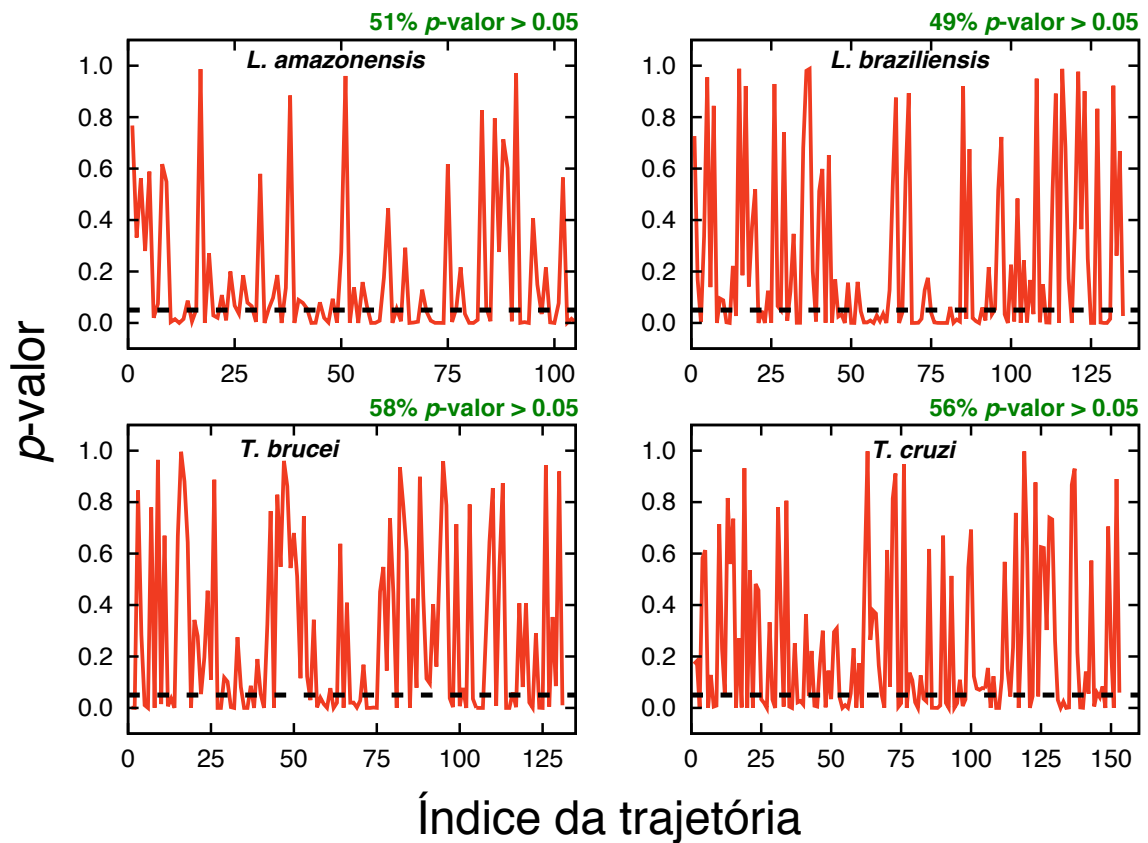


Figura 3.10: Teste de Kolmogorov-Smirnov. As linhas vermelhas mostram os p -valores do teste de Kolmogorov-Smirnov para a hipótese da gama generalizada descrever as distribuições de velocidades para todas as séries temporais de velocidades. Cerca de 50% das séries temporais de velocidade têm $p > 0,05$ (linha horizontal).

Tabela 3.2: Teste de Kolmogorov-Smirnov. Cada célula da tabela reporta o percentual de distribuições empíricas de velocidades que não foram rejeitadas pelo teste de Kolmogorov-Smirnov para uma dada distribuição (linhas) e uma dada espécie de protozoário (colunas), isto é, aquelas cujo p -valor do teste é maior que 0,05.

	<i>L. amazonensis</i>	<i>L. braziliensis</i>	<i>T. brucei</i>	<i>T. cruzi</i>
Gama generalizada	51,4	48,9	58,0	56,2
Weibull	25,7	28,9	26,7	28,1
q -exponencial	12,4	14,1	19,1	29,4
Gama	14,3	14,8	14,5	20,2
Maxwell-Boltzmann	0,9	1,5	0,8	0,6
Log-normal	0,0	0,0	0,0	0,0

3.4 Conclusões e perspectivas

Apresentamos uma descrição dos padrões de motilidade de quatro protozoários flagelados tripanosomatídeos. Ao analisarmos a evolução temporal das posições desses protozoários, identificamos que o espalhamento de suas trajetórias é superdifusivo para tempos curtos e caracterizado por distribuições de probabilidade autossimilares com caudas mais longas do que o esperado pelo movimento browniano usual. Também investigamos as velocidades desses protozoários e descobrimos que elas têm correlações de longo alcance e apresentam um forte comportamento persistente. Observamos ainda que as distribuições de velocidade não podem ser descritas por uma distribuição de Maxwell-Boltzmann bidimensional (a candidata natural no caso de um processo completamente aleatório). Em vez disso, a distribuição gama generalizada mostrou ser uma descrição melhor para as distribuições. Assim, os nossos resultados mostram que os padrões de motilidade desses protozoários são anômalos de várias maneiras e também revelam que esses quatro protozoários apresentam comportamentos semelhantes, apesar de suas diferenças morfológicas. Acreditamos que desvios em relação aos comportamentos relatados aqui podem ser utilizados como um indicador da atividade de medicamentos em testes de novas drogas.

Correlações de longo alcance e dinâmica fractal em *C. elegans*: mudanças com envelhecimento e estresse

O declínio mental e o controle motor reduzido são duas das características mais marcantes associadas ao envelhecimento e ao adoecimento [165–168]. Análises não lineares e fractais provaram ser úteis na investigação de alterações fisiológicas humanas com o envelhecimento e adoecimento [10, 169–172]. No entanto, resultados semelhantes não foram estabelecidos para organismos multicelulares menos complexos. Se a fisiologia de um organismo modelo mais simples exibisse as mesmas características, isso abriria uma nova janela de pesquisa sobre os mecanismos de controle que os organismos usam para regular os processos fisiológicos durante o envelhecimento e o estresse. Nesse capítulo, usamos uma tecnologia de rastreamento de animais recentemente introduzida para seguir simultaneamente dezenas de *Caenorhabditis elegans* durante várias horas [173] e usamos ferramentas da fisiologia fractal para avaliar quantitativamente os efeitos do envelhecimento e estresse de temperatura na motilidade desses nematóides ¹. De forma semelhante aos sinais fisiológicos humanos, a análise de escala revela correlações de longo alcance em variáveis de motilidade, propriedades fractais nas mudanças comportamentais e dinâmicas de flutuação em uma ampla gama de escalas temporais. Essas propriedades mudam como resultado de uma super-

¹Os nematóides (nemátodos) são um filo de animais cilíndricos e alongados.

posição de mecanismos de adaptação relacionados à idade e ao estresse que regulam a motilidade.

4.1 Introdução e apresentação do problema

As irregularidades do tipo fractal são uma marca registrada de sistemas fisiológicos saudáveis, tais como a frequência cardíaca e a dinâmica da marcha dos seres humanos [172]. A grande prevalência da dinâmica fractal nos processos fisiológicos refutou as teorias clássicas de controle fisiológico, que pressupunham que a saúde é mantida por meio da homeostase² estrita e que as flutuações longe da homeostase não devem ser correlacionadas. Em vez disso, sinais fisiológicos mostram padrões autossimilares em várias escalas e exibem correlações de longo alcance em suas flutuações.

Padrões fractais são difundidos no comportamento animal, tais como o tempo de movimentos específicos e padrões difusivos nos caminhos de animais que se deslocam através de um certo ambiente. Por exemplo, tem-se argumentado que os voos de Lévy são uma estratégia ótima para a exploração do ambiente na busca de alimentos, parceiros sexuais, objetos escondidos e assim por diante [174, 175]. Foram observados voos de Lévy no comportamento de forrageamento de formigas [176], albatrosses [177], macacos [178], tubarões, peixes ósseos, tartarugas marinhas e pinguins [179]. Padrões fractais foram observados também no sincronismo de comportamentos específicos, tais como a alimentação, comportamento sexual, social e vigilante em cabras espanholas [180], peixes (*Pimephales promelas*) [181], chimpanzés selvagens [182] e galinhas domésticas [182], respectivamente.

Examinamos se os comportamentos de um dos mais simples organismos-modelo multicelulares, *Caenorhabditis elegans*, é capaz de exibir dinâmicas tipo fractal. *C. elegans* é um organismo modelo proeminente com uma estrutura de corpo simples e uma linhagem de célula fixa contendo 302 neurônios de um total de 959 células somáticas. Aspectos da motilidade do *C. elegans* foram ligados a neurônios específicos [183], genes [184] e estímulos ambientais [168]. Muitas métricas comportamentais foram estudadas para o *C. elegans*, incluindo velocidade [185], frequência de ações particulares [168] e a configuração do corpo do verme ao longo do tempo [186]. Apesar de

²A homeostase é a habilidade de um organismo de manter um ambiente interno constante, um equilíbrio de condições, como a temperatura interna ou o conteúdo de um fluido, através da regulação de processos fisiológicos e ajustamentos às mudanças no ambiente externo.

ter um fundo quase isogênico³, nematóides individuais criados sob as mesmas condições podem ter um alto grau de variabilidade individual em comportamentos relacionados ao movimento [173]. Além disso, mesmo um único indivíduo *C. elegans* pode mostrar um comportamento altamente variável para períodos de tempo mais longos do que apenas alguns minutos (Figura 4.3).

4.2 Métodos e apresentação dos dados

4.2.1 *C. elegans*: cultura e experimentos de motilidade

Os experimentos descritos a seguir foram realizados no Morimoto-Lab do Departamento de Microbiologia da Northwestern University em parceria com o Amaral-Lab do Departamento de Engenharia Química e Biológica da Northwestern University.

Os métodos usados para adquirir todos os dados de motilidade para este trabalho foram previamente descritos em detalhe em Winter *et al.* [173]. Utilizamos o *Wild-type Bristol* isolado de *Caenorhabditis elegans* (N2), do Centro Genômico de *Caenorhabditis* (CGC) para todos os experimentos. Métodos padrões foram utilizados para cultivar e observar *C. elegans* [184]. Os nematóides foram sincronizados com a idade, por meio da postura de ovos e cresceram até a idade adulta a 20°C em placas de meio de crescimento de nematóides de 60 mm (*Nematode Growth Media*—NGM) semeadas com estirpe OP50 de *Escherichia coli*. Dez a quinze animais foram colocados dentro de uma armação de cobre com 2,5 × 1,5 cm de dimensões internas em uma placa NGM de 60 mm semeada com 200 μL de bactérias OP50. Todos os ensaios de motilidade foram realizados dentro de uma incubadora Percival I-36NL C8 para assegurar um ambiente de temperatura constante (Figura 4.1). Para criar um número suficiente de séries temporais de várias horas acompanhando o comportamento de animais individuais, usamos o *software* de aquisição de dados em tempo real *Multi-Worm Tracker* [187] e corrigimos os erros de imagem e de erro de identificação de vermes após a aquisição usando o *Worm Analysis for Live Detailed Observation* (WALDO) [173]. Nossa infraestrutura experimental e de *software* nos permite rastrear dezenas de animais por vez durante várias horas, mantendo a identidade de cada animal. Mostramos a trajetória de um dos vermes e algumas séries temporais extraídas de seus movimentos na Figura 4.2.

³Em um fundo isogênico os indivíduos são geneticamente semelhantes em relação a determinados pares de genes.

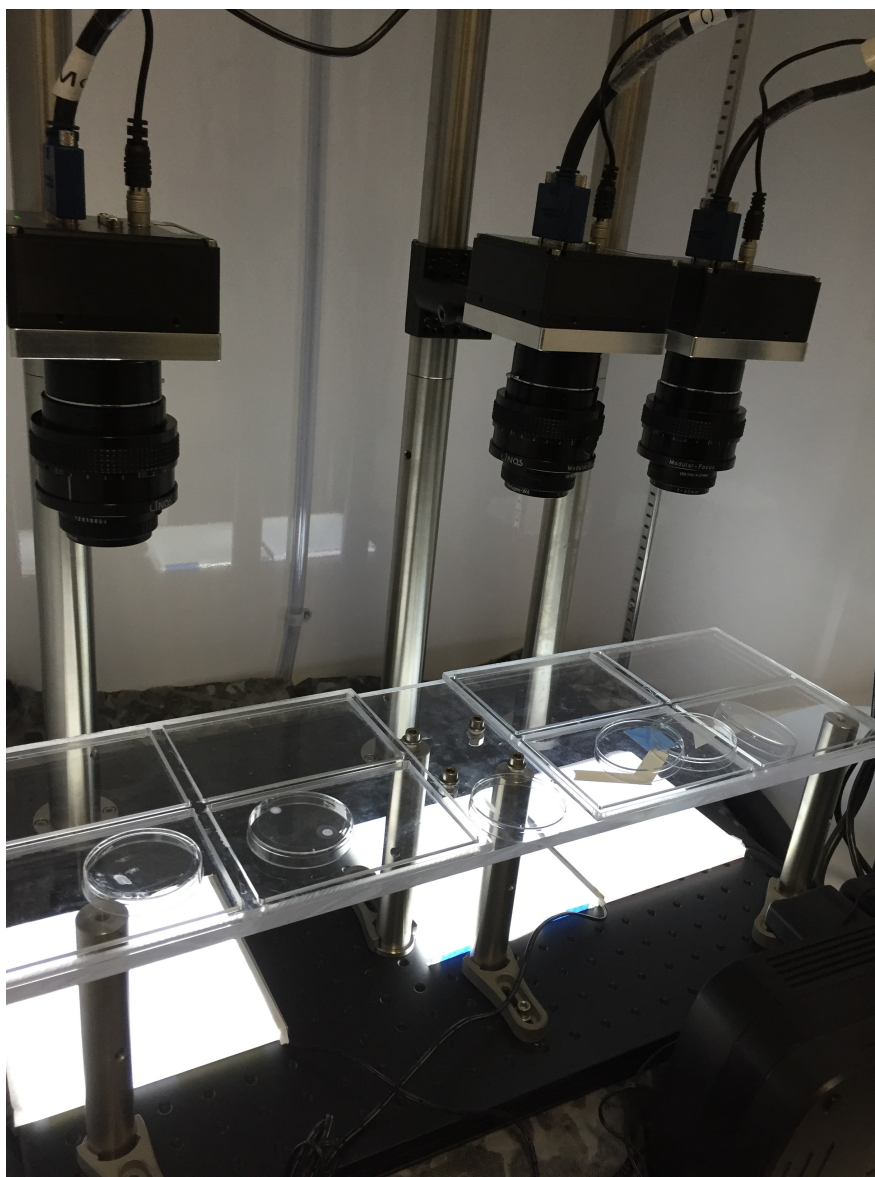


Figura 4.1: Foto da configuração experimental para os ensaios com *C. elegans*. Todas as gravações foram realizadas dentro de uma incubadora Percival I-36NL C8 para garantir um ambiente de temperatura constante. As gravações foram capturadas usando três câmeras Dalsa Falcon 4M30 (monocromáticas, 4 megapixel, 30 quadros por segundo) com lentes de aumento Rodenstock de 60 mm f/4.0 conectadas a computadores Dell Optiplex 790 (Intel i5-2400, 4 GB RAM) com cartões de câmera *National Instruments Link* executando o *Multi-Worm Tracker*, conforme descrito por Winter *et al.* [173]. As placas de Petri são posicionadas no foco dessas câmeras para a aquisição e o processamento das imagens.

Note que o problema de rastreamento de vermes desse capítulo difere um pouco do rastreamento de protozoários. Além do protozoário ser 10 vezes menor (o que exige o uso do microscópio

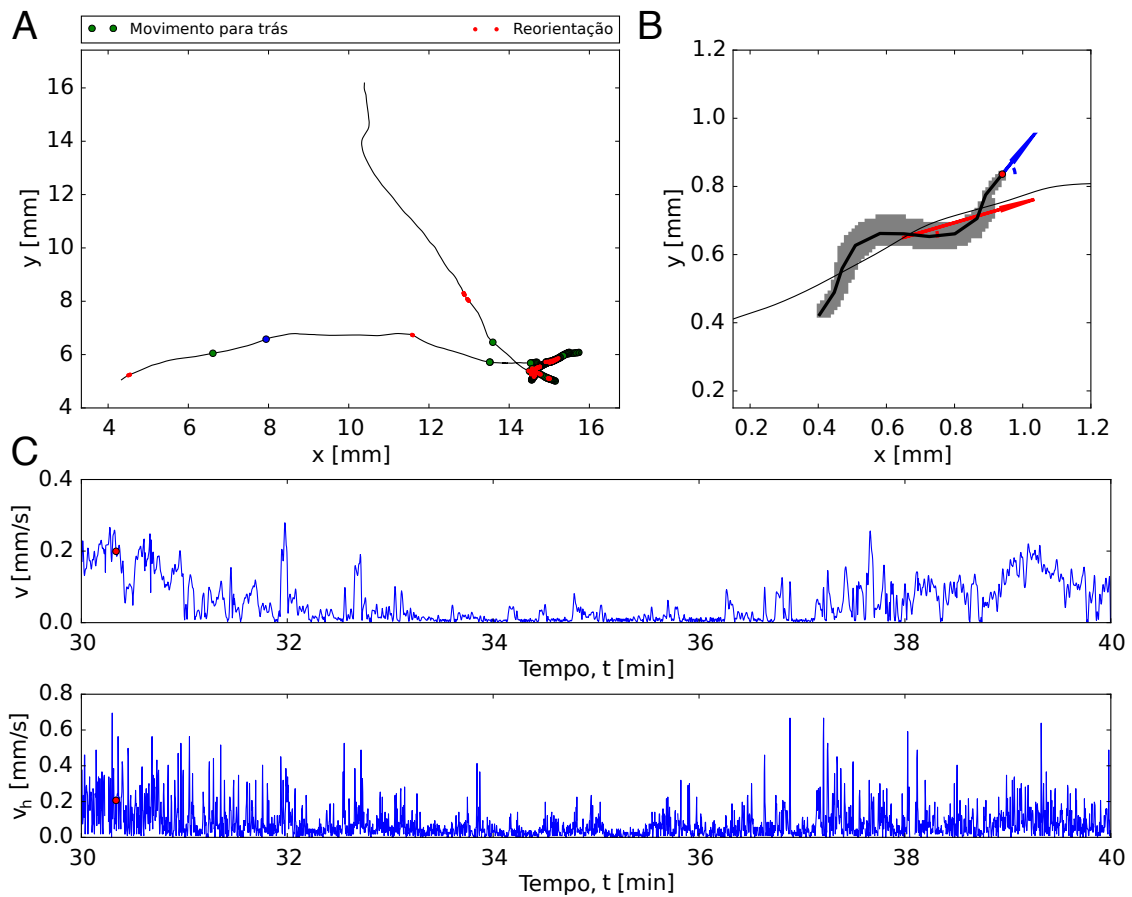


Figura 4.2: Rastreamento dos vermes. A) Dez minutos de trajetória de um verme. Pontos verdes indicam movimento para trás, pontos vermelhos representam eventos de reorientação e o ponto azul representa a posição atual. B) Forma do corpo do verme e sua trajetória ampliada. A seta vermelha representa a orientação do verme e a seta azul representa a orientação da cabeça. C) Série temporal da velocidade do centróide e velocidade da cabeça. O ponto vermelho indica a velocidade do verme para o tempo indicado em A).

e dificulta ainda mais o rastreamento dos microrganismos), as condições do experimento são diferentes: *i*) enquanto os protozoários estavam livres para sair do campo de visão da câmera, os vermes estão confinados em uma região que pode ser inteiramente visualizada; *ii*) enquanto os protozoários estavam em um meio totalmente líquido, os vermes ficam por cima de um substrato sólido na presença de estirpe OP50 de *Escherichia coli* (que serve como alimento para os vermes). O uso de correções semelhantes às propostas por Winter *et al* [173], no caso de protozoários, poderia melhorar a aquisição de dados e, portanto, melhorar os resultados sobre o movimento de protozoários. No entanto, esta tecnologia não estava disponível quando os resultados do capítulo anterior foram obtidos.

As séries temporais mostradas na Figura 4.3 exibem padrões irregulares que estão ligados à forma com que o organismo processa informações sobre seu estado interno e as sugestões químicas e mecânicas ao seu redor (Figura 4.2). Vamos nos concentrar em três tipos de análises usadas para detectar comportamentos fractais: deslocamento quadrático médio, dimensão fractal e correlação de longo alcance [7, 10, 13, 16, 60, 62, 171, 188]. Ao implementar todas essas abordagens, avaliamos se os vermes mudam sua posição, realizam ações e regulam os movimentos de uma maneira consistente com a fisiologia fractal (Figura 4.4).

