

UNIVERSIDADE DE SÃO PAULO
ESCOLA POLITÉCNICA - PECE

DANIEL GAZOLLI MARQUES BARBOSA

**ANÁLISE NÃO LINEAR DE SÉRIES TEMPORAIS APLICADA AO
PREÇO DA ENERGIA ELÉTRICA NO BRASIL**

São Paulo
2024

DANIEL GAZOLLI MARQUES BARBOSA

**ANÁLISE NÃO LINEAR DE SÉRIES TEMPORAIS APLICADA AO
PREÇO DA ENERGIA ELÉTRICA NO BRASIL**

Trabalho de Conclusão de Curso da
Universidade de São Paulo (PECE) como
requisito para obtenção do certificado do curso
de MBA em Engenharia Financeira

Orientador: Marcio Eisencraft

São Paulo

2024

AGRADECIMENTOS

A realização deste trabalho acadêmico não apenas representa um marco significativo em minha jornada educacional, mas também ressalta a importância fundamental do estudo e da pesquisa para o avanço do conhecimento e o desenvolvimento da sociedade. A academia desempenha um papel crucial na sociedade, servindo como um farol de inovação, pensamento crítico e solução de problemas complexos que enfrentamos no mundo contemporâneo.

Portanto, gostaria de expressar minha mais profunda gratidão ao meu orientador, Prof. Marcio Eisencraft, por sua paciência e orientação ao longo deste trabalho. Seu apoio acadêmico foi fundamental para a realização deste estudo, e suas valiosas contribuições enriqueceram minha jornada acadêmica.

Não posso deixar de agradecer à minha família pelo suporte incondicional e amor que me proporcionaram durante todo o processo. A meu pai, Antonio, e minha mãe, Selma, pelo incentivo constante e pela confiança em minha capacidade, mesmo nos momentos mais desafiadores.

E a minha amada namorada, Mayara, por sua compreensão, paciência e apoio emocional, que foram essenciais para manter meu equilíbrio e me incentivar nessa jornada.

Por último, mas não menos importante, agradeço aos meus colegas de MBA pela colaboração, pelas discussões enriquecedoras e pelo companheirismo durante nossos estudos. Essa experiência contribuiu imensamente para o meu crescimento pessoal e profissional. A todos que direta ou indiretamente fizeram parte desta jornada, cujos nomes não foram mencionados, mas cujas contribuições foram igualmente valiosas, minha sincera gratidão.

Este trabalho não seria o mesmo sem o apoio e encorajamento de cada um de vocês.

RESUMO

O setor de energia elétrica enfrenta desafios significativos relacionados à volatilidade dos preços, complexidade dos mercados e a necessidade de prever tendências de longo prazo. Explorar a aplicação do Expoente de Lyapunov, Expoente de Hurst e Entropia na análise dos preços de energia elétrica, fornece uma visão da complexidade e dinâmicas do mercado e da formação do preço da energia elétrica. O Expoente de Lyapunov revela a presença de comportamento caótico, destacando a imprevisibilidade dos preços. O Expoente de Hurst avalia a persistência ou reversão à média nos preços, informando estratégias de previsão e tendências de longo prazo. A Entropia quantifica a complexidade, indicando a variabilidade e estabilidade dos dados. Essas métricas capacitam tomadores de decisão no setor elétrico, oferecendo insights para a gestão de riscos, estratégias e compreensão do mercado de energia elétrica.

Palavras-chave: Expoente de Lyapunov, Expoente de Hurst, Entropia, Preço de Energia.

ABSTRACT

The electric power sector faces significant challenges related to price volatility, market complexity and the need to predict long-term trends. Exploring the application of the Lyapunov Exponent, Hurst Exponent and Entropy in the analysis of electricity prices provides an insight into the complexity and dynamics of the market and the formation of the price of electricity. The Lyapunov Exponent reveals the presence of chaotic behavior, highlighting the unpredictability of prices. The Hurst Exponent assesses persistence or mean reversion in prices, informing forecasting strategies and long-term trends. Entropy quantifies complexity, indicating data variability and stability. These metrics empower decision makers in the electricity sector, offering insights for risk management, strategies and understanding of the electricity market.

Keywords: Lyapunov Exponent, Hurst Exponent, Entropy, energy price.

LISTA DE FIGURAS

Figura 1 Retorno semanal do Preço de Energia, Câmbio e IBOVESPA	13
Figura 2 Matriz elétrica brasileira em 2019 e 2023, valores em MW e %	16
Figura 3 Complementariedade das fontes de geração [2]......	17
Figura 4 Fluxo temporal dos modelos de formação de preço [2]	18
Figura 5 Formação da Função de Custo Futuro do modelo Newave Fonte: [1].....	19
Figura 6 Gráfico de preços de energia elétrica, horário e semanal.	21
Figura 7 Histograma dos preços de energia elétrica, horário e semanal.....	21
Figura 8 Box-Plot dos preços de energia elétrica, horário e semanal.....	22
Figura 9 Mapa logístico para $r = 2,8$ e $r = 3,3$, com $x_0 = 0,2$	26
Figura 10 Mapas logísticos para $r = 4$ e com $x_0 = 0,2$ e $x_0 = 0,2 + 0,001$	27
Figura 11 Diagrama de bifurcação do mapa logístico	28
Figura 12 Diferença de duas trajetórias com condições iniciais diferentes em 0,001, para $r = 2,8$ (acima) e $r = 4$ (abaixo)	29
Figura 13 Diagrama de bifurcação e Expoente de Lyapunov vs r.....	31
Figura 14 Expoente de Lyapunov vs número de iterações n	32
Figura 15 Expoente de Lyapunov Esperado e calculado pelo método de Rosenstein	33
Figura 16 Expoente de Lyapunov Esperado e pelo método de Rosenstein, para diferentes valores de n	34
Figura 17 Expoente de Hurst Esperado e Diagrama de Bifurcação, para o mapa logístico	36
Figura 18 Entropia Amostral (SampEn) e Diagrama de Bifurcação, para o mapa logístico	38
Figura 19 Histogramas do expoente de Lyapunov para 5000 subséries, de tamanho 2190 para os preços horários e tamanho 52 para os preços semanais	40
Figura 20 Expoente de Hurst vs número de registros n.....	41
Figura 21 Histogramas do expoente de Hurst para 5000 subséries, de tamanho 2190 para os preços horários e tamanho 52 para os preços semanais	42
Figura 22 Histogramas da Entropia para 5000 subséries, de tamanho 2190 para os preços horários e tamanho 52 para os preços semanais.....	43
Figura 23 Tabela resumo dos valores médios dos indicadores analisados, para a série total e 5000 subséries	45