4.3 Resultados e discussão

O deslocamento quadrático médio quantifica como um animal se move da sua posição atual, tal como descrevemos no capítulo anterior. Consideramos as séries temporais das posições $\vec{r}_i(t)$ para medir a dependência temporal da variância da posição radial, isto é,

$$\sigma^2(t) = \langle [\vec{r}_i(t) - \langle \vec{r}_i(t) \rangle]^2 \rangle, \quad (4.1)$$

em que $\langle \vec{r}_i(t) \rangle$ é a posição radial média de todas as trajetórias i no tempo t . Conforme já descrevemos, para um processo aleatório (movimento browniano), a variância da posição de um indivíduo aumenta linearmente com o tempo. De um modo mais geral, a variância aumenta com o tempo de uma forma lei de potência [13, 16, 60, 62],

$$\sigma^2(t) \sim t^\gamma, \quad (4.2)$$

em que $0 < \gamma < 1$ corresponde a subdifusão, $\gamma > 1$ a superdifusão e $\gamma = 1$ é o regime de difusão browniano. Para os voos de Lévy em um espaço delimitado, a variância pode ser modelada como uma lei de potência que satura para longos períodos [189]. Matematicamente, isso pode ser escrito como,

$$\sigma^2(t) = \begin{cases} Dt^\gamma & t < t_c \\ C & t > t_c \end{cases}, \quad (4.3)$$

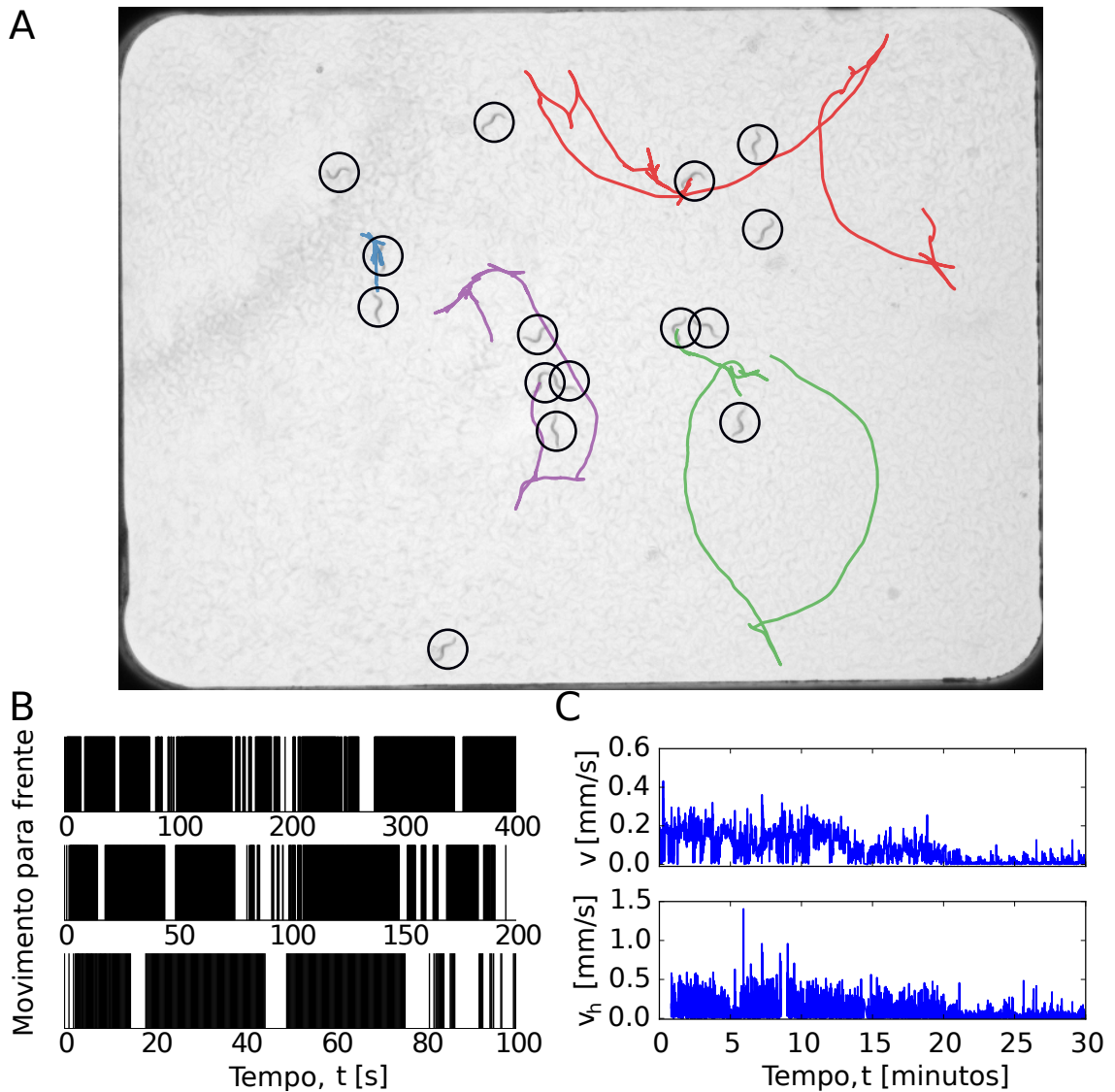


Figura 4.3: Autossimilaridade da motilidade de *C. elegans*. A) Nosso sistema experimental permite rastrear de dez a quinze vermes ao mesmo tempo e confinados dentro de um compartimento de $2,5 \text{ cm} \times 1,5 \text{ cm}$, o equivalente a uma quadra de basquete para humanos. Traçamos as trajetórias de quatro vermes durante um período de dez minutos. Observe a variabilidade nos comportamentos entre indivíduos. B) Comportamentos intermitentes, como "movimento para frente", exibem um padrão semelhante ao de conjuntos fractais. C) Séries temporais das velocidade do centróide v e das velocidade da cabeça v_h exibem flutuações em diferentes escalas de tempo.

em que γ é o expoente de difusão, D é uma constante relacionada ao coeficiente de difusão, t_c é o tempo necessário para atingir as bordas e C é constante decorrente do confinamento dentro de uma área delimitada.

Os resultados da Figura 4.4A mostram que o expoente γ é significativamente diferente de 1 para vermes de um dia de idade. Nossos dados experimentais permitem identificar o comportamento superdifusivo lei de potência ($\gamma > 1$) e o regime de saturação da variância para $t > 100$ s consistente com as previsões teóricas de voos de Lévy em uma área confinada [189].

As dimensões fractais de contagem de caixas — em inglês, *box-counting* — são utilizadas para quantificar a natureza fractal de comportamentos intermitentes. Os *C. elegans* engajam-se em vários tipos de comportamentos intermitentes, tais como movimento para frente e para trás, reorientação e enrolamento (Figura 4.3B). A série temporal desses eventos exibe uma geometria fractal, ou seja, a estrutura dos sinais parece semelhante em diferentes escalas de tempo. Para avaliar quantitativamente a fractalidade desses sinais, utilizamos o método de contagem de caixas [188] para calcular a dimensão fractal do comportamento intermitente de cada nematóide. Especificamente, contamos o número $N(s)$ de caixas de tamanho s contendo pelo menos um valor não nulo. Para um objeto fractal,

$$N(s) \sim s^{-d_f}, \quad (4.4)$$

em que d_f é a dimensão fractal de Hausdorff do objeto [188]. Para séries temporais unidimensionais de movimento (tais como a da Figura 4.3B), a dimensão fractal deve ser confinada entre $d_f = 0$, o comportamento é praticamente ausente, e $d_f = 1$, o comportamento ocorre com uma probabilidade uniforme ao longo do tempo. Na Figura 4.4B mostramos um gráfico do número de caixas $N(s)$ versus $1/s$ para um único verme. O expoente da dimensão fractal $d_f < 1$ é uma consequência da imprevisibilidade do comportamento do verme e como ele reage a sinais no ambiente, como alimentos ou a concentração de substâncias excretadas. O fato de encontrarmos $d_f < 1$ para o movimento para frente implica que este não é um comportamento regular, ou seja, existem intervalos de todos os comprimentos entre a ocorrência de eventos de movimento para frente.

Assim como no capítulo anterior, empregamos a análise de flutuação destendenciada – DFA – para caracterizar as correlações de longo alcance nas velocidades. A DFA mostra que ambas as séries temporais de velocidade do centróide e da cabeça exibem correlações de longo alcance e apresentam persistência em suas flutuações de velocidade para vermes no primeiro dia de vida

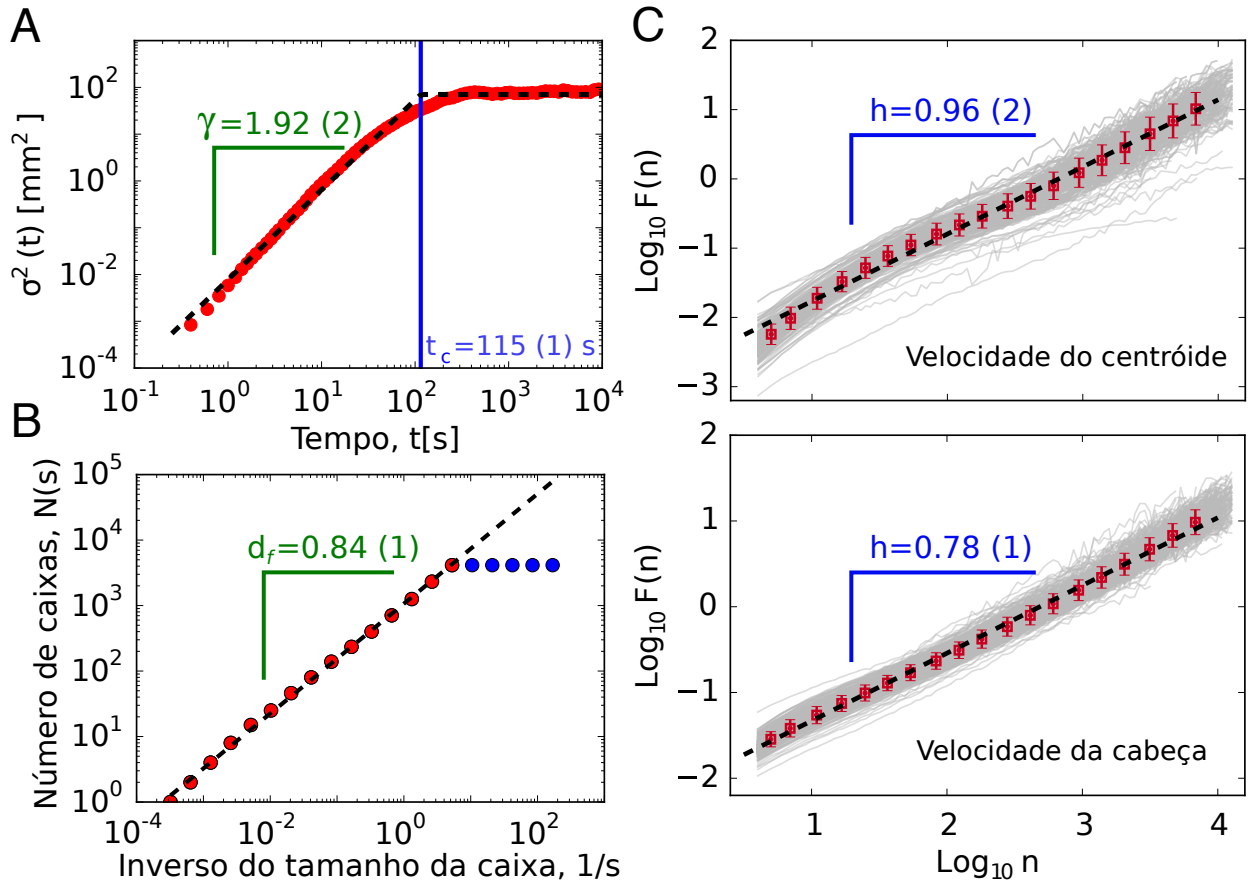


Figura 4.4: Fractalidade na motilidade de *C. elegans*. A) Determinação do comportamento difusivo dos vermes no primeiro dia de vida adulta. Pontos vermelhos são os dados e a linha tracejada preta é um ajuste para a Equação 4.3. Vale a pena notar que o valor de γ é significativamente maior do que 1, indicando que o movimento do verme não é descrito por uma difusão usual. A dimensão fractal d_f é menor do que 1, sugerindo que o movimento para frente não é um comportamento constante, ou seja, há períodos de todos os comprimentos entre os períodos consecutivos de movimento para a frente. C) Análise de correlações de longo alcance em séries temporais de velocidades. Cada linha cinza é a função de flutuação $\log_{10} F(n)$ em função da escala $\log_{10} n$ para uma série temporal das velocidades do centróide (topo) e da cabeça (abaixo). Os pontos vermelhos quadrados representam médias por janela sobre todas as curvas e as barras de erro são desvios padrão. A linha preta tracejada é um ajuste via método dos mínimos quadrados às médias. O expoente de Hurst $h > 1$ indica que as velocidades têm correlações persistentes de longo alcance. Os números entre parênteses em todas as figuras são o erro padrão no último dígito.

adulta. O comportamento da função de flutuação $\log_{10} F(n)$ versus $\log_{10} n$ para a série temporal de velocidade do centróide $v(t)$ e velocidade da cabeça $v_h(t)$ de todos os vermes são mostradas na Figura 4.4C. Uma tendência lei de potência é clara para todos os indivíduos.

Sabemos que o envelhecimento e o adoecimento podem alterar drasticamente as características fractais dos sinais da fisiologia humana. A seguir, testaremos se isso também é verdade para o *C. elegans*. Para explorar como o envelhecimento afeta a dinâmica da fisiologia do verme, repetimos a análise anterior para vermes de um a seis dias idade na fase adulta (Figura 4.5). Vale a pena notar que após o sexto dia de idade a maioria dos vermes já não apresentam movimento devido à deterioração dos músculos e à capacidade motora limitada. Usando deslocamento médio quadrático, observamos a prevalência de comportamento superdifusivo em todas as idades, embora com diferenças estatisticamente significativas entre idades (Figura 4.5A). O expoente de difusão γ é mostrado para cada idade na Figura 4.9 e as estimativas de γ obtidas via amostragem *bootstrapping* são mostradas na Figura 4.5A.

A dimensão fractal que caracteriza o movimento para a frente também muda com a idade. Na Figura 4.5B, mostramos a significância estatística das diferenças entre idades. Podemos observar que o segundo dia é ligeiramente menor do que os outros dias (excluindo os dias 5 e 6) e isso pode estar relacionado à postura de ovos, uma vez que quase 50% dos ovos são postos nesse dia [190]. De fato, uma diminuição na dimensão fractal de séries temporais comportamentais binárias foi observada durante a gravidez de cabras espanholas [180]. O expoente de Hurst para cada faixa (linhas cinzentas na Figura 4.4C) mostrou uma prevalência de correlações persistentes de longo alcance nas séries temporais de velocidade do centróide e da velocidade da cabeça em todas as idades, conforme mostrado na Figura 4.5C.

Nossos resultados mostram que as propriedades da motilidade do verme dependem de sua idade e estágio da vida. Alterações relacionadas à idade, tais como a postura de ovos, busca de parceiros ou alimentos e deterioração de órgãos e tecidos (sistema neuronal e muscular) podem estar relacionadas a mudanças no expoente de difusão γ , dimensão fractal d_f e expoente de correlação de Hurst h . A superposição desses efeitos é manifestada por pequenas (mas estaticamente significativas) alterações nos valores dos expoentes.

Assim como o envelhecimento, o estresse pode alterar as propriedades fractal de sistemas fisiológicos. Em vermes, usamos a temperatura para testar diferentes condições de estresse. A fim

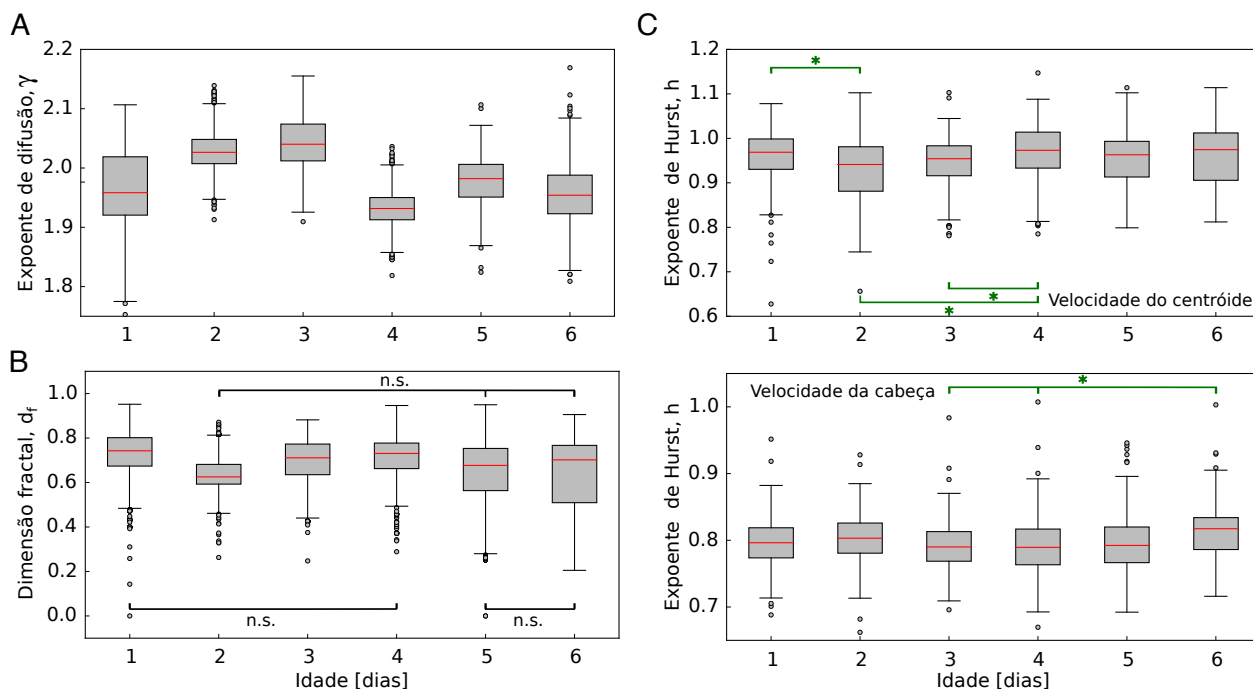


Figura 4.5: Mudanças nos expoentes de escala com o envelhecimento. A) Expoente de difusão γ como função da idade para amostras *bootstrapping*. Todas as diferenças são estatisticamente significativas para $p < 0,05/15$. B) Dimensão fractal d_f como função da idade. Todas as diferenças, exceto as marcadas com n.s., são estatisticamente significativas para $p < 0,05/15$. C) As correlações persistentes de longo alcance são prevalentes em todas as idades, apesar das mudanças no expoente de Hurst para as séries temporal da velocidade do centróide e da velocidade da cabeça. As diferenças que são estatisticamente significativas em $p < 0,05/15$ estão marcadas com *. A Figura 4.6 mostra os p -valores do teste t -Student para a hipótese nula de média idêntica com correções de Bonferroni para comparações múltiplas para cada expoente.

de fazer isso, analisamos as trajetórias de vermes crescidos em 20°C e os colocamos em uma temperatura mais fria (15°C) e em uma temperatura mais alta (25°C) e gravamos suas trajetórias. Em seguida, avaliamos como o expoente de difusão γ , a dimensão fractal d_f e o expoente de Hurst h mudam com a temperatura (Figura 4.7).

Temperaturas diferentes de 20°C parecem introduzir um comportamento mais aleatório nas trajetórias e os efeitos são maiores para temperaturas mais baixas (Figura 4.7A). Observe que a dimensão fractal d_f , para o comportamento do movimento para frente, diminui a medida que a temperatura aumenta (Figura 4.7B). O expoente de correlação h para a série de tempos de velocidades também muda com a temperatura (Figura 4.7C). Particularmente, podemos observar um aumento estatisticamente significativo no expoente de Hurst h da velocidade da cabeça à medida

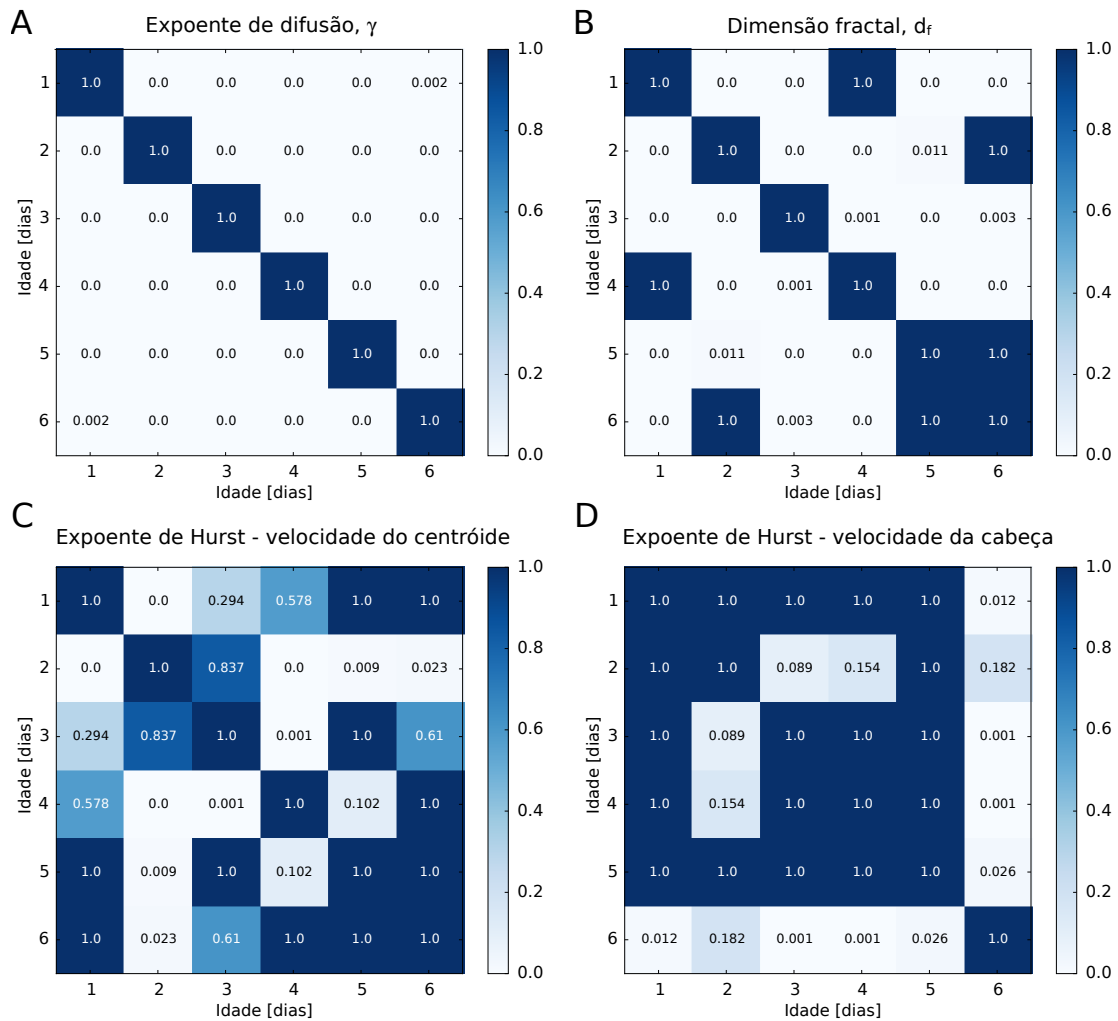


Figura 4.6: Comparação de valores de expoentes para diferentes idades. Mostramos os p -valores do teste t -Student da hipótese nula de médias de amostra idênticas. Usamos a correção de Bonferroni para contabilizar comparações múltiplas. Os painéis referem-se ao A) expoente de difusão γ , B) à dimensão fractal d_f e ao expoente de Hurst para a C) velocidade do centróide e para a D) velocidade da cabeça.

que a temperatura aumenta. A resposta ao estímulo à temperatura parece afetar mais a velocidade da cabeça, uma vez que os movimentos da cabeça estão associados à exploração/detecção durante o forrageamento.

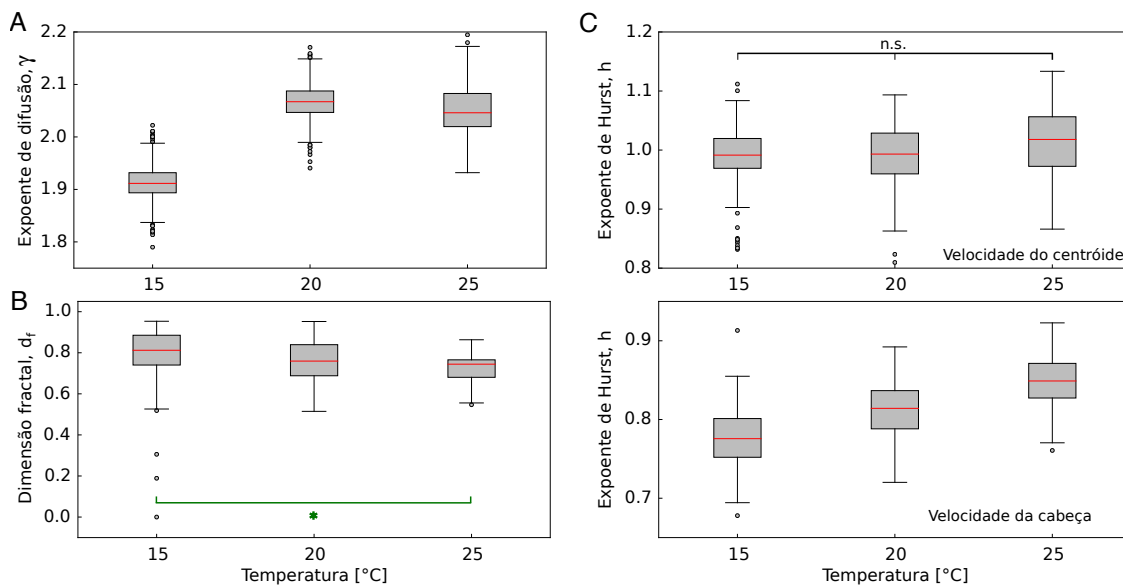


Figura 4.7: Mudanças nos expoentes de escala com o estresse. A) Expoente de difusão γ como função da temperatura para as amostras *bootstrapping*. Todas as diferenças são estatisticamente significativas para $p < 0,05/3$. B) Dimensão fractal d_f como função da temperatura. Somente a diferença entre 15°C e 25°C é estatisticamente significativa para $p < 0,05/3$. C) Correlações persistentes de longo alcance são prevalentes em todas as temperaturas apesar das mudanças no expoente de Hurst para a série temporal das velocidade do centróide e velocidade da cabeça. Todas as diferenças, exceto por aquelas marcadas com n.s., são estatisticamente significativas para $p < 0,05/3$. A Figura 4.8 mostra os p -valores do teste t -Student para a hipótese nula de média idêntica com correções de Bonferroni para comparações múltiplas para cada expoente.

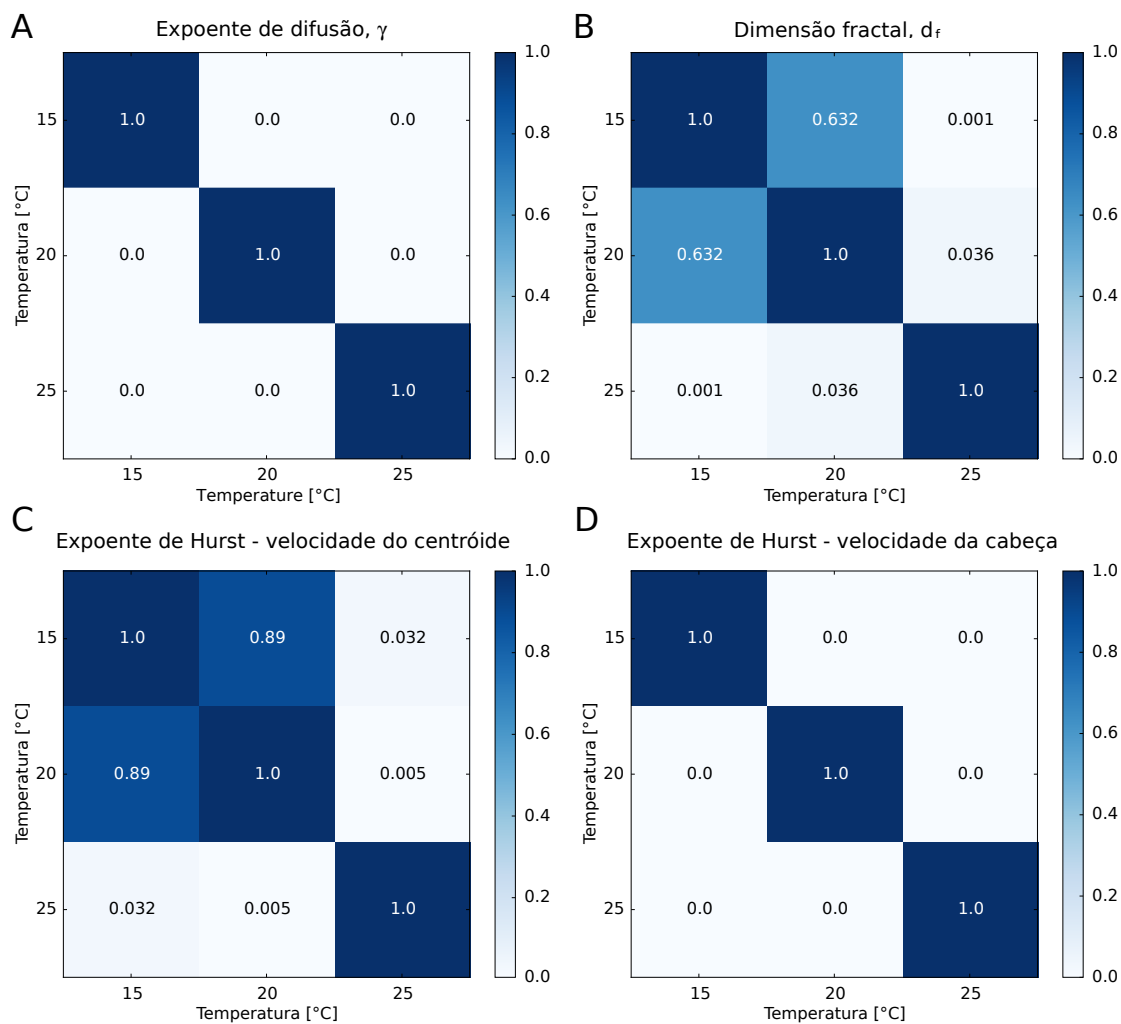


Figura 4.8: Comparação de valores de expoentes para diferentes temperaturas. Mostramos os p -valores do teste t -Student da hipótese nula de médias de amostra idênticas. Usamos a correção de Bonferroni para contabilizar comparações múltiplas. Os painéis referem-se a A) expoente de difusão γ , B) à dimensão fractal d_f e ao expoente de Hurst para a C) velocidade do centróide e para a D) velocidade da cabeça.

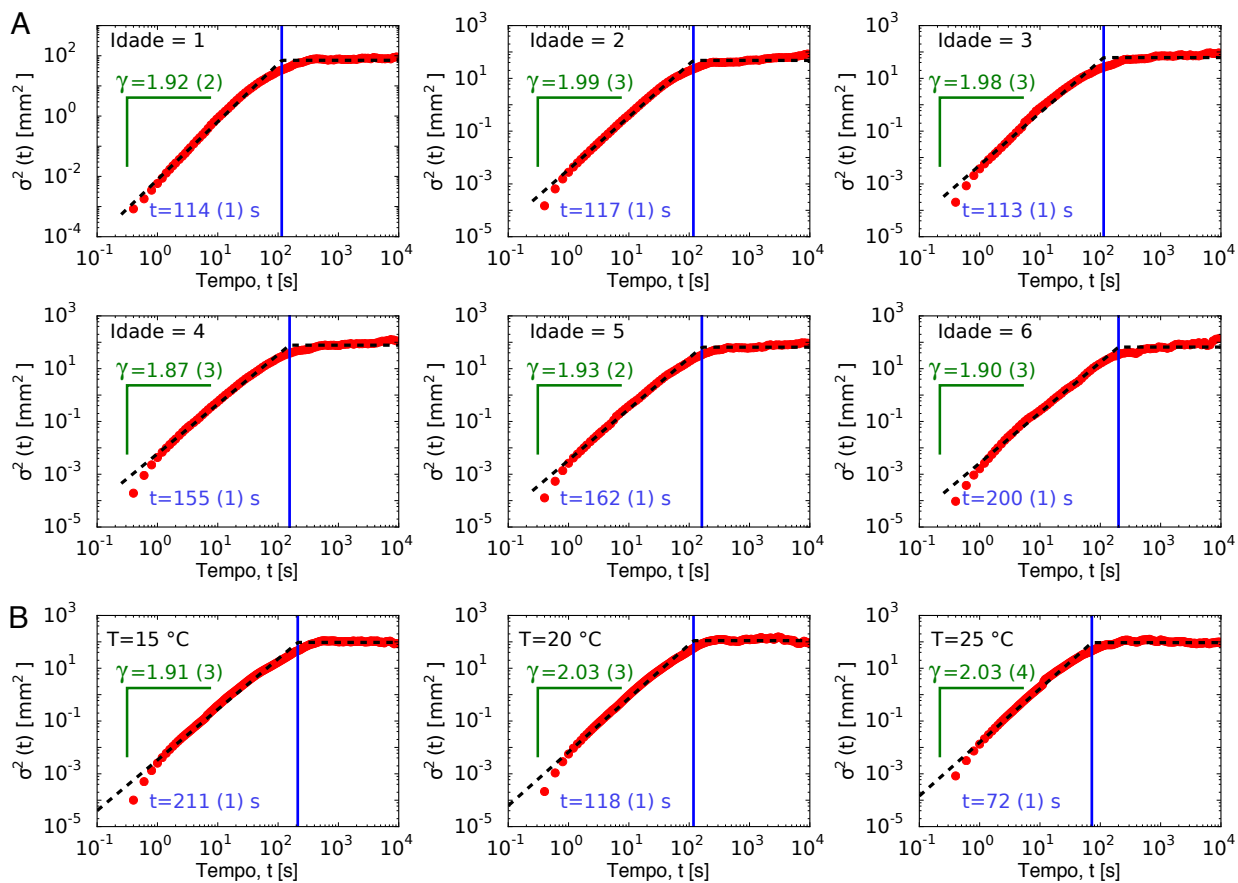


Figura 4.9: Alterações no expoente de difusão com envelhecimento e temperatura. Gráfico log-log do deslocamento quadrático médio como função do tempo para diferentes A) idades e B) temperatura. Os pontos vermelhos são os dados para as trajetórias de vermes e a linha tracejada preta é um ajuste da Equação 4.3. Os números entre parênteses em todos os gráficos são o erro padrão no último dígito.

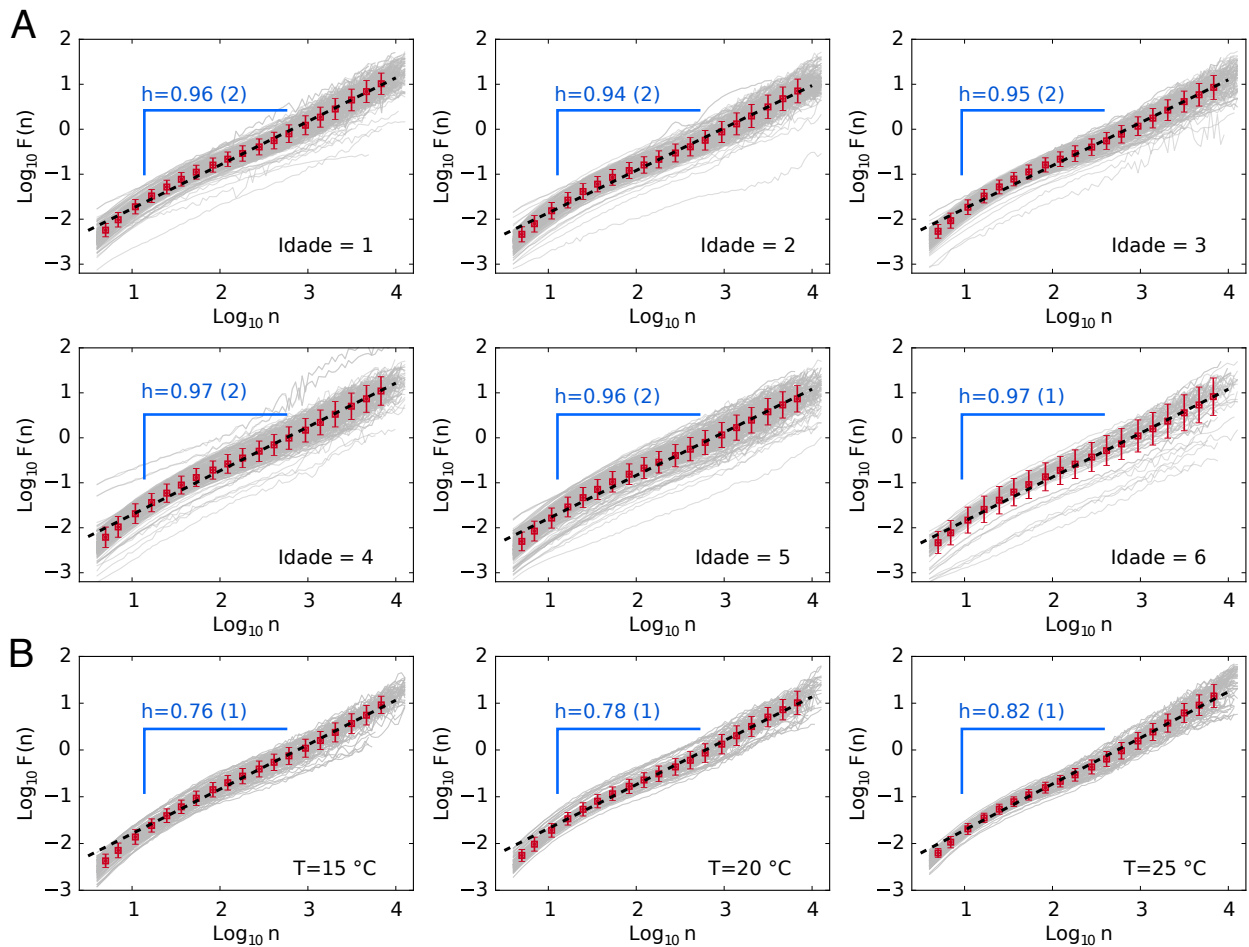


Figura 4.10: Correlação de longo alcance na velocidade do centróide dos *C. elegans*. Logaritmo da função de flutuação *versus* o logarítmico da escala de tempo para A) envelhecimento e B) temperatura. Cada linha cinza é a função de flutuação $\text{log}_{10} F(n)$ em função da escala $\text{log}_{10} n$ para uma série temporal de velocidade do centróide. Os pontos vermelhos representam médias por janela em todas as curvas e as barras de erro são um desvio padrão. A linha preta tracejada é um ajuste linear, via método dos mínimos quadrados às médias e h é o expoente Hurst. Os números entre parênteses em todos os gráficos são o erro padrão no último dígito.

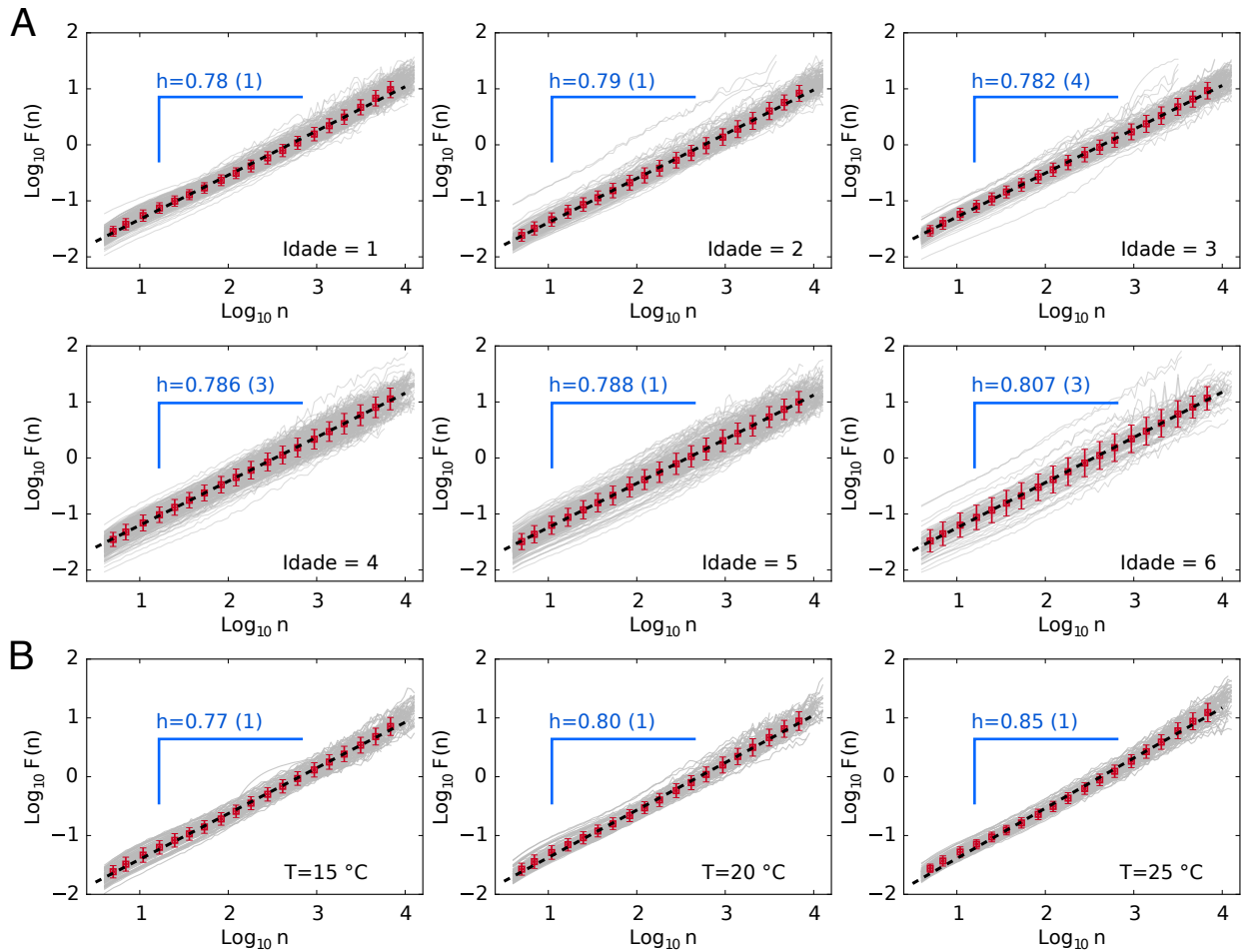


Figura 4.11: Correlação de longo alcance na velocidade da cabeça dos *C. elegans*. Logaritmo da função de flutuação *versus* o logarítmico da escala de tempo para A) envelhecimento e B) temperatura. Cada linha cinza é a função de flutuação $\text{log}_{10} F(n)$ em função da escala $\text{log}_{10} n$ para uma série temporal de velocidade da cabeça. Os pontos vermelhos representam médias por janela em todas as curvas e as barras de erro representam os desvios padrão. A linha preta tracejada é um ajuste linear, via método dos mínimos quadrados às médias e h é o expoente de Hurst. Os números entre parênteses em todos os gráficos são o erro padrão no último dígito.