LISTA DE ABREVIATURAS E SIGLAS

ONS	Operador Nacional do Sistema
CCEE	Câmara de Comercialização de Energia Elétrica
SIN	Sistema Interligado Nacional
PLD	Preço de Liquidação das Diferenças
CMO	Custo Marginal de Operação
MWh	Megawatt-hora
MW	Megawatt
SampEn	Sample Entropy

SUMÁRIO

1 INTRODUÇÃO.....	12
2 MERCADO DE ENERGIA ELÉTRICA BRASILEIRO	15
2.1 A matriz de geração e o planejamento da operação	15
2.2 Metodologia de determinação dos preços de energia	17
2.3 Comportamento dos dados históricos do preço de energia	20
3 CAOS EM SISTÊMAS DINÂMICOS E METODOLOGIA.....	24
3.1 Sistemas dinâmicos	24
3.2 Mapa logístico	25
3.3 Caos em sistemas dinâmicos	29
3.4 Exponente de Lyapunov.....	30
3.5 Exponente de Hurst	35
3.6 Entropia Amostral.....	37
4 RESULTADOS E DISCUSSÃO.....	39
5 CONCLUSÕES.....	45
REFERÊNCIAS	46
ANEXO A – CÓDIGOS COMPUTACIONAIS	48

1 INTRODUÇÃO

O Brasil possui um sistema de geração elétrica que predominantemente utiliza seus recursos hídricos. Nos últimos anos, em média, 74% da geração de energia elétrica corresponderam a geração em aproveitamentos hidroelétricos. Há, portanto, necessidade de uma cuidadosa coordenação da operação, tanto para que o sistema seja eletricamente seguro quanto para que os recursos sejam aproveitados de forma eficiente (ONS e CCEE).

A princípio, pode-se imaginar que a água armazenada nos reservatórios não possui valor associado, logo seria mais vantajoso atender a demanda utilizando primeiramente a água estocada nos reservatórios evitando o acionamento de usinas termoelétricas, que possuem um custo direto associado ao seu combustível. Porém, considerando que a capacidade de armazenamento em um sistema é limitada pela capacidade de seus reservatórios, e que a afluência futura aos aproveitamentos é desconhecida, pode-se dizer que há uma dependência entre a decisão operativa de hoje e os custos operativos do futuro (ONS, 2020).

Ainda segundo o Operador Nacional do Sistema - ONS, o planejamento da operação tem como objetivo a determinação ótima dos recursos hídricos e térmicos seguindo o critério de minimização dos custos globais de produção de energia elétrica associado ao atendimento da demanda e preservação dos reservatórios no longo prazo.

Para solução desse problema de otimização complexo foram desenvolvidos pelo Centro de Pesquisa de Energia Elétrica – CEPEL modelos computacionais que reduzem o problema em etapas menores e entregam o preço de energia, em R\$/MWh, para diferentes horizontes de tempo (mensal, semanal e horário).

Para visualizar o comportamento dos preços de energia ao longo do tempo, comparou-se na Figura 1 os retornos semanais do preço da energia no Brasil com os retornos semanais do dólar e IBOVESPA, de janeiro de 2010 a dezembro de 2023. É possível observar a grande volatilidade do preço de energia elétrica, até pela mudança da escala de seu gráfico. Enquanto dólar e IBOVESPA tem retornos semanais médios 0,2% e 0,1%, respectivamente, o preço de energia tem retorno médio 3,5%. E o desvio padrão do retorno semanal fica 2,1% para o câmbio, 3,1% para o IBOVESPA e o preço de energia elétrica 25,9%.