4.4 Conclusões e perspectivas

Nesse capítulo, apresentamos uma descrição dos padrões de motilidades do nematóide *C. elegans*. Assim como o no capítulo anterior, a evolução temporal das posições desse microrganismo tem um comportamento superdifusivo que satura para tempos longos. Além disso, observamos que os comportamentos intermitentes do verme, como movimento para frente e para trás, exibem padrões semelhantes aos conjuntos de fractais. De fato, verificamos que a dimensão fractal dessas séries de comportamentos intermitentes é menor que um, indicando a natureza fractal desse comportamento. Também verificamos que as velocidades desses vermes são correlacionadas de longo alcance, tanto para a velocidade da cabeça, quanto para a velocidade do centróide do nematóide. Encontramos diferenças estatisticamente significativas no comportamento fractal da motilidade do *C. elegans* para diferentes idades e níveis de estresse, similarmente ao que foi relatado para fisiologia humana [10, 170–172]. Embora o uso de *C. elegans* já esteja presente no estudo de envelhecimento [166, 168, 173, 183–187, 190], nossos resultados sugerem que a similaridade com o processo de envelhecimento humano é mais profunda do que se pensava anteriormente. Acreditamos que o *C. elegans* possa ser usado para estudar como a dinâmica fractal é gerada a partir de processos regulatórios de sistemas fisiológicos e fornecer ideias sobre os processos fundamentais necessários para manter uma fisiologia saudável, apesar do envelhecimento e estresse [10, 170–172].

Investigando a interação entre mecanismos de difusão anômala via caminhada browniana fracionária em uma estrutura de pente

O modelo de pente é uma descrição simplificada para a difusão anômala sob restrições geométricas. Ele representa partículas se espalhando em um espaço bidimensional, onde os movimentos na direção x só são permitidos quando a coordenada y da partícula é zero. Nesse capítulo, propomos uma extensão para o modelo de pente via equações de Langevin impulsionadas por ruídos gaussianos fracionários (com correlações de longo alcance). Por meio da realização de simulações computacionais, mostramos que as correlações na direção y afetam o comportamento difusivo na direção x de uma forma não trivial, resultando em um cenário difuso bastante rico, caracterizado por regimes usuais, superdifusivos ou subdifusivos para o segundo momento na direção x . Mostramos que as correlações de longo alcance afetam a distribuição de probabilidade das posições das partículas na direção x , fazendo suas caudas mais longas quando o ruído na direção y é persistente e mais curtas para ruídos antipersistentes. Nosso modelo combina e permite o estudo/análise da interação entre diferentes mecanismos de difusão anômala (restrições geométricas e correlações de longo alcance) e pode ter aplicações diretas para descrever a difusão em sistemas complexos, tais

como em células vivas.

5.1 Introdução e apresentação do problema

Compreender movimentos difusivos é um problema de longa data na agenda dos físicos. Desde os trabalhos de Einstein, Smoluchowski e Langevin, sabemos que um dos padrões mais marcantes da difusão livre clássica (usual) é o crescimento linear com o tempo do segundo momento das posições das partículas. Enquanto essa característica está onipresente na natureza, muitos outros padrões têm sido observados também para o segundo momento, tanto a partir de cálculos analíticos e dados experimentais [147], como os que reportamos nos dois capítulos anteriores. Conforme já mencionamos, esses desvios do comportamento usual são chamados de difusão anômala.

Hoje em dia, os esforços têm sido focados principalmente na compreensão dos mecanismos físicos subjacentes que levam a tais desvios, já que apenas a classificação dos comportamentos de difusão de um determinado sistema não é suficiente para caracterizá-la totalmente. Nesse contexto, as duas vias comuns que geram padrões de difusão anômala são correlações de longo alcance nas posições das partículas e restrições geométricas relacionadas à complexidade estrutural do ambiente onde as partículas estão se movendo. O primeiro pode ser relacionado à efeitos de memória e um exemplo é o movimento browniano fracionário [191, 192], para o qual o segundo momento tem uma dependência lei de potência com o tempo, ou seja, $\langle x^2(t) \rangle \propto t^{2h}$. Para esse caso, o expoente de Hurst h classifica a difusão anômala em subdifusão ($0 < h < 1/2$) ou superdifusão ($h > 1/2$) [193, 194] e também recupera as difusões usual ($h = 1/2$) e balística ($h = 1$) como casos limites. A difusão anômala devido à restrições geométricas é bem exemplificada pelo modelo de pente. Nesse modelo, as partículas difundem sobre um espaço bidimensional; no entanto, os movimentos na direção x só são permitidos quando a coordenada y das posições das partículas é zero. O resultado dessa restrição simples é um movimento subdifusivo (precisamente, $\langle x^2(t) \rangle \sim t^\alpha$ com $\alpha = 1/2$) e a aparência de uma espinha (em $y = 0$) e dentes (ao longo da direção do eixo y) que lembram um pente [195–197].

Principalmente devido à sua simplicidade e capacidade de imitar aspectos da difusão de sistemas altamente desordenados, o modelo pente tem sido extensivamente estudado e estendido por meio de métodos analíticos [198–204] e também empregado como uma descrição simplificada de

fenômenos naturais, tais como a proliferação de câncer [205, 206], transporte de dendrites espinhosas [207] e difusão de átomos ultrafrios [206, 208]. Além disso, como ambientes sobrecarregados tendem a abrandar o movimento difusivo de partículas por obstrução e/ou prendendo as partículas [209], o modelo de pente também pode ser utilizado como um modelo de simulação para descrever a difusão em processos intracelulares, nos quais vários pesquisadores relataram a existência de subdifusão e quebra de ergodicidade no passeio aleatório de diferentes compostos bioquímicos [209–216, 216–219] usando rastreamento de uma única partícula [220]. No caso de processos intracelulares, os expoentes da lei de potência α descrevendo o segundo momento variam de $\alpha \approx 0,1$ a $\alpha \approx 0,9$ (calculados tanto para um *ensemble* quanto para média no tempo) e, no caso de átomos ultrafrios, α aparece como função da profundidade da estrutura e assume valores maiores do que um ($\alpha > 1$, superdifusão).

Além da escala anômala do segundo momento, o trabalho de Weiss [221] também mostrou que os processos de difusão de fluidos intracelulares podem apresentar correções de longo alcance, como no movimento browniano fracionário, uma observação empírica que não tem sido explorada no âmbito do modelo de pente. Além disso, recentemente, Yamamoto *et al.* [222] forneceram evidências (de simulação de dinâmica molecular) que a difusão anômala de moléculas de água na superfície da membrana celular surge tanto na divergência do tempo médio nas armadilhas quanto nos ruídos com correlações de longo alcance. Nesse capítulo, estendemos o modelo de pente, considerando uma equação de Langevin na qual os ruídos são correlacionados de longo alcance. Assim, nosso modelo combina dois dos principais mecanismos de difusão anômala: restrições geométricas e correlações de longo alcance, permitindo uma descrição mais completa dos movimentos difusivos complexos e também a análise da interação entre esses dois mecanismos. Por meio da realização de simulações computacionais, mostramos que as correlações na direção y afetam o comportamento difusivo na direção x de uma forma não trivial e que um cenário difusivo bastante rico emerge dessas correlações de longo alcance. Relatamos que, esse processo difusivo na direção x dessa versão generalizada do modelo de pente pode ser usual, superdifusivo ou subdifusivo, dependendo de ambas as correções nas direções x e y . Mostramos ainda que as correlações de longo alcance afetam a distribuição de probabilidade das posições das partículas na direção x : a distribuição torna-se mais leptocúrtica quando o ruído na direção y é persistente enquanto que, para o ruído antipersistente, ela torna-se menos leptocúrtica e aproxima-se da distribuição gaussiana, como um caso limite.

5.2 Generalização do modelo de pente e correlações de longo alcance

Começamos escrevendo a equação de difusão generalizada descrevendo o movimento sob a estrutura de pente [223]:

$$\frac{\partial}{\partial t} \rho(x, y; t) = \delta(y) D_x \frac{\partial^2}{\partial x^2} \rho(x, y; t) + D_y \frac{\partial^2}{\partial y^2} \rho(x, y; t). \quad (5.1)$$

Aqui $\rho(x, y; t)$ é a distribuição de probabilidade conjunta das posições das partículas (coordenadas x e y) em função do tempo t , D_x e D_y são os coeficientes de difusão nas direções x e y . Note a presença da delta de Dirac $\delta(y)$ multiplicando a derivada espacial em relação a x e, consequentemente, limitando a difusão na direção x apenas para o valor de $y = 0$. Para investigar o papel das correlações de longo alcance nesse processo difusivo, propusemos as seguintes equações de Langevin (após desconsiderar o termo de inércia e discretizá-las) acopladas

$$\begin{aligned} x(t + dt) &= x(t) + \beta_x \delta(y) \eta_x(t) \\ y(t + dt) &= y(t) + \beta_y \eta_y(t), \end{aligned} \quad (5.2)$$

em que $\delta(y)$ é aproximado por uma banda estreita de largura ϵ (onde é permitido o movimento na direção x), β_x e β_y são constantes relacionadas a D_x e D_y , $\eta_x(t)$ e $\eta_y(t)$ são ruídos gaussianos fracionários [191, 192], ambos com média zero ($\langle \eta_x(t) \rangle = 0$ e $\langle \eta_y(t) \rangle = 0$), variância unitária ($\langle \eta_x^2(t) \rangle = 1$ e $\langle \eta_y^2(t) \rangle = 1$) e funções de correlações (para $t_1 \neq t_2$) dadas por

$$\begin{aligned} \langle \eta_x(t_1) \eta_x(t_2) \rangle &\sim h_x (2h_x - 1) |t_1 - t_2|^{2h_x - 1} \text{ e} \\ \langle \eta_y(t_1) \eta_y(t_2) \rangle &\sim h_y (2h_y - 1) |t_1 - t_2|^{2h_y - 1}. \end{aligned} \quad (5.3)$$

Note que os ruídos $\eta_x(t)$ e $\eta_y(t)$ são correlacionados por uma dependência lei de potência com expoentes de escala $(2h_x - 1)$ e $(2h_y - 1)$, em que h_x e h_y são os expoentes de Hurst. Para $h_{x,y} < 1/2$, as funções de correlação têm sinais negativos e os ruídos são antipersistentes, o que significa que valores positivos são seguidos por valores negativos (ou vice-versa) com mais frequência do que por acaso; enquanto para $h_{x,y} > 1/2$, as correlações são positivas, o que significa que os valores

positivos são seguidos por valores positivos e valores negativos são seguidos por valores negativos mais frequentemente do que por acaso. Ruídos completamente persistentes ocorrem como um caso limite quando $h_{x,y} = 1$, enquanto ruídos brownianos normais correspondem ao limite de $h_{x,y} = 1/2$.

Ao executar as simulações das equações de Langevin, consideramos uma banda estreita de espessura ϵ ao longo do eixo x , no interior do qual a difusão na direção x ocorre. Essa faixa imita o efeito da delta de Dirac $\delta(y)$ (que aparece multiplicando o ruído $\eta_x(t)$ na Equação 5.2). Verificamos que o valor ϵ não tem nenhuma influência no processo difusivo, desde que ϵ e as amplitudes dos ruídos β_x e β_y sejam da mesma ordem de grandeza. Essa restrição garante que a variável y não apresente dinâmica dentro dessa faixa. Para o nossos objetivos, fixamos $\epsilon = \beta_x = \beta_y = 0,1$. Os ruídos fracionários gaussianos $\eta_x(t)$ e $\eta_y(t)$ foram gerados seguindo o método de Hosking [224]. A Figura 5.1 mostra exemplos de trajetórias simuladas para três valores diferentes de h_y mantendo $h_x = 0,5$. Por inspeção visual, notamos que, dependendo se o movimento em y é antipersistente (Figura 5.1A) ou persistente (Figura 5.1C), o comportamento de difusão na direção x muda drasticamente. A antipersistência em $y(t)$ faz com que a partícula volte mais vezes para a faixa onde o movimento de difusão na direção x é permitido. Conseqüentemente, $x(t)$ abrange um intervalo maior ao longo do eixo x quando comparado ao modelo de pente usual ($h_x = h_y = 0,5$, Figura 5.1B). Por outro lado, a persistência em $y(t)$ faz a partícula retornar com menos frequência para a faixa e, conseqüentemente, $x(t)$ abrange um pequeno intervalo ao longo do eixo x .

Antes de quantificar os efeitos das correlações de longo alcance sobre as variâncias de $x(t)$ e $y(t)$, verificamos primeiramente se as equações de Langevin (Equação 5.2) são realmente equivalentes à equação de difusão generalizada (Equação 5.1) quando os ruídos $\eta_x(t)$ e $\eta_y(t)$ não estão correlacionados ($h_x = h_y = 0,5$). Simulamos um conjunto de 10^5 posições das partículas considerando $dt = 1$ e $h_x = h_y = 0,5$. Usando esses dados, podemos avaliar a evolução temporal das variâncias

$$\begin{aligned}\sigma_x^2(t) &= \langle (x(t) - \langle x(t) \rangle)^2 \rangle \text{ e} \\ \sigma_y^2(t) &= \langle (y(t) - \langle y(t) \rangle)^2 \rangle,\end{aligned}\tag{5.4}$$

em que $\langle \dots \rangle$ denota a média de *ensemble*. As Figuras 5.2A e B mostram o comportamento de $\sigma_x^2(t)$ e $\sigma_y^2(t)$ *versus* tempo para vários valores de tempo de integração t_{\max} nas equações de Lan-

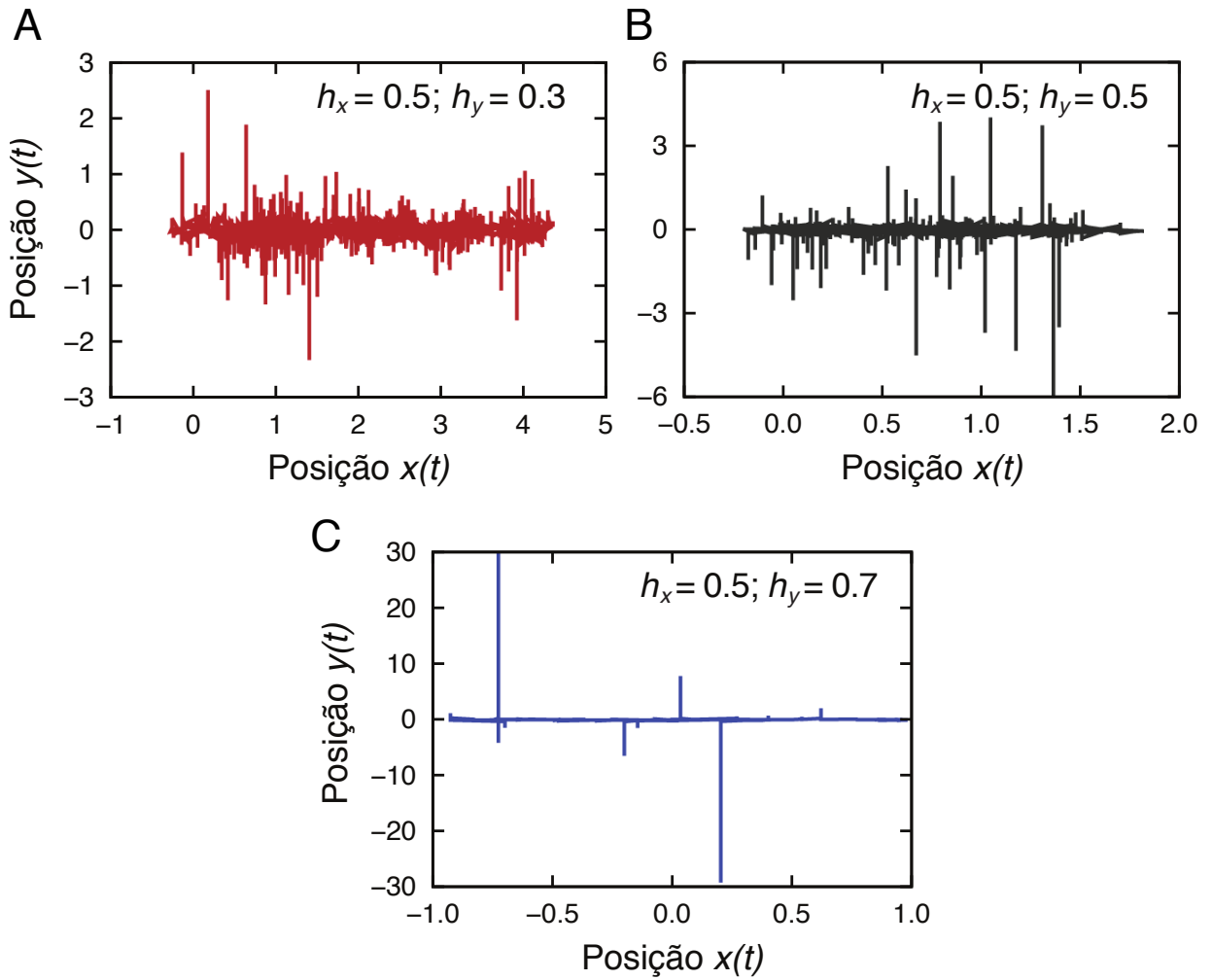


Figura 5.1: Simulação do modelo de pente via a equação de Langevin. Exemplos de trajetórias simuladas do modelo de pente generalizado descrito pela Equação 5.2 com $\epsilon = \beta_x = \beta_y = 0,1$. Mantivemos $h_x = 0,5$ e consideramos três valores do expoente de Hurst h_y : A) $h_y = 0,3$, B) $h_y = 0,5$ e C) $h_y = 0,7$. Notamos que o caminhante retorna com menos frequência para $y \approx 0$ quando o ruído é persistente na direção y ($h_y > 0,5$); enquanto que, para um ruído antipersistente ($h_y < 0,5$), o caminhante retorna mais vezes para $y \approx 0$ quando comparado ao modelo de pente usual ($h_x = h_y = 0,5$).

gevin, abrangendo valores de 2^7 à 2^{14} como indicado nesses gráficos. Em ambos os casos, notamos um efeito de tamanho finito caracterizado por um tempo de *crossover* t_x em que $\sigma_x^2(t)$ e $\sigma_y^2(t)$ mudam seus comportamentos. Apesar disso, e também descartando um regime transiente inicial ($t \lesssim 10$), observamos que os perfis de variância de $\sigma_x^2(t)$ e $\sigma_y^2(t)$ estão em boa concordância com as leis de potência previstas pelo modelo da equação de difusão generalizada do modelo de pente usual (Equação 5.1), isto é, $\sigma_x^2(t) \sim t^{\alpha_x}$ e $\sigma_y^2(t) \sim t^{\alpha_y}$ com $\alpha_x = 0,5$ e $\alpha_y = 1$ [223, 225, 226].