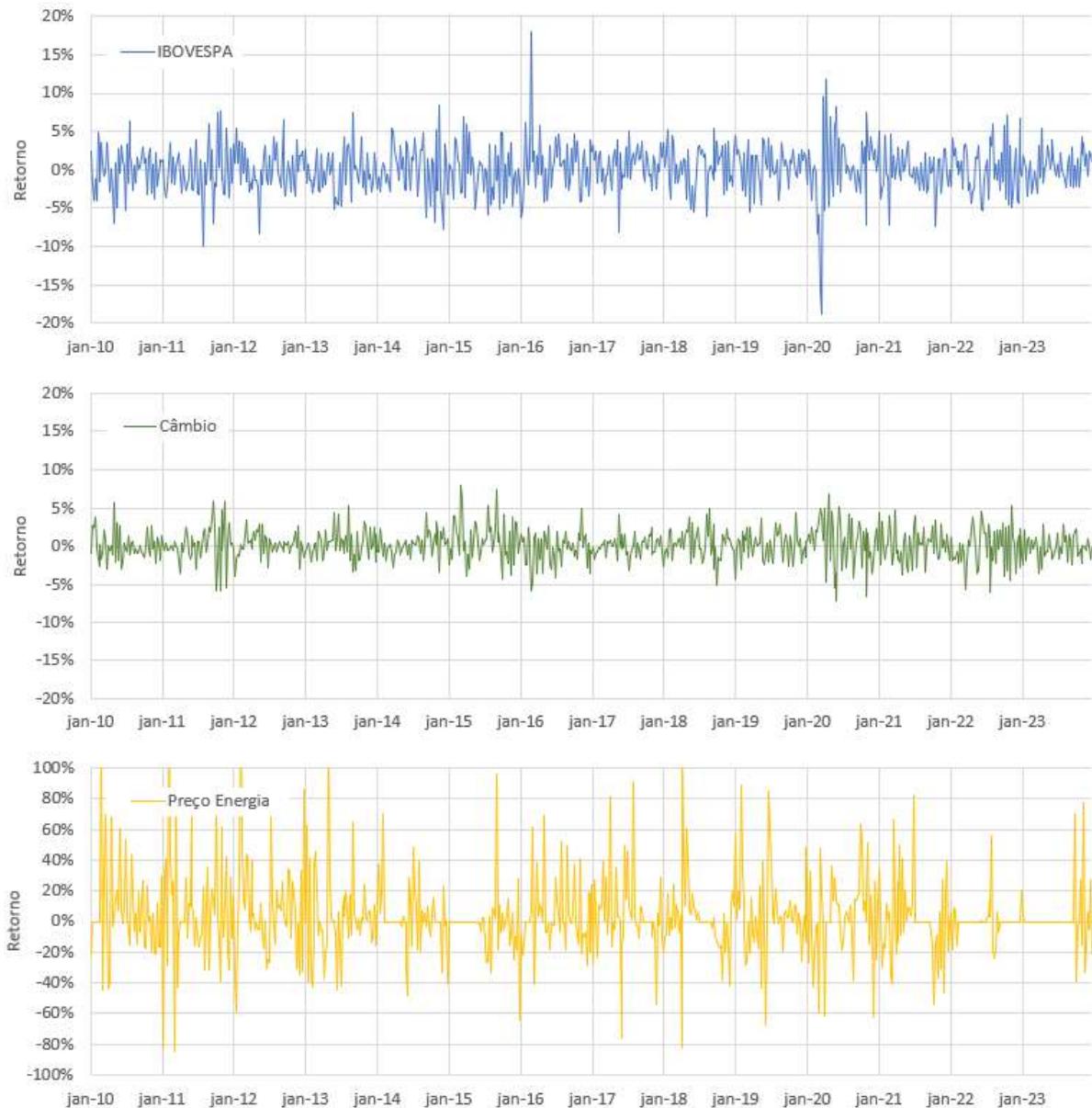


Figura 1 Retorno semanal do Preço de Energia, Câmbio e IBOVESPA

Esse comportamento volátil do preço da energia no Brasil tem influência da composição da matriz da geração brasileira e deve-se considerar que o sistema elétrico é constantemente adaptado às novas exigências de padrões de consumo e/ou padrões de geração. Entender a dinâmica do comportamento de determinação do preço de energia tem por objetivo a compreensão da correta sinalização econômica aos agentes que participam do mercado de energia.

E a teoria dos sistemas dinâmicos pode auxiliar nesta análise de sinalização econômica do preço de energia. Para este trabalho, escolheu-se apresentar uma introdução ao

comportamento caótico de sistemas dinâmicos pela análise do Preço de Liquidação das Diferenças - PLD, de acordo com três métricas:

- O coeficiente de Lyapunov que mede a taxa com que trajetórias próximas no espaço de fase de um sistema dinâmico divergem com o tempo, como explicado em Alligood (1996). A ideia de sua aplicação é poder entender quão sensível os preços são a pequenas mudanças, sugerindo um potencial de volatilidade;
- O Exponente de Hurst que mede a persistência ou memória de longo prazo de uma série temporal, como explicado em Mandelbrot e Wallis (1969). Com este conceito será avaliado o comportamento futuro do preço, indicando se seguirão tendências passadas ou mudarão de direção;
- Entropia estima a desordem ou complexidade de um sistema, conforme Pincus (1990). Pode ser aplicada nos preços de energia para entender o risco associado aos preços de energia;

Este trabalho busca compreender a aplicação da teoria citada aos preços de energia, bem como sua utilidade para o auxílio da tomada de decisão em investimentos que envolvam o preço de energia.

No Capítulo 2 deste texto será trazido um breve histórico do setor de energia brasileiro e a métrica adotada para definição dos preços de energia elétrica, abordando as principais características que explicam o comportamento dos preços.

O Capítulo 3 abordará a revisão bibliográfica das métricas apresentadas e quais metodologias serão utilizadas para a mensuração dos indicadores da série temporal dos preços de energia, utilizando como exemplo o mapa logístico.

Posteriormente, serão apresentados no Capítulo 4 os resultados e discussão da aplicação do expoente de Lyapunov, expoente de Hurst e entropia na série de preços de energia, no Capítulo 4 e no Capítulo 5 será feita a citação de trabalhos futuros a serem realizados.

Veremos, então, a seguir as características da formação de preço de energia.

2 MERCADO DE ENERGIA ELÉTRICA BRASILEIRO

Diferentemente da maioria dos países que passaram por reformas no desenho do mercado de eletricidade na década de 1990, o Brasil optou pela formação de preços a partir de custos auditados. Nesse modelo, os agentes declaram sua disponibilidade e seus custos de produção, que são utilizados por um operador central para despachar o sistema de forma ótima. Para tanto, o operador dispõe de modelos computacionais de otimização, que calculam o “despacho ótimo” com base em parâmetros de risco pré-estabelecidos, restrições de transmissão, dados do sistema e expectativas de comportamento futuro — de carga e, no caso brasileiro, de hidrologia. O custo marginal de geração determina também o preço do mercado de curto prazo (Castro; Borges e Simone, 2023).

Neste capítulo será apresentado um histórico da formação de preço do Brasil, desde a composição de nossa matriz até a metodologia adotada atualmente para definição do preço de energia elétrica.

2.1 A matriz de geração e o planejamento da operação

Como publicado no livro “O planejamento da operação energética no sistema interligado nacional” (ONS, 2020), o Operador Nacional do Sistema - ONS relata que a partir de meados da década de 1960 o crescimento do consumo de energia elétrica trouxe uma necessidade da expansão da capacidade instalada geradora do Brasil e da malha de transmissão. Isso exigiu que os sistemas elétricos brasileiros funcionassem de maneira integrada, para aproveitamento racional das fontes energéticas disponíveis e buscando um equilíbrio econômico.

Ainda segundo o ONS em seu livro, em razão de suas características naturais, o Brasil desenvolveu sua matriz energética com predominância da hidroeletricidade. O operador ainda ressalta que apesar do predomínio hidrelétrico, a recente expansão do Sistema Interligado Nacional - SIN tem sido marcada pelo crescimento na participação de novas formas de geração, com destaque para as fontes eólica e fotovoltaica, e de transmissão, especialmente a utilização mais intensa da transmissão em corrente contínua. Todos esses movimentos, de acordo com os autores, aumentam a complexidade do sistema e o desafio do ONS de garantir que a operação do SIN se dê da forma mais segura e econômica, requisito cada vez mais essencial à medida que avança na sociedade brasileira a sua dependência da energia elétrica.

Conforme Figura 2, os dados disponíveis no *site* do ONS¹ demonstram que a expansão recente da matriz brasileira vem ocorrendo pelas fontes renováveis. De 2019 a 2023, a participação eólica e solar passou de 11% para 18%, sendo que outros 12% de Micro e Mini Geração Distribuída – MMGD são predominantemente solar.

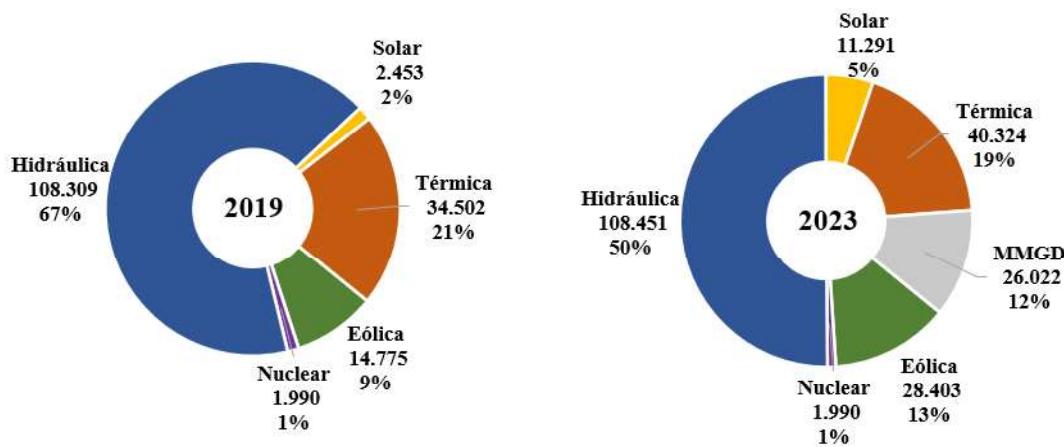


Figura 2 Matriz elétrica brasileira em 2019 e 2023, valores em MW e %

O fato da expansão de outras fontes, além da hidroelétrica e térmica, é importante devido a complementariedade da geração. Essa diversificação traz maior segurança de suprimento para a carga. Esse fato é apresentado pelo ONS (2020) e pode ser constatado na Figura 3, que mostra a complementariedade das fontes eólica, solar, hidrelétrica e PCT (Pequenas Centrais Térmicas, por exemplo Biomassa).

¹ <https://www.ons.org.br/paginas/sobre-o-sin/o-sistema-em-numeros>

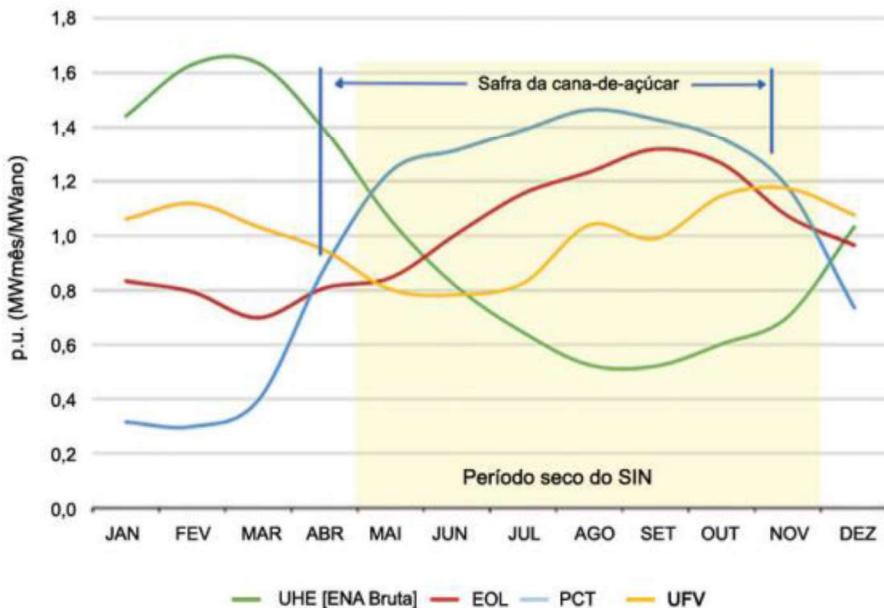


Figura 3 Complementariedade das fontes de geração [2].