Também analisamos o comportamento do tempo de *crossover* t_x em função do tempo de integração máximo. As Figuras 5.2C e D mostram que t_x cresce linearmente com t_{\max} para $\sigma_x^2(t)$ e $\sigma_y^2(t)$, confirmando que os regimes lei de potência simulados prevalecem quando t_{\max} vai a infinito. Para fortalecer ainda mais a equivalência entre nossas equações de Langevin (Equa-

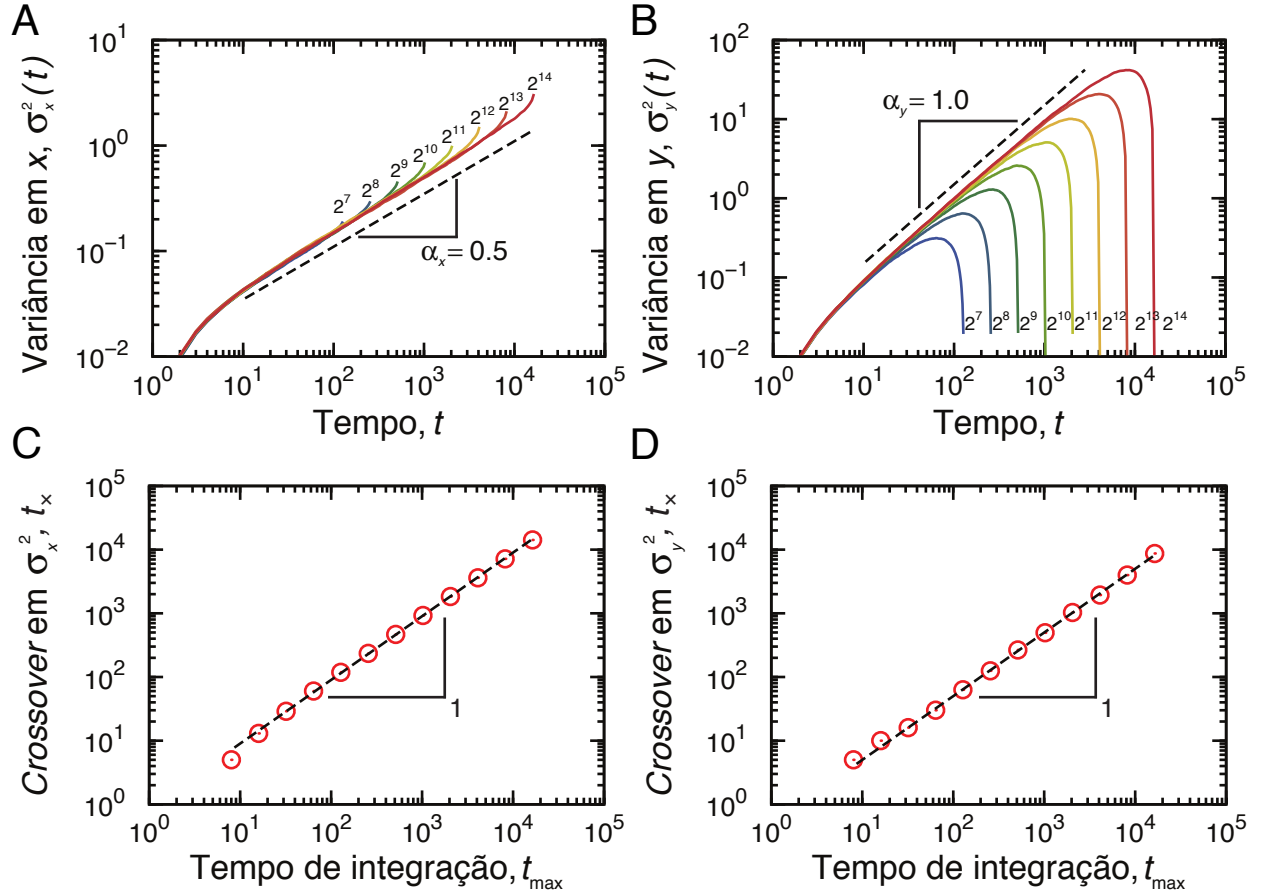


Figura 5.2: Equivalência entre as variâncias simuladas pelas equações de Langevin e os resultados analíticos obtidos a partir da equação de difusão generalizada no caso usual. Dependência temporal das variâncias A) $\sigma_x^2(t)$ na direção x B) $\sigma_y^2(t)$ na direção y para vários valores do tempo de integração máximo t_{\max} (como indicado nas figuras). As linhas tracejadas representam as previsões da equação de difusão generalizada, ou seja, $\sigma_x^2(t) \sim t^{\alpha_x}$ com $\alpha_x = 0,5$ e $\sigma_y^2(t) \sim t^{\alpha_y}$ com $\alpha_y = 1$ [223, 225, 226]. Notamos um efeito de tamanho finito em ambos os gráficos. Em C) e D), mostramos como o tempo de *crossover* t_x em que as variâncias $\sigma_x^2(t)$ e $\sigma_y^2(t)$ começam a mudar os seus comportamentos em função do tempo de integração máxima t_{\max} . Observamos que t_x cresce linearmente com t_{\max} em ambos os casos. Portanto, no limite em que t_{\max} se aproxima de infinito, os regimes lei de potência devem prevalecer.

ção 5.2) e a equação de difusão (Equação 5.1), também calculamos as funções de probabilidade marginal $P(x;t) = \int \rho(x,y;t)dy$ e $P(y;t) = \int \rho(x,y;t)dx$ para os dados simulados. As Figu-

ras 5.3A e 5.3B mostram essas distribuições para vários valores de t nas quais notamos o espalhamento dessas distribuições. Para uma melhor comparação com as previsões analíticas da Equação 5.1, também calculamos as distribuições marginais considerando as posições normalizadas $\xi_x = [x(t) - \langle x(t) \rangle] / \sigma_x(t)$ e $\xi_y = [y(t) - \langle y(t) \rangle] / \sigma_y(t)$. As Figuras 5.3C e 5.3D revelam um bom colapso das distribuições empíricas e também um bom acordo com as distribuições analíticas (linhas tracejadas) previstas pela Equação 5.1, a qual é uma gaussiana padrão para $P(\xi_y)$ e pode ser expressa em termos da função de Fox H para $P(x; t)$ [223, 225, 226] ou pelo seguinte somatório [194, 218]

$$P(\xi_x) = \frac{1}{\pi^{1/4}} \sum_{n=0}^{\infty} \frac{(-1)^n}{n! \Gamma[1 - (n+1)/4]} \left(\frac{\xi_x^2}{\sqrt{\pi}/4} \right)^{n/2}. \quad (5.5)$$

Assim, para o caso em que não há correlações presentes nos ruídos $\eta_x(t)$ e $\eta_y(t)$, os resultados mostrados na Figura 5.3 confirmam a equivalência entre nossas simulações das equações de Langevin e a equação de difusão para o modelo de pente.

Convencidos da equivalência no caso usual, abordamos agora o papel das correlações de longo alcance sobre as propriedades de difusão de um caminhante aleatório descrito por nossas equações de Langevin (Equação 5.2). Em primeiro lugar, investigamos como os efeitos de memória na direção y modificam a difusão nas direções x e y . A fim de fazer isso, construímos um conjunto de 10^5 posições das partículas considerando $dt = 1$ e $h_x = 0,5$ para cada um dos dezenove valores de h_y homogeneamente distribuídos entre 0,05 e 0,95. Para cada valor de h_y , podemos avaliar a evolução temporal das variâncias $\sigma_x^2(t)$ e $\sigma_y^2(t)$. As Figuras 5.4A e 5.4B mostram os perfis de variância nas direções x e y , respectivamente. Nesses gráficos, as linhas tracejadas representam o comportamento para o modelo de pente usual, isto é, quando não há correlações presentes. No caso de $\sigma_x^2(t)$, podemos constatar que quanto mais o ruído em y é antipersistente ($h_y < 0,5$), maior é o expoente da relação lei de potência α_x descrevendo o comportamento de $\sigma_x^2(t)$; em contraste, o ruído persistente em y retarda a difusão e diminui o expoente da relação lei de potência α_x . Os casos extremos em que $h_y = 0,05$ e $h_y = 0,95$ são caracterizados por $\alpha_x \approx 0,95$ e $\alpha_x \approx 0,19$, respectivamente. Assim, como notado pela primeira vez na Figura 5.1, o ruído antipersistente em y obriga o caminhante voltar com mais frequência para a faixa onde ele pode mover-se na direção x e, conseqüentemente, a difusão é mais rápida quando comparada ao caso do modelo de

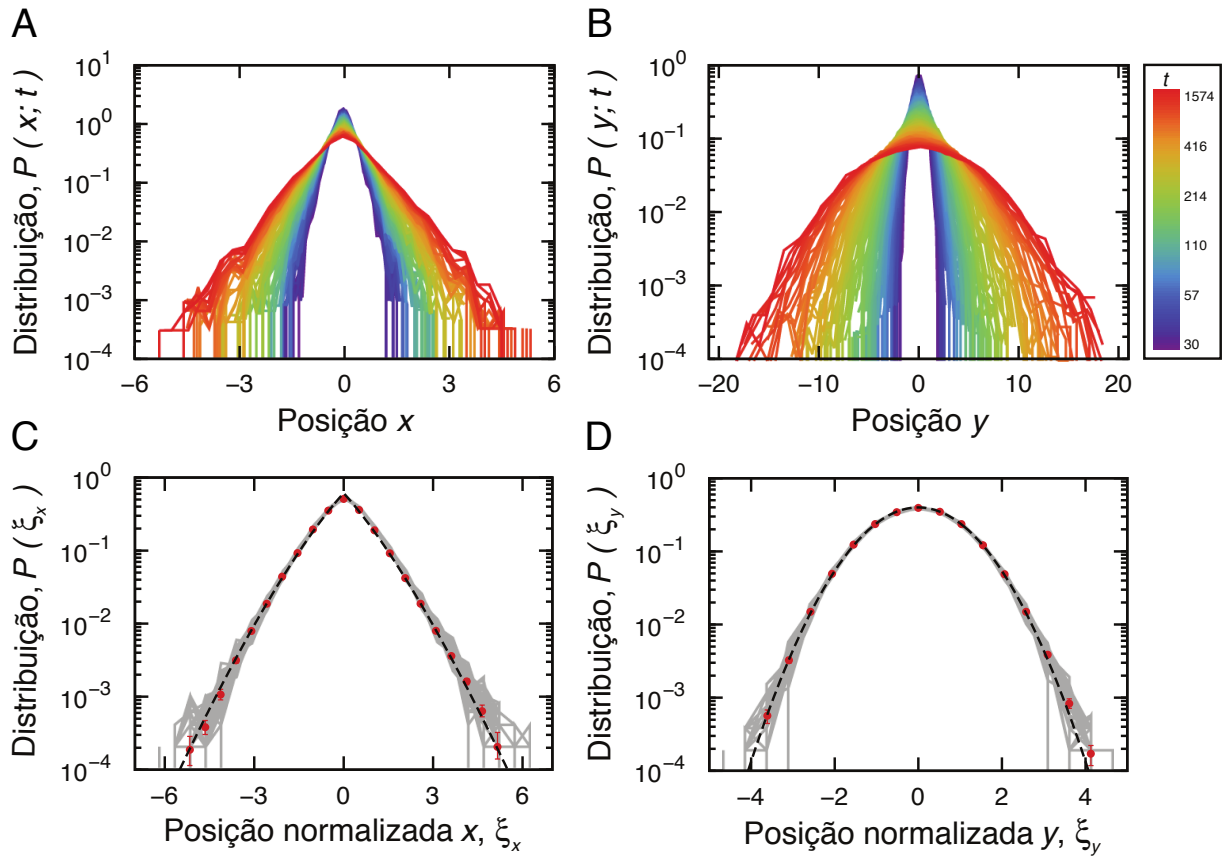


Figura 5.3: Equivalência entre as distribuições marginais $P(x; t)$ e $P(y; t)$ simuladas por meio das equações de Langevin e os resultados analíticos obtidos a partir da equação de difusão generalizada no caso usual. Distribuições marginais A) $P(x; t)$ e B) $P(y; t)$ para vários valores de t , como indicado pelo código de cores. Notamos o alargamento dessas distribuições e a normalidade de $P(y; t)$. Distribuições marginais das posições normalizadas C) $P(\xi_x)$ e D) $P(\xi_y)$ para os mesmos valores de t das figuras anteriores. As linhas cinzas são as distribuições marginais para cada valor de t e os círculos vermelhos são os valores médios por janela sobre todas as distribuições (as barras de erro são intervalos de 95% de confiança obtidos via *bootstrapping*). As linhas tracejadas são as previsões analíticas da Equação (5.1). Especificamente, em C), a linha tracejada representa a distribuição da Equação 5.5 e em D) é uma gaussiana padrão $P(\xi_y) = (1/\sqrt{2\pi}) \exp(-\xi_y^2/2)$. Além da boa concordância (visual) entre as simulações e expressões analíticas, descobrimos que os p -valores do teste de Kolmogorov-Smirnov (que testa a igualdade das distribuições empíricas e analíticas) são todas maiores do que 0,05, indicando que a hipótese das distribuições analíticas descreverem nossos dados não pode ser rejeitada em um nível de confiança de 95%.

penete usual ($h_y = 0,5$). Por outro lado, o ruído persistente em y impulsiona o caminhante longe da faixa e, conseqüentemente, a difusão desacelera em x . No caso de $\sigma_y^2(t)$, o comportamento é simples: o ruído antipersistente em y retarda a difusão, enquanto o ruído persistente a torna mais rápida (veja a Figura 5.4B). Para quantificar o papel das correlações de longo alcance em y

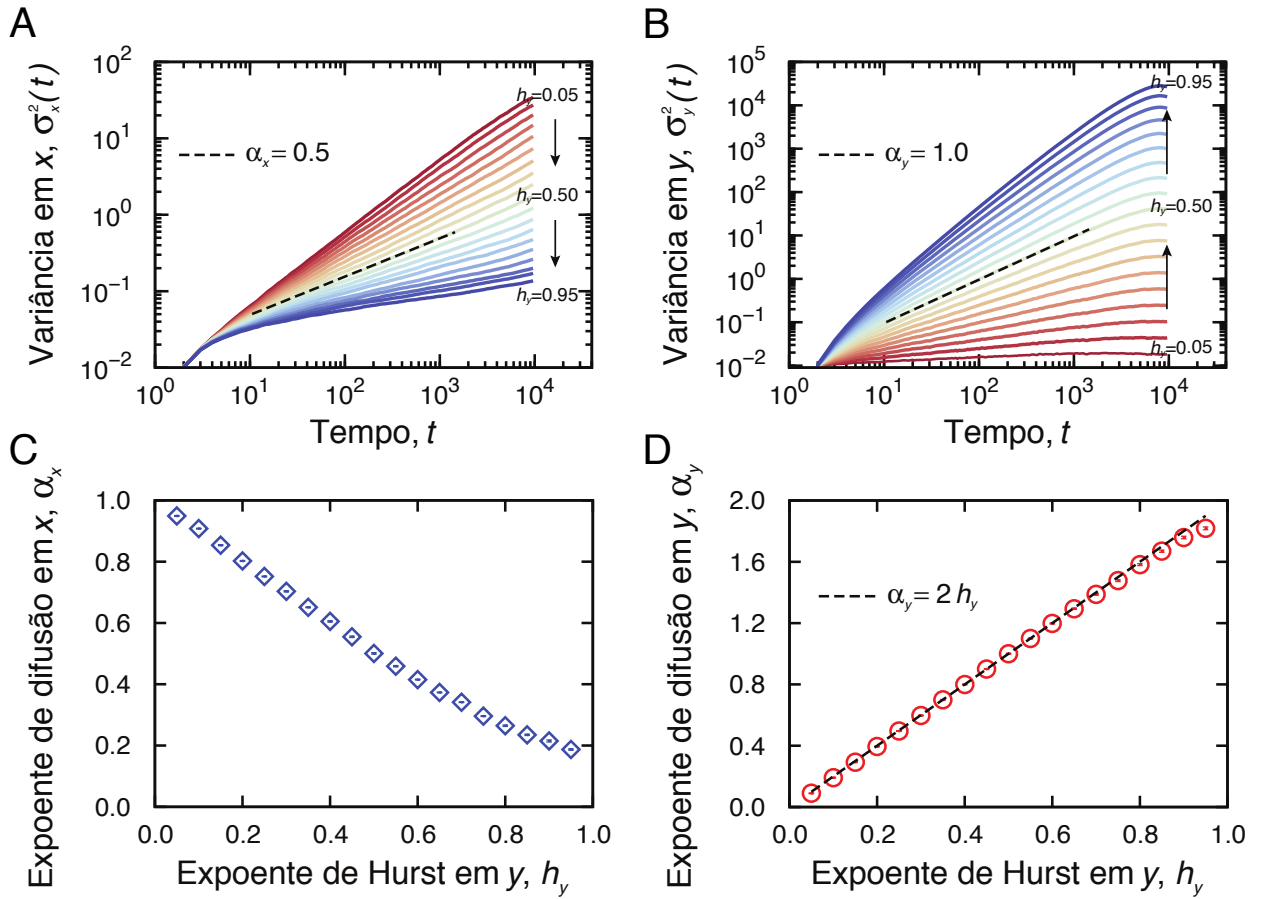


Figura 5.4: O papel do ruído correlacionado de longo alcance na direção y sobre as variâncias $\sigma_x^2(t)$ e $\sigma_y^2(t)$. Perfil de A) $\sigma_x^2(t)$ e B) $\sigma_y^2(t)$ para dezenove valores de h_y homogeneamente distribuídos entre 0,05 e 0,95. Em ambos os casos, as linhas tracejadas representam o comportamento do modelo de pente usual: $\sigma_x^2(t) \sim t^{\alpha_x}$ com $\alpha_x = 0,5$ e $\sigma_y^2(t) \sim t^{\alpha_y}$ com $\alpha_y = 1$. Aqui, omitimos o efeito de tamanho finito nas variâncias para melhor visualização do comportamento lei de potência. Para $\sigma_x^2(t)$, observamos que a difusão é acelerada (quando comparada ao modelo de pente usual) para um ruído antipersistente em y ($h_y < 0,5$) e que é mais lenta quando o ruído em y é persistente. No caso de $\sigma_y^2(t)$, o ruído antipersistente retarda a difusão, enquanto o ruído persistente a acelera. Em C) e D), mostramos a dependência de α_x e α_y em h_y . Os valores de α_x e α_y foram obtidos via mínimos quadrados ajustando um modelo linear para as relações em escala log-log entre as variâncias e o tempo t . As barras de erro (que são muito pequenas) são os erros padrão de α_x e α_y . Ao ajustar os dados, selecionamos apenas os valores de $\sigma_x^2(t)$ e $\sigma_y^2(t)$ dentro do intervalo de $30 \leq t \leq 1600$, a fim de evitar o regime transiente e o efeito de tamanho finito. Notamos um decaimento monótono, mas não linear de α_x em função de h_y e um comportamento linear de α_y versus h_y descrito por $\alpha_y = 2h_y$ (linha tracejada).

em $\sigma_x^2(t)$ e $\sigma_y^2(t)$, calculamos a dependência dos expoentes das relações leis de potência α_x e α_y em h_y . Na Figura 5.4C, observamos um decaimento monótono, mas não linear de α_x quando h_y

aumenta, enquanto que a Figura 5.4D confirma o comportamento simples de α_y como função de h_y ($\alpha_y = 2h_y$).

Outra questão interessante é se as combinações de ruídos correlacionados de longo alcance nas direções x e y afetam os comportamentos de difusão de uma forma não trivial. Para responder a essa pergunta, construímos um conjunto de 10^5 posições das partículas para cada possível par de h_x e h_y , homoganeamente distribuído entre 0,05 e 0,95, com tamanho de passo de 0,05. Para cada combinação de h_x e h_y , calculamos o perfil de $\sigma_x^2(t)$ e $\sigma_y^2(t)$ e os expoentes lei de potência α_x e α_y que descrevem a principal tendência dessas curvas (usando o mesmo processo da Figura 5.4). A Figura 5.5A mostra um gráfico de α_x como função de h_x e h_y e a Figura 5.5B representa o mesmo para α_y . Para a Figura 5.5A, notamos um cenário difusivo bastante rico, em que, dependendo dos valores de h_x e h_y , podemos ter subdifusão mais lenta ($\alpha_x < 0,5$) ou mais rápida ($0,5 < \alpha_x < 1$) do que o modelo de pente usual, a difusão usual ($\alpha_x = 1$) e até mesmo superdifusão ($\alpha_x > 1$). Observamos ainda que α_x não pode ser escrito como uma combinação linear de h_x e h_y (que seria representado por linhas retas nessas figuras) e, assim, as correlações em x e y apresentam uma interação não trivial no comportamento de α_x . Por outro lado, o caso da Figura 5.5B é bastante simples, porque as correlações em x não afetam o comportamento difusivo na direção y , isto se torna evidente observando que as curvas de nível de α_y são linhas retas horizontais.

É importante investigar também se as correlações de longo alcance afetam os perfis das distribuições das posições das partículas. A fim de fazer isso, primeiramente calculamos as distribuições marginais $P(\xi_x)$ e $P(\xi_y)$ para diferentes valores de h_y (0,05, 0,50 e 0,95) e mantendo $h_x = 0,5$. As Figuras 5.6A e 5.6B mostram essas distribuições. Notamos que as distribuições $P(\xi_x)$ exibem caudas mais longas do que aquelas do modelo de pente usual (região escura sombreada) quando o ruído na direção y é persistente ($h_y = 0,95$), enquanto que, para ruído antipersistente ($h_y = 0,05$), as caudas são mais curtas do que as do modelo de pente usual e estão próximas da distribuição gaussiana (região sombreada mais clara). As distribuições $P(\xi_y)$ das posições normalizadas ξ_y não são afetadas pelos ruídos correlacionados de longo alcance na direção y . Portanto, confirmamos que as correlações de longo alcance em y tem influência sobre os perfis de $P(\xi_x)$. Para uma caracterização melhor e quantitativa do papel do ruído correlacionado $\eta_y(t)$ na forma de $P(\xi_x)$, avaliamos o coeficiente de curtose $\kappa_x = \langle \xi_x^4 \rangle / \langle \xi_x^2 \rangle^2$ em função de h_x e h_y . Para $\kappa_x > 3$ a distribuição é leptocúrtica, enquanto para $\kappa_x < 3$ a distribuição é platicúrtica, para $\kappa_x = 3$ temos uma distribuição gaussiana e

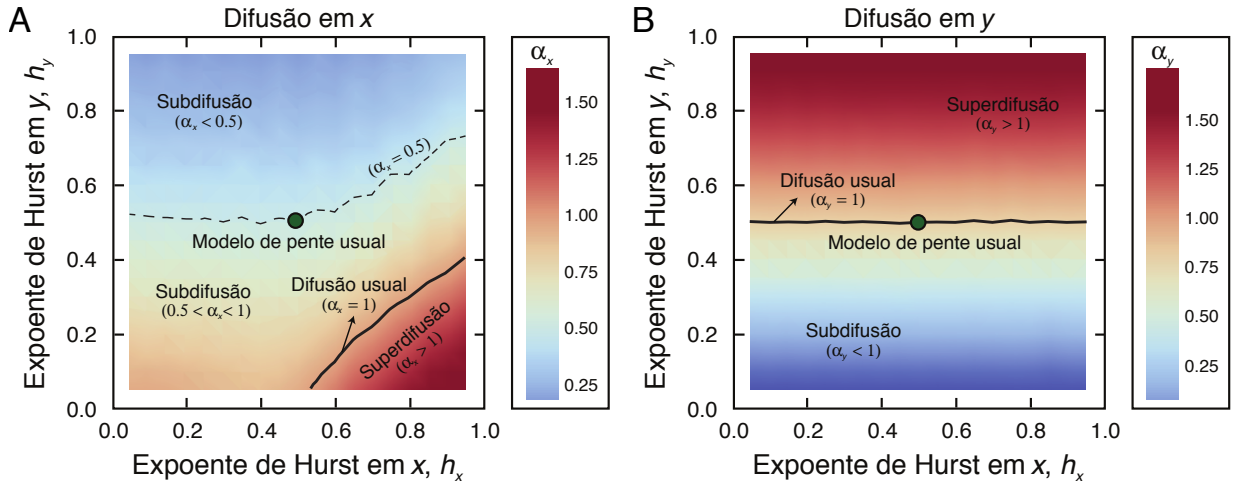


Figura 5.5: Ação combinada entre os ruídos correlacionados de longo alcance nas direções x e y sobre as variâncias $\sigma_x^2(t)$ e $\sigma_y^2(t)$. Gráficos de A) α_x e B) α_y em função de h_x e h_y . Notamos um cenário difusivo rico nas direções x onde podemos ter subdifusão mais lenta ($\alpha_x < 0,5$) ou mais rápida ($0,5 < \alpha_x < 1$) do que o modelo de pente usual, difusão usual ($\alpha_x = 1$) e até mesmo superdifusão ($\alpha_x > 1$). Observamos ainda que o expoente da lei de potência α_x não pode ser expresso como uma combinação linear de h_x e h_y , sugerindo um acoplamento mais complexo entre os expoentes de Hurst e difusão na direção x .

para $\kappa_x \approx 4,71$ recuperamos o modelo de pente usual (distribuição da Equação 5.5). A Figura 5.6C mostra um gráfico de κ_x como função de h_x e h_y , no qual observamos que os intervalos de curtose de aproximadamente 3 ($\kappa_x = 3,04 \pm 0,04$ para $h_x = h_y = 0,05$) até quase 13 ($\kappa_x = 12,55 \pm 0,87$ para $h_x = h_y = 0,95$). Vale a pena notar que, semelhante ao caso da variância (Figura 5.5A), a curtose κ_x não é uma combinação trivial de h_x e h_y .

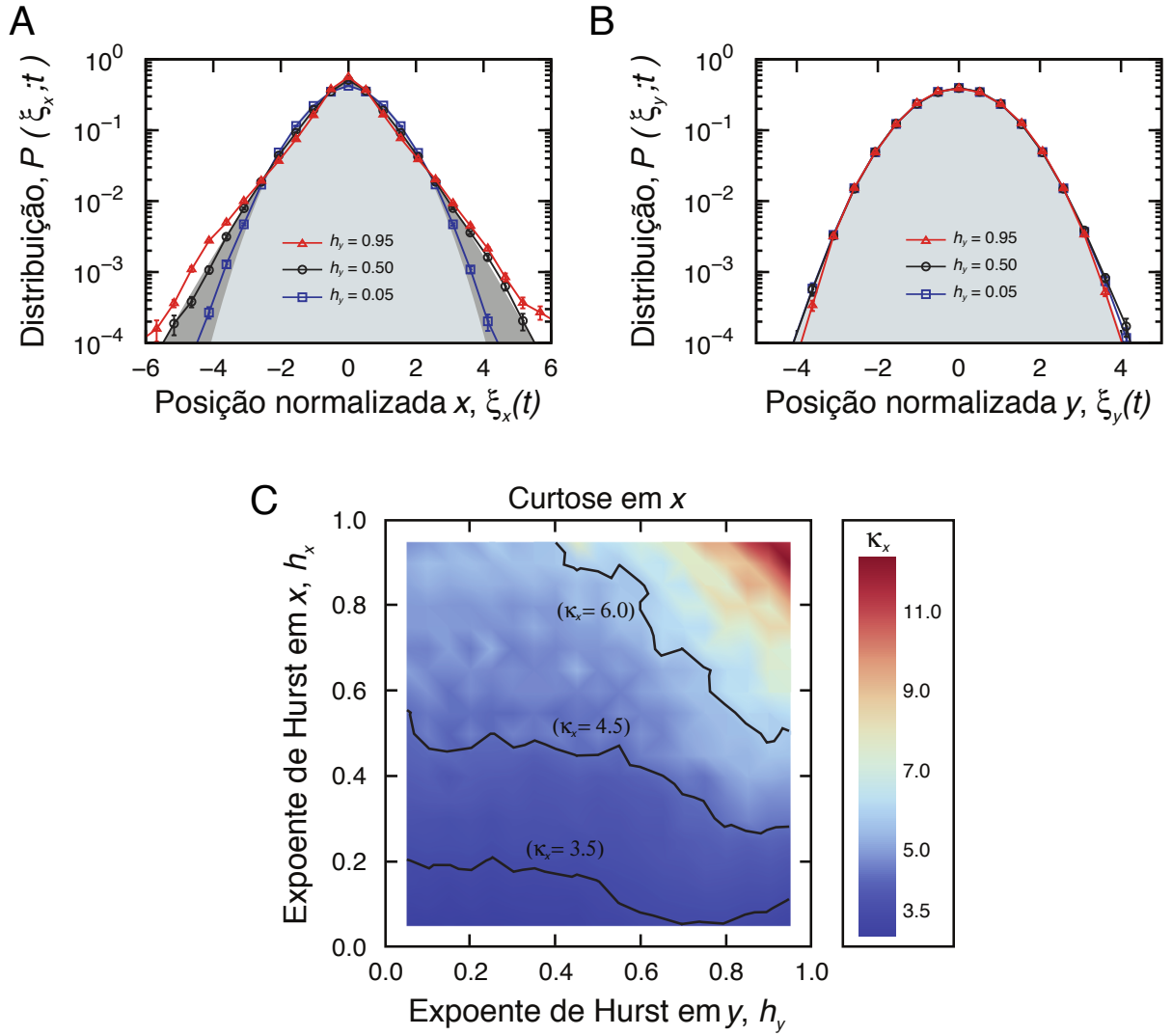


Figura 5.6: Efeito dos ruídos correlacionados de longo alcance sobre as distribuições marginais. Distribuições marginais das posições normalizadas A) $P(\xi_x)$ e B) $P(\xi_y)$ para alguns valores de h_y (indicados nos gráficos) e $h_x = 0,5$. Em ambos os gráficos, as regiões sombreadas mais claras representam uma gaussiana padrão, enquanto que a região escura sombreada em A) é o perfil da distribuição marginal para ξ_x do modelo de pente usual (Equação 5.5). Para $P(\xi_x)$, notamos que as caudas das distribuição decaem mais lentamente do que aquelas do modelo de pente usual quando o ruído na direção y é persistente ($h_y = 0,95$); enquanto que para o ruído antipersistente ($h_y = 0,05$), as caudas são mais curtas e aproximam-se da distribuição gaussiana. No caso $P(\xi_y)$, as correlações de longo alcance em y praticamente não tem efeito sobre o perfil dessas distribuições. C) Gráfico de nível dos coeficientes de curtose $\kappa_x = \langle \xi_x^4 \rangle / \langle \xi_x^2 \rangle^2$ para a posição normalizada ξ_x em função de h_x e h_y . O valor de κ_x é uma média em t dentro do intervalo de $30 \leq t \leq 1600$. Notamos que, dependendo dos valores de h_x e h_y , as distribuições $P(\xi_x)$ exibem perfis diferentes, podendo ser mesocúrticas ($\kappa_x = 3$, como no caso da gaussiana), leptocúrtica ($\kappa_x > 3$) ou platicúrtica ($\kappa_x < 3$). Também observamos que, semelhante ao caso da variância, κ_x não pode ser expresso em termos de uma combinação linear de h_x e h_y .

5.3 Conclusões e perspectivas

Propusemos uma extensão para o modelo de pente por meio das equações de Langevin governadas por ruídos correlacionados de longo alcance (ruídos gaussianos fracionários). Inicialmente mostramos que nossas equações de Langevin são equivalentes à equação de difusão generalizada (descrevendo o modelo de pente usual) por meio da comparação das variâncias e distribuições marginais obtidas numericamente com os resultados analíticos previstos pela equação de difusão. A seguir, apresentamos uma extensa caracterização das propriedades de difusão das posições das partículas em função dos expoentes de Hurst h_x e h_y definindo os ruídos correlacionados. Nossos resultados mostram que o ruído na direção y afeta as variâncias e as distribuições marginais ao longo da direção x . Especificamente, a persistência de longo alcance em y retarda a difusão em x reduzindo o expoente da lei de potência α_x e fazendo as caudas da distribuição marginal $P(\xi_x)$ mais curtas com o aumento de $h_y > 0,5$; enquanto a antipersistência de longo alcance em y acelera a difusão em x , aumentando o valor de α_x e fazendo as caudas de $P(\xi_x)$ mais curtas quanto menor $h_y < 0,5$. Investigamos também a interação entre as correlações de longo alcance nas direções x e y , na qual encontramos que tanto α_x quanto o coeficiente de curtose κ_x de $P(\xi_x)$ não são funções triviais (uma combinação linear) de h_x e h_y , mas, em vez disso, os nossos resultados sugerem um acoplamento mais complexo entre os expoentes de Hurst e a difusão na direção x . Em resumo, um cenário difusivo muito rico emerge do nosso modelo que pode ser aplicado diretamente para descrever (pelo menos numa primeira aproximação) alguns resultados empíricos relacionadas à difusão em sistemas complexos, incluindo células vivas. Além disso, o nosso modelo combina e permite o estudo/análise da interação entre diferentes mecanismos de difusão anômala, restrições geométricas e correlações de longo alcance, um fato importante que não havia sido explorado no âmbito do modelo de pente e que pode proporcionar uma melhor compreensão desses processos como uma combinação de diferentes mecanismos.

Visão geral

Nessa tese de doutorado estudamos sistemas complexos urbanos e biológicos via métodos de física estatística, simulação computacional e análise de dados. Várias das técnicas aqui empregadas são reminiscentes da física estatística e termodinâmica, incluindo conceitos de invariância de escala, autossimilaridade, fractalidade, transições de fases e difusão anômala. Os resultados obtidos por meio dos estudos desenvolvidos, nessa tese, são apresentados a seguir.

No primeiro capítulo, discutimos aspectos sobre invariância de escala (leis de escalas) entre indicadores urbanos e a população. Mostramos que os indicadores urbanos tais como trabalho infantil, crime, analfabetismo, desemprego, renda e indicadores de população escalam com o tamanho da população de uma maneira não linear, seguindo leis de escala com expoentes característicos que variam de regimes sublineares (*e.g.*, expoente $\approx 0,3$ para renda familiar) a regimes superlineares (*e.g.*, expoente $\approx 1,3$ para crimes de homicídio). Devido a essas não linearidades, propomos uma métrica ajustada à escala para a classificação e previsão de indicadores urbanos nas cidades brasileiras. Com a nova métrica, modelos de regressão simples descrevem de 31% a 97% da variância observada dependendo do indicador urbano em análise. O alto poder de descrição permitiu, então, fazer previsões nas quais verificamos aumentos e diminuições nos valores dos indicadores, dependendo da região do território brasileiro.

Ainda a despeito de métricas urbanas, no segundo capítulo, estudamos correlações espaciais, aglomeração/*cluster* e transições de fase tipo percolação no indicador de crime de homicídios *per*

capita. Crime de homicídios *per capita* em cidades exibe correlações de curto alcance (decaimento exponencial) e um comprimento de correlação característico que apresenta uma tendência crescente com o tempo. Devido às correlações espaciais, observamos a formação de agrupamentos/*clusters* de cidades similares. Por isso, usando uma análise de percolação, investigamos como o maior *cluster* se comporta em função de um limiar de número de homicídios *per capita*. Nesse cenário, observamos um comportamento tipo transição de fase na função do tamanho do maior *cluster* *versus* o limiar do número de homicídios *per capita*. Observamos também que esse comportamento tipo transição de fase tem ocorrido para limiares cada vez maiores e magnitudes de transição cada vez menores.

No terceiro capítulo, abordamos a questão da motilidade de quatro espécie de protozoários flagelados causadores de doenças tropicais negligenciadas. Analisamos vários aspectos relacionados à difusão dos protozoários, como a evolução da variância com o tempo, autossimilaridades nas distribuições de probabilidades das posições e correlações de longo alcance nas séries temporais das velocidades. Foi encontrada uma dinâmica anômala superdifusiva com distribuições de probabilidades das posições autossimilares e de cauda mais alongada (gaussiana alongada). Também observamos que as séries das velocidades são correlacionadas de longo alcance (expoente de Hurst próximo de um) com distribuições de velocidades que podem ser aproximadas por uma generalização da distribuição gama. Apesar das diferenças morfológicas das quatro espécies de protozoários, padrões aproximadamente universais parecem governar a dinâmica desses organismos unicelulares. Sugerimos que desvios dos padrões relatados aqui poderiam ser usados como indicativo de viabilidade celular no estudo de novas drogas.

No quarto capítulo, utilizamos técnicas similares àquelas utilizadas no capítulo 3 para o estudo do movimento de vermes. Os resultados desse capítulo estabelecem uma analogia frutífera entre a fisiologia humana e o comportamento do movimento do *C. elegans*. Da mesma forma que a fisiologia humana, descobrimos que as propriedades fractal do movimento desses nematóides mudam com a idade e com o estresse ambiental. Nossos resultados sugerem que a fisiologia do *C. elegans* compartilha uma similaridade mais profunda com a fisiologia humana do que se poderia esperar. Essa maior similaridade abre novos espaços para o estudo da fisiologia e do envelhecimento por causa das ferramentas disponíveis para alterar a genética e as condições ambientais desses invertebrados.

Apresentamos um resultado teórico sobre difusão no quinto capítulo. Propomos uma extensão do modelo de pente via equações de Langevin governadas por ruídos fracionários correlacionados de longo alcance. Mostramos que as equações de Langevin propostas são equivalentes à equação de difusão generalizada analítica descrevendo o caso usual no modelo de pente. Mostramos também que o ruído na direção y afeta os desvios e distribuições marginais ao longo da direção x . Especificamente, a correlação persistente de longo alcance em y retarda a difusão em x , enquanto a correlação antipersistente de longo alcance em y melhora (aumenta) a difusão na direção x . Esse modelo combina e permite o estudo/análise da interação entre diferentes mecanismos de difusão anômala (restrições geométricas e correlações de longo alcance) e pode ter aplicações diretas para descrever a difusão em sistemas complexos, como as células vivas.

Em resumo, nessa tese encontramos vários padrões e leis em sistemas complexos urbanos e biológicos, por meio da aplicação de métodos de física estatística. Além das contribuições específicas para cada problema descrito acima, as técnicas e resultados aqui encontrados podem ser aplicados a outros problemas de sistemas complexos. Esperamos que os resultados e as técnicas aqui descritas possam ser usados em trabalhos futuros em sistemas urbanos e biológicos, assim como em outros sistemas complexos.

Referências Bibliográficas

- [1] Shlesinger, M. F. Fractal time and $1/f$ noise in complex systems. *Annals of the New York Academy of Sciences* **504**, 214–228 (1987).
- [2] Bar-Yam, Y. *Dynamics of complex systems*, vol. 213 (Addison-Wesley Reading, MA, 1997).
- [3] Mantegna, R. N. & Stanley, H. E. *Introduction to econophysics: Correlations and complexity in finance* (Cambridge University Press, 1999).
- [4] Albert, R. & Barabási, A.-L. Statistical mechanics of complex networks. *Reviews of Modern Physics* **74**, 47 (2002).
- [5] Boccarra, N. *Modeling Complex Systems*, vol. 58 (Springer, 2005).
- [6] Castellano, C., Fortunato, S. & Loreto, V. Statistical physics of social dynamics. *Reviews of Modern Physics* **81**, 591–646 (2009).
- [7] Peng, C.-K. K., Buldyrev, S. V., Havlin, S., Simons, M., Stanley, H. E. & Goldberger, A. L. Mosaic organization of DNA nucleotides. *Physical Review E* **49**, 1685–1689 (1994).
- [8] Stanley, M. H. R., Amaral, L. A. N., Buldyrev, S. V., Havlin, S., Leschhorn, H., Maass, P., Salinger, M. A. & Stanley, H. E. Scaling behaviour in the growth of companies. *Nature* **379**, 804–806 (1996).

- [9] Conte, R., Gilbert, N., Bonelli, G., Cioffi-Revilla, C., Deffuant, G., Kertesz, J., Loreto, V., Moat, S., Nadal, J.-P., Sanchez, A., Nowak, A., Flache, A., San Miguel, M. & Helbing, D. Manifesto of computational social science. *European Physical Journal: Special Topics* **214**, 325–346 (2012).
- [10] Goldberger, A. L., Amaral, L. A. N., Glass, L., Hausdorff, J. M., Ivanov, P. C., Mark, R. G., Mietus, J. E., Moody, G. B., Peng, C.-K. & Stanley, H. E. PhysioBank, PhysioToolkit, and PhysioNet - Components of a new research resource for complex physiologic signals. *Circulation* **101**, 215–220 (2000).
- [11] Mendes, R. S., Ribeiro, H. V., Freire, F. C. M., Tateishi, A. A. & Lenzi, E. K. Universal patterns in sound amplitudes of songs and music genres. *Physical Review E* **83**, 017101 (2011).
- [12] Amancio, D. R., Oliveira Jr, O. N. & da Fontoura Costa, L. Identification of literary movements using complex networks to represent texts. *New Journal of Physics* **14**, 043029 (2012).
- [13] Ribeiro, H. V., Tateishi, A. A., Alves, L. G. A., Zola, R. S. & Lenzi, E. K. Investigating the interplay between mechanisms of anomalous diffusion via fractional Brownian walks on a comb-like structure. *New Journal of Physics* **16**, 093050 (2014).
- [14] Alves, L. G. A., Mendes, R. S., Lenzi, E. K. & Ribeiro, H. V. Scale-adjusted metrics for predicting the evolution of urban indicators and quantifying the performance of cities. *PLoS ONE* **10**, e0134862 (2015).
- [15] Alves, L. G. A., Lenzi, E. K., Mendes, R. S. & Ribeiro, H. V. Spatial correlations, clustering and percolation-like transitions in homicide crimes. *Europhysics Letters* **111**, 18002 (2015).
- [16] Alves, L. G. A., Scariot, D. B., Guimarães, R. R., Nakamura, C. V., Mendes, R. S. & Ribeiro, H. V. Transient superdiffusion and long-range correlations in the motility patterns of trypanosomatid flagellate protozoa. *PLoS ONE* **11**, e0152092 (2016).