Enquanto o planejamento da expansão visa atender de maneira otimizada o consumo de energia elétrica e fornecer sinais para minimizar os custos totais futuros para a sociedade, o objetivo principal de cada um dos agentes que participam desse setor é a maximização de seus resultados.

Cada agente privado desenvolve seu plano de expansão empresarial, com objetivos que podem ser bastante distintos daqueles do planejamento governamental. Ao governo cabe buscar as utilizações adequadas, racionais e otimizadas dos recursos naturais.

Para solucionar esse problema, cabe ao Operador do sistema fornecer o justo sinal econômico. Então, segundo o ONS em seu livro sobre o planejamento da operação, o problema de planejamento energético do sistema brasileiro consiste em determinar metas de geração para as usinas hidroelétricas e termoelétricas para cada estágio de tempo ao longo do horizonte de estudo, atendendo à demanda de energia elétrica, às restrições operativas das usinas e às restrições elétricas do sistema.

2.2 Metodologia de determinação dos preços de energia

A determinação da alocação ótima dos recursos hídricos e térmicos é norteada segundo um determinado critério, como o critério de minimização dos custos globais de produção da energia elétrica. A princípio, pode-se imaginar que a água armazenada nos reservatórios não possui valor associado, logo seria mais vantajoso atender a demanda utilizando primeiramente

a água estocada nos reservatórios evitando o acionamento de usinas termoelétricas, que possuem um custo direto associado ao seu combustível. Porém, considerando que a capacidade de armazenamento em um sistema é limitada pela capacidade de seus reservatórios, e que a afluência futura aos aproveitamentos é desconhecida, pode-se dizer que há uma dependência entre a decisão operativa de hoje e os custos operativos do futuro (ONS e CCEE).

A estratégia do Planejamento da Operação resume-se em tomar decisões no presente para que os reservatórios tenham como se manter em níveis seguros para atendimento da carga no futuro e que garantam o menor custo.

Como mencionado em seu manual de treinamento (ONS e CCEE), o ONS dispõe de uma cadeia de modelos computacionais para a otimização da operação e minimização do custo de atendimento à carga. Estes modelos são desenvolvidos pelo Centro de Pesquisas de Energia Elétrica – CEPEL. O modelo NEWAVE é utilizado para projeção de médio prazo (5 anos), enquanto os modelos DECOMP e DESSEM agregam mais detalhes das usinas e projetam os preços semanais e horários, respectivamente. A Figura 4, abaixo, detalha o fluxo temporal da cadeia de modelos. Cabe destacar que todos os programas computacionais de formação de preço passam por processos de validação no âmbito da Agência Nacional de Energia Elétrica - ANEEL, com a participação de agentes de geração, transmissão, distribuição e comercialização associados ao ONS e à CCEE.



Figura 4 Fluxo temporal dos modelos de formação de preço [2]

De acordo com ONS (2020) o planejamento da operação energética de médio e curto prazo é um problema de natureza essencialmente estocástica, dado que não temos o conhecimento perfeito da disponibilidade das afluências futuras, combustível da principal fonte geradora do SIN. Uma alternativa para resolver o problema estocástico é escrevê-lo na sua

forma equivalente determinística, onde as restrições e variáveis de todos os períodos e cenários são consideradas de uma só vez em um único problema de otimização determinística.

Devido ao porte e complexidade desse problema, é necessária a sua divisão em diferentes etapas, abrangendo estudos nos horizontes de médio prazo, curto prazo e programação diária (ONS, 2020). Pode-se então estudar o comportamento do sistema com diversos cenários de vazões. E essas vazões afluentes, ainda seguindo o Operador do Sistema, são representadas de forma equivalente, com o ajuste de modelos estocásticos PAR(p) (autorregressivos periódicos) e a utilização de uma modelagem probabilística na simulação das vazões afluentes futuras. Dessa forma, através da Programação Dinâmica Dual Estocástica, o modelo NEWAVE é responsável pela construção da Função de Custo Futuro (FCF), que valora a água armazenada nos reservatórios do SIN em cada estágio de simulação, trazendo esse custo ao valor presente. A Figura 5, extraída do livro do ONS (2020), ilustra esse processo. Sendo a linha mais grossa a média dos cenários futuros que resulta na formação dos custos de energia.

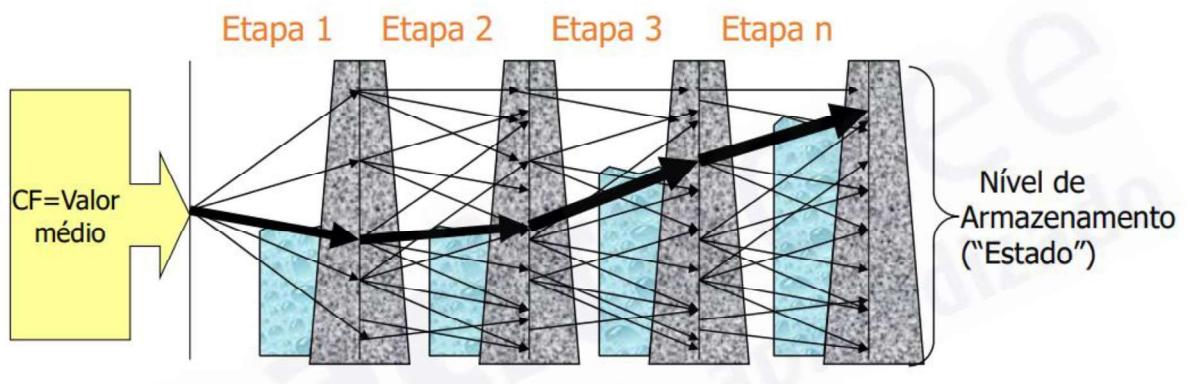


Figura 5 Formação da Função de Custo Futuro do modelo Newave Fonte: [1]

Ainda segundo o ONS (2020) o DECOMP realiza uma otimização linear para o despacho hidrotérmico, considerando restrições operativas e de transmissão do sistema. O modelo busca minimizar o custo total da operação, incluindo custos de geração térmica e penalidades por não atendimento da demanda, respeitando limites operativos das usinas e linhas de transmissão. E o DESSEM leva em conta as condições operativas do sistema em intervalos de tempo curtos (hora a hora). As equações consideram a demanda de carga, geração termelétrica, intercâmbios energéticos entre regiões, e outras restrições técnicas.

O problema a ser resolvido, então, é a otimização do atendimento do curto prazo e a preservação dos reservatórios no longo prazo. Isto é, depende de decisões imediatas e decisões

futuras. Segundo o ONS (2020), o problema de despacho ótimo resolvido que faz parte do planejamento da operação energética pode ser escrito de forma sucinta, de acordo com a equação de minimização descrita. Para resolver este problema, utiliza-se métodos de programação linear que implica em minimizar o custo total da operação considerando as restrições de balanço de energia, atendimento a demanda, função de produção, hidráulicas, elétricas, de custo futuro e de caixa.

Os modelos de formação de preço de energia, então, são fundamentais para o funcionamento eficiente do mercado de energia elétrica. Eles são projetados para refletir os custos de produção, transmissão e distribuição de energia, bem como para incorporar fatores como a demanda de mercado, a disponibilidade de recursos, condições ambientais, e políticas regulatórias. A seguir será apresentado o comportamento do preço histórico de energia elétrica, horário e semanal.

2.3 Comportamento dos dados históricos do preço de energia

Como apresentado anteriormente, a precificação da energia elétrica é influenciada por uma série de variáveis de entrada, cada uma contribuindo de maneira única para a determinação dos preços finais. Entre estas variáveis, destacam-se fatores como a disponibilidade de recursos naturais (por exemplo, água para hidrelétricas, vento para turbinas eólicas), custos de produção, demanda do mercado, políticas governamentais e restrições elétricas. A disponibilidade de recursos naturais é particularmente crítica, pois influencia diretamente a capacidade de geração de energia.

Por exemplo, em períodos de seca, a capacidade das hidrelétricas pode ser significativamente reduzida, levando a um aumento no preço da energia devido à escassez de oferta. A Figura 6 apresenta o comportamento do preço da energia elétrica ao longo do histórico analisado, nos formatos horário e semanal.

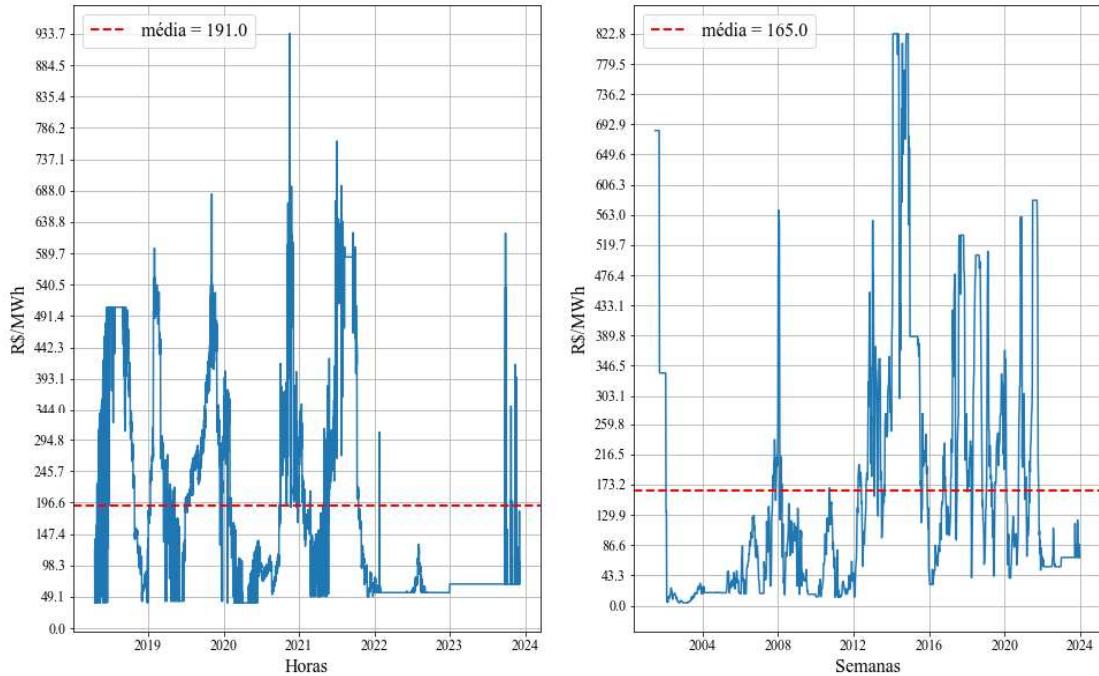


Figura 6 Gráfico de preços de energia elétrica, horário e semanal.