- [17] Alves, L. G. A., Winter, P. B., Ferreira, L. N., Briemann, B. M., Morimoto, R. I. & Amaral, L. A. N. Long-range correlations and fractal dynamics in *C. elegans*: Changes with aging and stress. *Em preparação* 1–21 (2016).
- [18] Bernoulli, D. *Hydrodynamica sive de viribus et motibus fluidorum commentarii* [Commentary on hydrodynamics from the forces and movements of bodies of solids and liquids] (Johann Reinhold Dulsecker, 1738).
- [19] Maxwell, J. C. Illustrations of the dynamical theory of gases. Part I. On the motions and collisions of perfectly elastic spheres. *Philosophical Magazine* **19**, 19–31 (1860).
- [20] Maxwell, J. C. Illustrations of the dynamical theory of gases. Part II. On the process of diffusion of two or more kinds of moving particles among one another. *Philosophical Magazine* **20**, 21–37 (1860).
- [21] Boltzmann, L. Über die mechanische bedeutung des zweiten hauptsatzes der wärmetheorie [On the mechanical meaning of the second law of thermodynamic theory]. *Sitzungsberichte der Kaiserlichen Akademie der Wissenschaften* **53** (1866).
- [22] Gibbs, J. W. *Elementary principles in statistical mechanics* (New York: Charles Scribner's Sons, 1902).
- [23] Reif, F. *Fundamentals of statistical and thermal physics* (McGraw-Hill, 1965).
- [24] Reichl, L. E. *A modern course in statistical physics*, vol. 71 (University of Texas press Austin, 1980).
- [25] Callen, H. *Thermodynamics and an introduction to thermostatistics* (Wiley, 1985).
- [26] Kubo, R., Ichimura, H. & Usui, T. *Statistical mechanics: An advanced course with problems and solutions* (North-Holland, 1990).
- [27] Pathria, R. & Beale, P. *Statistical Mechanics* (Elsevier Science, 1996).
- [28] Wallace, D. *Statistical physics of crystals and liquids: A guide to highly accurate equations of state* (World Scientific, 2002).

- [29] Stanley, H. *Introduction to phase transitions and critical phenomena* (Oxford University Press, 1971).
- [30] Yeomans, J. *Statistical mechanics of phase transitions* (Clarendon Press, 1992).
- [31] Tang, C. & Bak, P. Critical exponents and scaling relations for self-organized critical phenomena. *Physical Review Letters* **60**, 2347–2350 (1988).
- [32] Stanley, H. Scaling, universality, and renormalization: Three pillars of modern critical phenomena. *Reviews of Modern Physics* **71**, S358–S366 (1999).
- [33] Sornette, D. *Critical Phenomena in natural sciences: Chaos, fractals, self-organization and disorder: Concepts and tools* (Springer Berlin Heidelberg, 2006).
- [34] Rothbard, M. *An Austrian perspective on the history of economic thought* (Edward Elgar, 1995).
- [35] Mandelbrot, B. How long is the coast of Britain? Statistical self-similarity and fractional dimension. *Science* **156** (1967).
- [36] Kleiber, M. Body size and metabolism. *Hilgardia: A Journal of Agricultural Science* **6** (1932).
- [37] Kleiber, M. Body size and metabolic rate. *Physiological Reviews* **27**, 511–41 (1947).
- [38] West, G. B., Brown, J. H. & Enquist, B. J. A general model for the origin of allometric scaling laws in biology. *Science* **276**, 122–126 (1997).
- [39] West, G. B., Woodruff, W. H. & Brown, J. H. Allometric scaling of metabolic rate from molecules and mitochondria to cells and mammals. *Proceedings of the National Academy of Sciences* **99**, 2473–2478 (2002).
- [40] West, G. B. & Brown, J. H. The origin of allometric scaling laws in biology from genomes to ecosystems: towards a quantitative unifying theory of biological structure and organization. *Journal of Experimental Biology* **208**, 1575–1592 (2005).

- [41] Bettencourt, L. M. A., Lobo, J., Helbing, D., Kühnert, C. & West, G. B. Growth, innovation, scaling, and the pace of life in cities. *Proceedings of the National Academy of Sciences* **104**, 7301–7306 (2007).
- [42] Bettencourt, L. M. & West, G. A unified theory of urban living. *Nature* **467**, 912–913 (2010).
- [43] Alves, L. G. A., Ribeiro, H. V. & Mendes, R. S. Scaling laws in the dynamics of crime growth rate. *Physica A*. **392**, 2672–2679 (2013).
- [44] Alves, L. G. A., Ribeiro, H. V., Lenzi, E. K. & Mendes, R. S. Distance to the scaling law: A useful approach for unveiling relationships between crime and urban metrics. *PLoS ONE* **8**, e0069580 (2013).
- [45] Arbesman, S., Kleinberg, J. M. & Strogatz, S. H. Superlinear scaling for innovation in cities. *Physical Review E* **79**, 016115 (2009).
- [46] Bettencourt, L. M. A., Lobo, J., Strumsky, D. & West, G. B. Urban scaling and its deviations: Revealing the structure of wealth, innovation and crime across cities. *PLoS ONE* **5**, e0013541 (2010).
- [47] Gomez-Lievano, A. & Youn, B. L. M. A., H. The statistics of urban scaling and their connection to Zipf’s law. *PLoS ONE* **7**, e0040393 (2012).
- [48] Alves, L. G. A., Ribeiro, H. V., Lenzi, E. K. & Mendes, R. S. Empirical analysis on the connection between power-law distributions and allometries for urban indicators. *Physica A*. **409**, 175–182 (2014).
- [49] Ignazzi, C. A. Scaling laws, economic growth, education and crime: Evidence from Brazil [Lois d’échelle, croissance économique, éducation et crime au Brésil]. *Espace Géographique* **43**, 324–337 (2014).
- [50] Melo, H. P. M., Moreira, A. A., Makse, H. A. & Andrade, J. S. Statistical signs of social influence on suicides. *Scientific Reports* **4**, 06239 (2014).
- [51] Mantovani, M. C., Ribeiro, H. V., Moro, M. V., Picoli, S. & Mendes, R. S. Scaling laws and universality in the choice of election candidates. *Europhysics Letters* **96**, 48001 (2011).

- [52] Mantovani, M. C., Ribeiro, H. V., Lenzi, E. K., Picoli, S. & Mendes, R. S. Engagement in the electoral processes: Scaling laws and the role of political positions. *Physical Review E* **88**, 024802 (2013).
- [53] Samaniego, H. & Moses, M. E. Cities as organisms: Allometric scaling of urban road networks. *Journal of Transport and Land Use* **1**, 21–39 (2008).
- [54] Louf, R., Roth, C. & Barthelemy, M. Scaling in transportation networks. *PLoS ONE* **9**, e0102007 (2014).
- [55] Pumain, D., Paulus, F., Vacchiani-Marcuzzo, C. & Lobo, J. An evolutionary theory for interpreting urban scaling laws. *Cybergeo* **219**, 2519 (2006).
- [56] Pan, W., Ghoshal, G., Krumme, C., Cebrian, M. & Pentland, A. Urban characteristics attributable to density-driven tie formation. *Nature Communications* **4**, 2961 (2013).
- [57] Brown, R. A brief account of microscopical observations made in the months of June, July, and August, 1827, on the particles contained in the pollen of plants; and on the general existence of active molecules in organic and inorganic bodies. *Philosophical Magazine* **4**, 161 (1828).
- [58] Einstein, A. Über die von der molekularkinetischen Theorie der Wärme geforderte Bewegung von in ruhenden Flüssigkeiten suspendierten Teilchen [On the motion of small particles suspended in liquids at rest required by the molecular-kinetic theory of heat]. *Annalen der Physik* **17**, 549 (1905).
- [59] Smoluchowski, M. Zur kinetischen Theorie der Brownschen Molekularbewegung und der Suspensionen [On the kinetic theory of Brownian molecular motion and suspensions]. *Annalen der Physik* **21**, 756 (1906).
- [60] Metzler, R., Jeon, J.-H., Cherstvy, A. G. & Barkai, E. Anomalous diffusion models and their properties: Non-stationarity, non-ergodicity, and ageing at the centenary of single particle tracking. *Physical Chemistry Chemical Physics* **16**, 24128–24164 (2014).

- [61] Upadhyaya, A., Rieu, J.-P., Glazier, J. A. & Sawada, Y. Anomalous diffusion and non-Gaussian velocity distribution of *Hydra* cells in cellular aggregates. *Physica A*. **293**, 549–558 (2001).
- [62] Reverey, J. F., Jeon, J.-H., Bao, H., Leippe, M., Metzler, R. & Selhuber-Unkel, C. Superdiffusion dominates intracellular particle motion in the supercrowded cytoplasm of pathogenic *Acanthamoeba castellanii*. *Scientific Reports* **5**, 11690 (2015).
- [63] Ribeiro, H. V., Costa, L. S., Alves, L. G. A., Santoro, P. A., Picoli, S., Lenzi, E. K. & Mendes, R. S. Analogies between the cracking noise of ethanol-dampened charcoal and earthquakes. *Physical Review Letters* **115**, 025503 (2015).
- [64] Alves, L. G. A., Ribeiro, H. V., Santos, M. A. F., Mendes, R. S. & Lenzi, E. K. Solutions for a q -generalized Schrödinger equation of entangled interacting particles. *Physica A*. **429**, 35–44 (2015).
- [65] Braga, A. C., Alves, L. G. A., Costa, L. S., Ribeiro, A. A., De Jesus, M. M., Tateishi, A. A. & Ribeiro, H. V. Characterization of river flow fluctuations via horizontal visibility graphs. *Physica A*. **444**, 1003–1011 (2016).
- [66] United Nations. *World Urbanization Prospects: The 2014 Revision, Highlights (ST/ESA/SER.A/352)* (United Nations, 2014). URL <http://esa.un.org/unpd/wup/Highlights/WUP2014-Highlights.pdf>.
- [67] Louf, R. & Barthelemy, M. How congestion shapes cities: From mobility patterns to scaling. *Scientific Reports* **4**, 5561 (2014).
- [68] Caminha, C., Furtado, V., Pequeno, T. H. C., Ponte, C., Melo, H. P. M., Oliveira, E. A. & Andrade Jr, J. S. Human mobility in large cities as a proxy for crime. *arXiv preprint arXiv:1610.09936* (2016).
- [69] Oliveira, E. A., Andrade Jr., J. S. & Makse, H. A. Large cities are less green. *Scientific Reports* **4**, 4235 (2014).
- [70] Louf, R. & Barthelemy, M. Scaling: Lost in the smog. *Environment and Planning B: Planning and Design* **41**, 767 (2014).

- [71] Rybski, D., Reusser, D. E., Winz, A.-L., Fichtner, C., Sterzel, T. & Kropp, J. P. Cities as nuclei of sustainability? *Environment and Planning B: Planning and Design* 0265813516638340 (2016).
- [72] Gordon, M. B. A random walk in the literature on criminality: A partial and critical view on some statistical analyses and modelling approaches. *European Journal of Applied Mathematics* **21**, 283–306 (2010).
- [73] Pumain, D. Scaling laws and urban systems. *SFI Working Paper* (2004).
- [74] Bettencourt, L. M. A. The origins of scaling in cities. *Science* **340**, 1438–1441 (2013).
- [75] Podobnik, B., Horvatic, D., Kenett, D. Y. & Stanley, H. E. The competitiveness versus the wealth of a country. *Scientific Reports* **2**, 00678 (2012).
- [76] Lobo, J., Bettencourt, L. M. A., Strumsky, D. & West, G. B. Urban scaling and the production function for cities. *PLoS ONE* **8**, e0058407 (2013).
- [77] Department of Data Processing. *Brazil's Public healthcare System (SUS)* (DATASUS, 2011). URL <http://www.datasus.gov.br/>.
- [78] Angel, S. & Sheppard, S. *The dynamics of global urban expansion* (The World Bank, 2005). URL http://www.citiesalliance.org/sites/citiesalliance.org/files/CA_Docs/resources/upgrading/urban-expansion/1.pdf.
- [79] Schläpfer, M., Bettencourt, L. M. A., Grauwin, S., Raschke, M., Claxton, R., Smoreda, Z., West, G. B. & Ratti, C. The scaling of human interactions with city size. *Journal of the Royal Society Interface* **11**, 20130789 (2014).
- [80] Arcaute, E., Hatna, E., Ferguson, P., Youn, H., Johansson, A. & Batty, M. Constructing cities, deconstructing scaling laws. *Journal of the Royal Society Interface* **12**, 20140745 (2015).
- [81] Masucci, A. P., Arcaute, E., Hatna, E., Stanilov, K. & Batty, M. On the problem of boundaries and scaling for urban street networks. *Journal of The Royal Society Interface* **12** (2015).

- [82] Arcaute, E., Molinero, C., Hatna, E., Murcio, R., Vargas-Ruiz, C., Masucci, A. P. & Batty, M. Cities and regions in Britain through hierarchical percolation. *Royal Society Open Science* **3** (2016).
- [83] Rozenfeld, H. D., Rybski, D., Andrade, J. S., Batty, M., Stanley, H. E. & Makse, H. A. Laws of population growth. *Proceedings of the National Academy of Sciences* **105**, 18702–18707 (2008).
- [84] Instituto Brasileiro de Geografia e Estatística. *Brazilian metropolitan areas* (IBGE, 2015). URL <http://www.ibge.gov.br/>.
- [85] Bettencourt, L. M. A., Lobo, J. & Youn, H. The hypothesis of urban scaling: Formalization, implications and challenges. *SFI Working Paper* (2013).
- [86] Ortman, S. G., Cabaniss, A. H. F., Sturm, J. O. & Bettencourt, L. M. A. The pre-history of urban scaling. *PLoS ONE* **9**, e0087902 (2014).
- [87] Oliphant, T. E. & Peterson, P. SciPy: Open source scientific tools for Python (2001). URL <http://www.scipy.org/>.
- [88] Christakis, N. A. & Fowler, J. H. The spread of obesity in a large social network over 32 years. *New England Journal of Medicine* **357**, 370–379 (2007).
- [89] Hesketh, T. & Xing, Z. W. Abnormal sex ratios in human populations: Causes and consequences. *Proceedings of the National Academy of Sciences* **103**, 13271–13275 (2006).
- [90] Galam, S. *Sociophysics: A physicist's modeling of psycho-political phenomena* (Springer, 2012).
- [91] Vespignani, A. Modelling dynamical processes in complex socio-technical systems. *Nature Physics* **8**, 32–39 (2012).
- [92] D'Orsogna, M. R. & Perc, M. Statistical physics of crime: A review. *Physics of Life Reviews* **12**, 1–21 (2015).
- [93] Wilson, J. Q. & Kelling, G. L. Broken windows. *Atlantic Monthly* **3**, 29–38 (1982).

- [94] Short, M. B., D’Orsogna, M. R., Pasour, V. B., Tita, G. E., Brantingham, P. J., Bertozzi, A. L. & Chayes, L. B. A statistical model of criminal behavior. *Mathematical Models and Methods in Applied Sciences* **18**, 1249–1267 (2008).
- [95] Keizer, K., Lindenberg, S. & Steg, L. The spreading of disorder. *Science* **322**, 1681–1685 (2008).
- [96] Short, M. B., D’Orsogna, M. R., Brantingham, P. J. & Tita, G. E. Measuring and modeling repeat and near-repeat burglary effects. *Journal of Quantitative Criminology* **25**, 325–339 (2009).
- [97] Short, M. B., Brantingham, P. J., Bertozzi, A. L. & Tita, G. E. Dissipation and displacement of hotspots in reaction-diffusion models of crime. *Proceedings of the National Academy of Sciences* **107**, 3961–3965 (2010).
- [98] Rodriguez, N. & Bertozzi, A. Local existence and uniqueness of solutions to a PDE model for criminal behavior. *Mathematical Models and Methods in Applied Sciences* **20**, 1425–1457 (2010).
- [99] Mohler, G. O., Short, M. B., Brantingham, P. J., Schoenberg, F. P. & Tita, G. E. Self-exciting point process modeling of crime. *Journal of the American Statistical Association* **106**, 100–108 (2011).
- [100] Mohler, G. O. & Short, M. B. Geographic profiling from kinetic models of criminal behavior. *SIAM Journal on Applied Mathematics* **72**, 163–180 (2012).
- [101] Lewis, E., Mohler, G., Brantingham, P. J. & Bertozzi, A. L. Self-exciting point process models of civilian deaths in Iraq. *Security Journal* **25**, 244–264 (2012).
- [102] Harries, K. Mapping Crime . Principle and practice. *National Institute of Justice* **164**, 206 (1999). URL <https://www.ncjrs.gov/App/Publications/abstract.aspx?ID=178919>.
- [103] Gallos, L. K., Barttfeld, P., Havlin, S., Sigman, M. & Makse, H. A. Collective behavior in the spatial spreading of obesity. *Scientific Reports* **2**, 00454 (2012).

- [104] Edwards, S. F. & Wilkinson, D. R. The surface statistics of a granular aggregate. *Proceedings of the Royal Society A* **381**, 17–31 (1982).
- [105] Nattermann, T. & Tang, L.-H. Kinetic surface roughening. I. the Kardar-Parisi-Zhang equation in the weak-coupling regime. *Physical Review A* **45**, 7156–7161 (1992).
- [106] Barabási, A.-L. & Stanley, H. E. *Fractal concepts in surface growth* (Cambridge University Press, 1995).
- [107] Borghesi, C. & Bouchaud, J.-P. Spatial correlations in vote statistics: A diffusive field model for decision-making. *European Physical Journal B* **75**, 395–404 (2010).
- [108] Borghesi, C., Raynal, J.-C. & Bouchaud, J.-P. Election turnout statistics in many countries: Similarities, differences, and a diffusive field model for decision-making. *PLoS ONE* **7**, e0036289 (2012).
- [109] Ribeiro, H. V., Antonio, F. J., Alves, L. G. A., Lenzi, E. K. & Mendes, R. S. Long-range spatial correlations and fluctuation statistics of lightning activity rates in Brazil. *Europhysics Letters* **104**, 69001 (2013).
- [110] Makse, H. A., Havlin, S. & Stanley, H. E. Modelling urban growth patterns. *Nature* **377**, 608–612 (1995).
- [111] Makse, H. A., Andrade Jr., J. S., Batty, M., Havlin, S. & Eugene Stanley, H. Modeling urban growth patterns with correlated percolation. *Physical Review E* **58**, 7054–7062 (1998).
- [112] Pedregosa, F. & Varoquaux, G. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research* **12**, 2825–2830 (2011).
- [113] Tortora, G., Funke, B. & Case, C. *Microbiology: An introduction* (Benjamin Cummings, 2001).
- [114] Nussbaum, K., Honek, J., Cadmus, C. M. & Efferth, T. Trypanosomatid parasites causing neglected diseases. *Current Medicinal Chemistry* **17**, 1594–1617 (2010).

- [115] Manderson, L., Aagaard-Hansen, J., Allotey, P., Gyapong, M. & Sommerfeld, J. Social research on neglected diseases of poverty: Continuing and emerging themes. *PLoS Neglected Tropical Diseases* **3** (2009).
- [116] Anônimo. Parasites-leishmaniasis. *Epidemiology & Risk Factors* (2015).
- [117] World Health Organization. Neglected tropical diseases (2015). URL <http://www.who.int/neglected{ }diseases/diseases/en/>.
- [118] Baron, S. *Medical Microbiology* (University of Texas Medical Branch at Galveston, Galveston, 1996).
- [119] Kennedy, P. G. E. Human African trypanosomiasis of the CNS: Current issues and challenges. *Journal of Clinical Investigation* **113**, 496–504 (2004).
- [120] Grab, D. J. & Kennedy, P. G. E. Traversal of human and animal trypanosomes across the blood-brain barrier. *Journal of NeuroVirology* **14**, 344–351 (2008).
- [121] Dianqing, W. Signaling mechanisms for regulation of chemotaxis. *Cell Research* **15**, 52–56 (2005).
- [122] Langousis, G. & Hill, K. L. Motility and more: The flagellum of *Trypanosoma brucei*. *Nature Reviews Microbiology* **12**, 505–518 (2014).
- [123] Forestier, C.-L., MacHu, C., Loussert, C., Pescher, P. & Späth, G. F. Imaging host cell-*Leishmania* interaction dynamics implicates parasite motility, lysosome recruitment, and host cell wounding in the infection process. *Cell Host and Microbe* **9**, 319–330 (2011).
- [124] Hill, K. L. Biology and mechanism of trypanosome cell motility. *Eukaryotic Cell* **2**, 200–208 (2003).
- [125] Ginger, M. L., Portman, N. & McKean, P. G. Swimming with protists: Perception, motility and flagellum assembly. *Nature Reviews Microbiology* **6**, 838–850 (2008).
- [126] Rodríguez, J. A., Lopez, M. A., Thayer, M. C., Zhao, Y., Oberholzer, M., Chang, D. D., Kisalu, N. K., Penichet, M. L., Helguera, G., Bruinsma, R., Hill, K. L. & Miao, J. Propulsion

of African trypanosomes is driven by bihelical waves with alternating chirality separated by kinks. *Proceedings of the National Academy of Sciences* **106**, 19322–19327 (2009).