Observando ainda a Figura 7, dos histogramas de preços horários e semanais, vemos uma elevada frequência de observações próximas aos intervalos de menores preços de energia. A série horária tem mediana R\$ 104,31/MWh e a série semanal mediana R\$ 87,51/MWh.

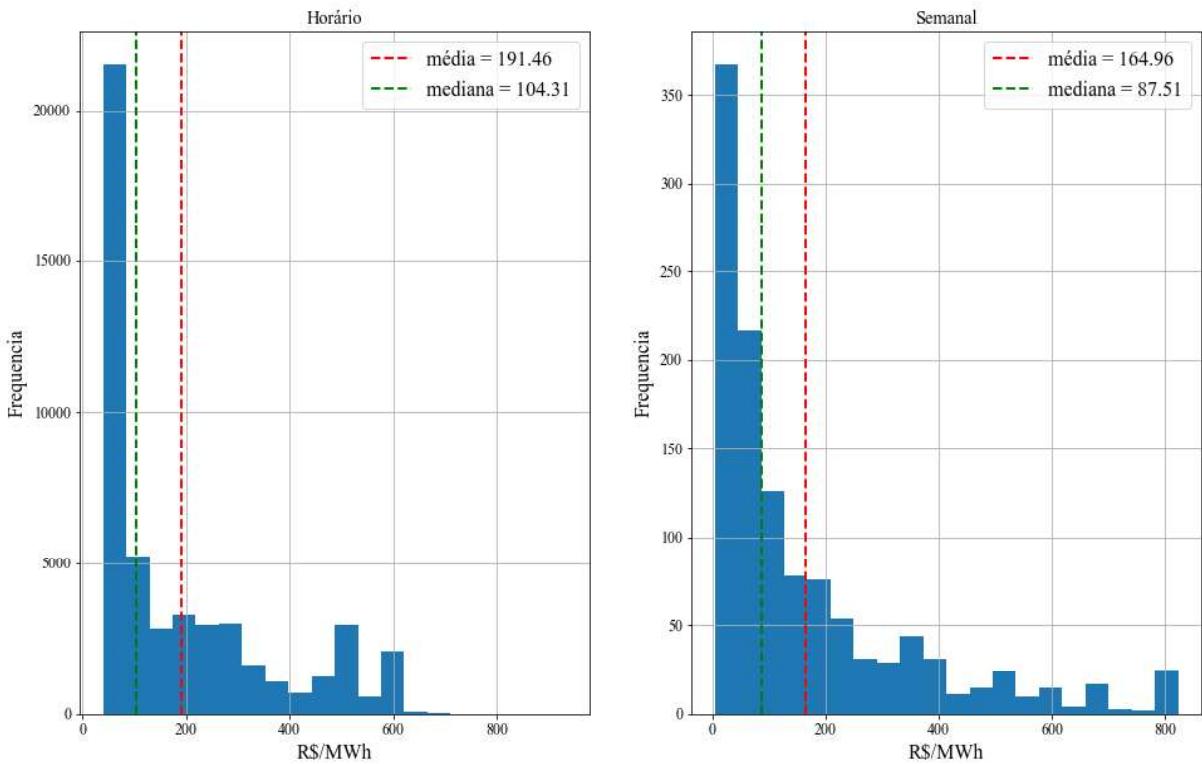


Figura 7 Histograma dos preços de energia elétrica, horário e semanal

Uma análise Box-Plot, Figura 8, nos permite visualizar que para os preços horários e semanais, 75% dos dados ficam abaixo de R\$ 275,86/MWh e R\$ 220,52/MWh, respectivamente. A mediana dos dados é R\$ 104,31/MWh e R\$ 87,51/MWh. E o intervalo entre o primeiro e terceiro quartis é R\$ 206,82/MWh e R\$ 190,03/MWh, a diferença relevante entre os quartis mostra que o preço de energia tem um comportamento volátil.