- [127] Alizadehrad, D., Krüger, T., Engstler, M. & Stark, H. Simulating the complex cell design of *Trypanosoma brucei* and its motility. *PLoS Computational Biology* **11** (2015).
- [128] Thurner, S., Wick, N., Hanel, R., Sedivy, R. & Huber, L. Anomalous diffusion on dynamical networks: A model for interacting epithelial cell migration. *Physica A*. **320**, 475–484 (2003).
- [129] Song, L., Nadkarni, S. M., Bödeker, H. U., Beta, C., Bae, A., Franck, C., Rappel, W.-J., Loomis, W. F. & Bodenschatz, E. *Dictyostelium discoideum* chemotaxis: Threshold for directed motion. *European Journal of Cell Biology* **85**, 981–989 (2006).
- [130] Amselem, G., Theves, M., Bae, A., Bodenschatz, E. & Beta, C. A stochastic description of *Dictyostelium* chemotaxis. *PLoS ONE* **7**, e0037213 (2012).
- [131] Theves, M., Taktikos, J., Zaburdaev, V., Stark, H. & Beta, C. A bacterial swimmer with two alternating speeds of propagation. *Biophysical Journal* **105**, 1915–1924 (2013).
- [132] Makarava, N., Menz, S., Theves, M., Huisinga, W., Beta, C. & Holschneider, M. Quantifying the degree of persistence in random amoeboid motion based on the Hurst exponent of fractional Brownian motion. *Physical Review E* **90**, 042703 (2014).
- [133] Theves, M., Taktikos, J., Zaburdaev, V., Stark, H. & Beta, C. Random walk patterns of a soil bacterium in open and confined environments. *Europhysics Letters* **109**, 28007 (2015).
- [134] Tsallis, C. Possible generalization of Boltzmann-Gibbs statistics. *Journal of Statistical Physics* **52**, 479–487 (1988).
- [135] Plastino, A. R. & Plastino, A. Non-extensive statistical mechanics and generalized Fokker-Planck equation. *Physica A*. **222**, 347–354 (1995).
- [136] Zarley, J. H., Britigan, B. E. & Wilson, M. E. Hydrogen peroxide-mediated toxicity for *Leishmania donovani* chagasi promastigotes: Role of hydroxyl radical and protection by heat shock. *Journal of Clinical Investigation* **88**, 1511–1521 (1991).

- [137] Broadhead, R., Dawe, H. R., Farr, H., Griffiths, S., Hart, S. R., Portman, N., Shaw, M. K., Ginger, M. L., Gaskell, S. J., McKean, P. G. & Gull, K. Flagellar motility is required for the viability of the bloodstream trypanosome. *Nature* **440**, 224–227 (2006).
- [138] Uppaluri, S., Nagler, J., Stellamanns, E., Heddergott, N., Herminghaus, S., Engstler, M. & Pfohl, T. Impact of microscopic motility on the swimming behavior of parasites: Straighter Trypanosomes are more directional. *PLoS Computational Biology* **7**, e1002058 (2011).
- [139] Heddergott, N., Krüger, T., Babu, S. B., Wei, A., Stellamanns, E., Uppaluri, S., Pfohl, T., Stark, H. & Engstler, M. Trypanosome motion represents an adaptation to the crowded environment of the vertebrate bloodstream. *PLoS Pathogens* **8**, e1003023 (2012).
- [140] Ballesteros-Rodea, G., Santillán, M., Martínez-Calvillo, S. & Manning-Cela, R. Flagellar motility of *Trypanosoma cruzi* epimastigotes. *Journal of Biomedicine and Biotechnology* **2012**, 520380 (2012).
- [141] Hoare, C. A. & Wallace, F. G. Developmental stages of trypanosomatid flagellates: a new terminology. *Nature* **212**, 1385 – 1386 (1966).
- [142] Rozenberg, G. *Microscopic haematology: A practical guide to laboratory* (Churchill Livingstone, London, 2003).
- [143] Farrar, J., Hotez, P. J., Junghanss, T., Kang, G., Lalloo, D. & White, N. J. *Manson's tropical diseases* (W. B. Saunders, 2013).
- [144] De Noya, B. A., González, O. N. & Robertson, L. J. *Trypanosoma cruzi as a foodborne pathogen* (Springer International Publishing, 2015).
- [145] Wauthier, F. *Motion tracking in image sequences* (Department of Statistics, University of Oxford, 2015). URL <http://www.stats.ox.ac.uk/~wauthier/tracker/>.
- [146] Lauga, E., DiLuzio, W. R., Whitesides, G. M. & Stone, H. A. Swimming in circles: Motion of bacteria near solid boundaries. *Biophysical Journal* **90**, 400–412 (2006).
- [147] Klafter, J. & Sokolov, I. M. Anomalous diffusion spreads its wings. *Physics World* **18**, 29–32 (2005).

- [148] David Barry, J. & McCulloch, R. Antigenic variation in trypanosomes: Enhanced phenotypic variation in a eukaryotic parasite. *Advances in Parasitology* **49**, 2–70 (2001).
- [149] Donelson, J. E. Antigenic variation and the African trypanosome genome. *Acta Tropica* **85**, 391–404 (2003).
- [150] Lythgoe, K. A., Morrison, L. J., Read, A. F. & Barry, J. D. Parasite-intrinsic factors can explain ordered progression of trypanosome antigenic variation. *Proceedings of the National Academy of Sciences* **104**, 8095–8100 (2007).
- [151] Pays, E., Vanhamme, L. & Pérez-Morga, D. Antigenic variation in *Trypanosoma brucei*: Facts, challenges and mysteries. *Current Opinion in Microbiology* **7**, 369–374 (2004).
- [152] Souza, W. Basic cell biology of *Trypanosoma cruzi*. *Current pharmaceutical design* **8**, 269–285 (2002).
- [153] Bates, P. A. Transmission of *Leishmania* metacyclic promastigotes by phlebotomine sand flies. *International Journal for Parasitology* **37**, 1097–1106 (2007).
- [154] Rogers, M., Kropf, P., Choi, B.-S., Dillon, R., Podinovskaia, M., Bates, P. & Müller, I. Proteophosphoglycans regurgitated by *Leishmania*-infected sand flies target the L-arginine metabolism of host macrophages to promote parasite survival. *PLoS Pathogens* **5**, e1000555 (2009).
- [155] Rogers, M. E., Chance, M. L. & Bates, P. A. The role of promastigote secretory gel in the origin and transmission of the infective stage of *Leishmania mexicana* by the sandfly *Lutzomyia longipalpis*. *Parasitology* **124**, 495–507 (2002).
- [156] Rogers, M. E. The role of *Leishmania* proteophosphoglycans in sand fly transmission and infection of the mammalian host. *Frontiers in Microbiology* **3**, 00223 (2012).
- [157] Richardson, L. F. Atmospheric diffusion shown on a distance-neighbour graph. *Proceedings of the Royal Society A* **110**, 709–737 (1926).
- [158] Kantelhardt, J. W., Koscielny-Bunde, E., Rego, H. H. A., Havlin, S. & Bunde, A. Detecting long-range correlations with detrended fluctuation analysis. *Physica A*. **295**, 441–454 (2001).

- [159] Atkins, P. & Paula, J. D. *Physical chemistry: Thermodynamics, structure, and change* (Oxford University Press, 2006).
- [160] Silverberg, J. L., Bierbaum, M., Sethna, J. P. & Cohen, I. Collective motion of humans in mosh and circle pits at heavy metal concerts. *Physical Review Letters* **110**, 228701 (2013).
- [161] Morgan, E. C., Lackner, M., Vogel, R. M. & Baise, L. G. Probability distributions for offshore wind speeds. *Energy Conversion and Management* **52**, 15–26 (2011).
- [162] Kiss, P. & Jánosi, I. M. Comprehensive empirical analysis of ERA-40 surface wind speed distribution over Europe. *Energy Conversion and Management* **49**, 2142–2151 (2008).
- [163] Carta, J. A., Ramírez, P. & Velázquez, S. A review of wind speed probability distributions used in wind energy analysis. Case studies in the Canary Islands. *Renewable and Sustainable Energy Reviews* **13**, 933–955 (2009).
- [164] Van Der Auwera, L., De Meyer, F. & Malet, L. M. The use of the Weibull three-parameter model for estimating mean wind power densities. *Journal of Applied Meteorology* **19**, 819–825 (1980).
- [165] Shaw, R. F. & Bercaw, B. L. Temperature and life-span in poikilothermous animals. *Nature* **196**, 454–457 (1962).
- [166] Herndon, L. A., Schmeissner, P. P. J., Dudaronek, J. M. J., Brown, P. a., Listner, K. M. K., Sakano, Y., Paupard, M. C. M., Hall, D. H. D. & Driscoll, M. Stochastic and genetic factors influence tissue-specific decline in ageing *C. elegans*. *Nature* **419**, 808–814 (2002).
- [167] Lu, T., Pan, Y., Kao, S.-Y., Li, C., Kohane, I., Chan, J. & Yankner, B. A. Gene regulation and DNA damage in the ageing human brain. *Nature* **429**, 883–891 (2004).
- [168] Luo, L., Cook, N., Venkatachalam, V., Martinez-Velazquez, L. A., Zhang, X., Calvo, A. C., Hawk, J., MacInnis, B. L., Frank, M., Ng, J. H. R., Klein, M., Gershow, M., Hammarlund, M., Goodman, M. B., Colón-Ramos, D. A., Zhang, Y. & Samuel, A. D. T. Bidirectional thermotaxis in *Caenorhabditis elegans* is mediated by distinct sensorimotor strategies driven by the AFD thermosensory neurons. *Proceedings of the National Academy of Sciences* **111**, 2776–2781 (2014).

- [169] Walleczek, J. *Self-organized biological dynamics and nonlinear control: Toward understanding complexity, chaos and emergent function in living systems* (Cambridge University Press, 2006).
- [170] Ivanov, P. C., Amaral, L. A. N., Goldberger, A. L., Havlin, S., Rosenblum, M. G., Struzik, Z. R. & Stanley, H. E. Multifractality in human heartbeat dynamics. *Nature* **399**, 461–465 (1999).
- [171] Goldberger, A. L., Amaral, L. A. N., Hausdorff, J. M., Ivanov, P. C., Peng, C.-K. & Stanley, H. E. Fractal dynamics in physiology: Alterations with disease and aging. *Proceedings of the National Academy of Sciences* **99**, 2466–2472 (2002).
- [172] Hausdorff, J. M. Gait dynamics, fractals and falls: Finding meaning in the stride-to-stride fluctuations of human walking. *Human Movement Science* **26**, 555–589 (2007).
- [173] Winter, P. B., Brielmann, R. M., Timkovich, N. P., Helio, T., Teixeira-Castro, A., Morimoto, D., Richard I & Amaral, L. A. N. A network approach to discerning the identities of visually indistinguishable *C. elegans* in a free moving population. *Scientific Reports* **6**, 34859 (2016).
- [174] Mandelbrot, B. B. *The fractal geometry of nature*, vol. 51 (American Journal of Physics, 1983).
- [175] Credidio, H. F., Teixeira, E. N., Reis, S. D., Moreira, A. A. & Andrade Jr, J. S. Statistical patterns of visual search for hidden objects. *Scientific Reports* **2**, 920 (2012).
- [176] Stanley, H. E. & Ostrowsky, N. *On growth and form: Fractal and non-fractal patterns in physics*. Nato Science Series E: (Springer Netherlands, 1985).
- [177] Viswanathan, G. M., Afanasyev, V., Buldyrev, S. V., Murphy, E. J., Prince, P. A. & Stanley, H. E. Lévy flight search patterns of wandering albatrosses. *Nature* **381**, 413–415 (1996).
- [178] Ramos-Fernández, G., Mateos, J. L., Miramontes, O., Cocho, G., Larralde, H. & Ayala-Orozco, B. Lévy walk patterns in the foraging movements of spider monkeys. *Behavioral Ecology and Sociobiology* **55**, 223–230 (2004).

- [179] Sims, D., Southall, E. J., Humphries, N. E., Hays, G. C., Bradshaw, C. J. a., Pitchford, J. W., James, A., Ahmed, M. Z., Brierley, A. S., Hindell, M. a., Morritt, D., Musyl, M. K., Righton, D., Shepard, E. L. C., Wearmouth, V. J., Wilson, R. P., Witt, M. J. & Metcalfe, J. D. Scaling laws of marine predator search behaviour. *Nature* **451**, 1098–1102 (2008).
- [180] Alados, C. L., Escos, J. M. & Emlen, J. M. Fractal structure of sequential behaviour patterns: An indicator of stress. *Animal Behaviour* **51**, 437–443 (1996).
- [181] Alados, C. L. & Weber, D. N. Lead effects on the precitability of reproductive behavior in fathead minnows (*Pimephales promelas*): A mathematical model. *Environmental Toxicology and Chemistry* **18**, 2392–2399 (1999).
- [182] Rutherford, K. M. D., Haskell, M. J., Glasbey, C., Jones, R. B. & Lawrence, A. B. Detrended fluctuation analysis of behavioural responses to mild acute stressors in domestic hens. *Applied Animal Behaviour Science* **83**, 125–139 (2003).
- [183] Haspel, G., O'Donovan, M. J. & Hart, A. C. Motoneurons dedicated to either forward or backward locomotion in the nematode *Caenorhabditis elegans*. *The Journal of neuroscience: The official journal of the Society for Neuroscience* **30**, 11151–11156 (2010).
- [184] Brenner, S. The genetics of *Caenorhabditis elegans*. *Genetics* **77**, 71–94 (1974).
- [185] Hahm, J.-H., Kim, S., DiLoreto, R., Shi, C., Lee, S.-J. V., Murphy, C. T. & Nam, H. G. *C. elegans* maximum velocity correlates with healthspan and is maintained in worms with an insulin receptor mutation. *Nature communications* **6**, 8919 (2015).
- [186] Stephens, G. J., Johnson-Kerner, B., Bialek, W. & Ryu, W. S. Dimensionality and dynamics in the behavior of *C. elegans*. *PLoS Computational Biology* **4**, e1000028 (2008).
- [187] Swierczek, N. A., Giles, A. C., Rankin, C. H. & Kerr, R. A. High-throughput behavioral analysis in *C. elegans*. *Nature Methods* **8**, 592–598 (2011).
- [188] Falconer, K. *Fractal geometry mathematical foundations and applications*, vol. 46 (Biometrics, 1990).

- [189] Vahabi, M., Schulz, J. H. P., Shokri, B. & Metzler, R. Area coverage of radial Lévy flights with periodic boundary conditions. *Physical Review E* **87**, 1–10 (2013).
- [190] McMullen, P. D., Aprison, E. Z., Winter, P. B., Amaral, L. A. N., Morimoto, R. I. & Ruvinsky, I. Macro-level modeling of the response of *C. elegans* reproduction to chronic heat stress. *PLoS Computational Biology* **8**, e1002338 (2012).
- [191] Mandelbrot, B. B. & Van Ness, J. W. Fractional Brownian motions, fractional noises and applications. *SIAM Reviews* **10**, 422–437 (1968).
- [192] Gripenberg, G. & Norros, I. On the prediction of fractional Brownian motion. *Journal of Applied Probability* **33**, 400–410 (1996).
- [193] Bouchaud, J.-P. & Georges, A. Anomalous diffusion in disordered media: Statistical mechanisms, models and physical applications. *Physics Reports* **195**, 127–293 (1990).
- [194] Metzler, R. & Klafter, J. The random walk’s guide to anomalous diffusion: A fractional dynamics approach. *Physics Reports* **339**, 1–77 (2000).
- [195] White, S. R. & Barma, M. Field-induced drift and trapping in percolation networks. *Journal of Physics A: General Physics* **17**, 2995–3008 (1984).
- [196] Ben-Avraham, D. & Havlin, S. *Diffusion and reactions in fractals and disordered systems* (Cambridge University Press, Cambridge, 2000).
- [197] Weiss, G. H. & Havlin, S. Some properties of a random walk on a comb structure. *Physica A*. **134**, 474–482 (1986).
- [198] Baskin, E. & Iomin, A. Superdiffusion on a comb structure. *Physical Review Letters* **93**, 120603 (2004).
- [199] Arkhincheev, V. E. Random walks on the comb model and its generalizations. *Chaos* **17**, 043102 (2007).
- [200] Iomin, A. Subdiffusion on a fractal comb. *Physical Review E* **83**, 052106 (2011).

- [201] Tateishi, A. A., Lenzi, E. K., Ribeiro, H. V., Evangelista, L. R., Mendes, R. S. & Da Silva, L. R. Solutions for a diffusion equation with a backbone term. *Journal of Statistical Mechanics* **2011**, P02022 (2011).
- [202] Villamaina, D., Sarracino, A., Gradenigo, G., Puglisi, A. & Vulpiani, A. On anomalous diffusion and the out of equilibrium response function in one-dimensional models. *Journal of Statistical Mechanics* **2011**, L01002 (2011).
- [203] Tateishi, A. A., Michels, F. S., Dos Santos, M. A. F., Lenzi, E. K. & Ribeiro, H. V. First passage time for a diffusive process under a geometric constraint. *Journal of Statistical Mechanics* **2013**, 9017 (2013).
- [204] Lenzi, E. K., Da Silva, L. R., Tateishi, A. A., Lenzi, M. K. & Ribeiro, H. V. Diffusive process on a backbone structure with drift terms. *Physical Review E* **87**, 012121 (2013).
- [205] Iomin, A. Superdiffusion of cancer on a comb structure. *Journal of Physics: Conference Series* **7**, 57 (2005).
- [206] Iomin, A. Toy model of fractional transport of cancer cells due to self-entrapping. *Physical Review E* **73**, 061918 (2006).
- [207] Méndez, V. & Iomin, A. Comb-like models for transport along spiny dendrites. *Chaos, Solitons and Fractals* **53**, 46–51 (2013).
- [208] Sagi, Y., Brook, M., Almog, I. & Davidson, N. Observation of anomalous diffusion and fractional self-similarity in one dimension. *Physical Review Letters* **108**, 093002 (2012).
- [209] Höfling, F. & Franosch, T. Anomalous transport in the crowded world of biological cells. *Reports on Progress in Physics* **76**, 046602 (2013).
- [210] Condamin, S., Tejedor, V., Voituriez, R., Bénichou, O. & Klafter, J. Probing microscopic origins of confined subdiffusion by first-passage observables. *Proceedings of the National Academy of Sciences* **105**, 5675–5680 (2008).
- [211] Selhuber-Unkel, C., Berg-Sørensen, K. & Oddershede, L. B. Variety in intracellular diffusion during the cell cycle. *Physical Biology* **6** (2009).

- [212] Jeon, J.-H., Tejedor, V., Burov, S., Barkai, E., Selhuber-Unkel, C., Berg-Sørensen, K., Oddershede, L. & Metzler, R. In vivo anomalous diffusion and weak ergodicity breaking of lipid granules. *Physical Review Letters* **106**, 048103 (2011).
- [213] Burov, S., Jeon, J.-H., Metzler, R. & Barkai, E. Single particle tracking in systems showing anomalous diffusion: The role of weak ergodicity breaking. *Physical Chemistry Chemical Physics* **13**, 1800–1812 (2011).
- [214] Leijnse, N., Jeon, J.-H., Loft, S., Metzler, R. & Oddershede, L. B. Diffusion inside living human cells. *European Physical Journal: Special Topics* **204**, 75–84 (2012).
- [215] Sokolov, I. M. Models of anomalous diffusion in crowded environments. *Soft Matter* **8**, 9043–9052 (2012).
- [216] Cherstvy, A. G., Chechkin, A. V. & Metzler, R. Particle invasion, survival, and non-ergodicity in 2D diffusion processes with space-dependent diffusivity. *Soft Matter* **10**, 1591–1601 (2014).
- [217] Brackley, C. A., Cates, M. E. & Marenduzzo, D. Intracellular facilitated diffusion: Searchers, crowdors, and blockers. *Physical Review Letters* **111**, 108101 (2013).
- [218] Girst, S., Hable, V., Drexler, G. A., Greubel, C., Siebenwirth, C., Haum, M., Friedl, A. A. & Dollinger, G. Subdiffusion supports joining of correct ends during repair of DNA double-strand breaks. *Scientific Reports* **3**, 02511 (2013).
- [219] Tabei, S. M. A., Burov, S., Kim, H. Y., Kuznetsov, A., Huynh, T., Jureller, J., Philipson, L. H., Dinner, A. R. & Scherer, N. F. Intracellular transport of insulin granules is a subordinated random walk. *Proceedings of the National Academy of Sciences* **110**, 4911–4916 (2013).
- [220] Braüchle, C., Lamb, C. D. & Michaelis, J. *Single particle tracking and single molecule energy transfer* (Wiley-VCH, Weinheim, 2010).
- [221] Weiss, M. Single-particle tracking data reveal anticorrelated fractional Brownian motion in crowded fluids. *Physical Review E* **88**, 010101 (2013).

- [222] Yamamoto, E., Akimoto, T., Yasui, M. & Yasuoka, K. Origin of subdiffusion of water molecules on cell membrane surfaces. *Scientific Reports* **4**, 04720 (2014).
- [223] Arkhincheev, V. E. & Baskin, E. M. Anomalous diffusion and drift in a comb model of percolation clusters. *Zh. Éksp. Teor. Fiz.* **73**, 292 (1991).
- [224] Hosking, J. R. M. Modeling persistence in hydrological time series using fractional differencing. *Water Resources Research* **20**, 1898–1908 (1984).
- [225] Arkhincheev, V. E. Random walk on hierarchical comb structures. *Journal of Experimental and Theoretical Physics* **88**, 710–715 (1999).
- [226] da Silva, L. R., Tateishi, A. A., Lenzi, M. K., Lenzi, E. K. & da Silva, P. C. Green function for a non-Markovian Fokker-Planck equation: Comb-model and anomalous diffusion. *Brazilian Journal of Physics* **39**, 483–487 (2009).