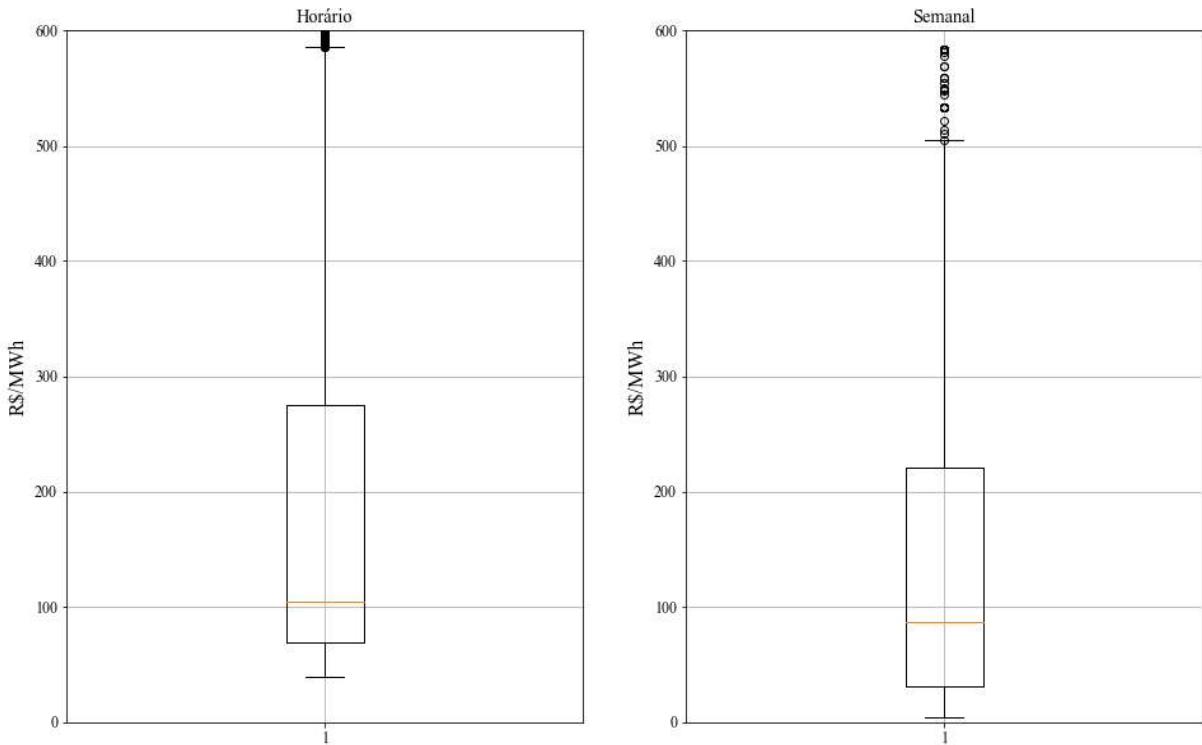


Figura 8 Box-Plot dos preços de energia elétrica, horário e semanal

Em virtude das características do sistema elétrico brasileiro — com elevada participação de usinas hidrelétricas — alguma variabilidade nos preços é esperada, uma vez que as afluências futuras são desconhecidas e os modelos meteorológicos são imprecisos. A volatilidade dos preços na indústria de energia elétrica é um fator de risco percebido tanto para os geradores quanto para os consumidores. Um gerador corre o risco de que o preço seja muito baixo para cobrir todos os custos. Um consumidor corre o risco de que, às vezes, o preço da eletricidade seja muito alto em comparação com os benefícios que ela cria (Castro; Borges e Simone, 2023).

Em suma, a análise do histórico do setor de energia do Brasil, com suas características únicas de grande participação hidrelétrica, nos leva a entender a metodologia de formação de preços. Os modelos computacionais de formação de preços são responsáveis pela sinalização econômica ao mercado e seus agentes, refletindo os custos de geração, condições climáticas e

políticas regulatórias. O resultado disso é, como observado, que os preços de energia elétrica apresentam variações significativas em intervalos horários e semanais.

No próximo capítulo deste trabalho será analisada a metodologia escolhida para determinar quão caótico pode ser o comportamento dos preços de energia.

3 CAOS EM SISTÊMAS DINÂMICOS E METODOLOGIA

Neste capítulo, exploraremos as metodologias relacionadas a três importantes conceitos de análise de sistemas não lineares utilizados na análise de séries temporais: o Expoente de Hurst, a Entropia Amostral e o Expoente de Lyapunov. Este capítulo não apenas detalha a teoria por trás desses conceitos, mas também discute as metodologias utilizadas para aplicações em séries reais, além de uma contextualização da dinâmica dos sistemas, do mapa logístico e do caos.

Inicialmente, será apresentado o conceito básico de um sistema dinâmico.

3.1 Sistemas dinâmicos

Um sistema dinâmico consiste em um conjunto de possíveis estados, junto com uma regra que determina o estado presente em termo dos estados passados (Alligod, 1996). Ainda de acordo com o autor, esse tipo de sistema é determinístico, significando que podemos determinar o estado presente unicamente de acordo com os estados passados.

Existem dois tipos principais de sistemas dinâmicos: equações diferenciais e mapas iterados (também conhecidos como equações de diferenças). As equações diferenciais descrevem a evolução dos sistemas em tempo contínuo, enquanto os mapas iterados surgem em problemas onde o tempo é discreto (Strogatz, 2015).

E a compreensão dos sistemas passa pelo conceito de estabilidade, conforme apresentada por Boyce e Di Prima (2009), envolve a análise do comportamento das soluções de equações diferenciais em relação ao tempo. Os pontos onde a expressão tem resultado nulo são chamados muitas vezes de soluções de equilíbrio ou pontos críticos (fixos). Ainda segundo os autores, a equação ainda segue um caminho, ou trajetória, e o plano com todas as fases da trajetória é o plano de fases.

Um ponto fixo estável tem a propriedade de que os pontos próximos a ele se movam ainda mais perto do ponto fixo no sistema dinâmico. Para um ponto fixo instável, os pontos próximos se afastam com o passar do tempo. Uma boa analogia é que uma bola no fundo de um vale é estável, enquanto uma bola equilibrada no topo de uma montanha é instável. (Alligood, 1996)

De acordo com Alligood (1996) a quantificação da trajetória do sistema, para análise de sua estabilidade e compreensão de suas propriedades, é seu mapa. Que pode ser entendido como uma função cujas entradas e saídas pertencem ao mesmo domínio.

‘Seja x um ponto e f um mapa. A órbita de x em f é o conjunto de pontos $\{x, f(x), f^2(x), \dots\}$. O ponto de início x da órbita é chamado de valor inicial (ou condição inicial) da órbita. Um ponto p é um ponto fixo do mapa f se $f(p)=p$.’

Boyce e Di Prima (2009) ainda destacam que o objetivo principal da caracterização das equações de acordo com o padrão geométrico formado por suas trajetórias é fundamental na teoria qualitativa de equações diferenciais.

E para Alligood (1996) questão da estabilidade é significativa porque um sistema do mundo real está constantemente sujeito a pequenas perturbações. Portanto, um estado estacionário observado num sistema realista deve corresponder a um ponto fixo estável. Se o ponto fixo for instável, pequenos erros ou perturbações no estado fariam com que a órbita se afastasse do ponto fixo, o que então não seria observado.

Portanto, pontos fixos e a estabilidade são conceitos fundamentais no estudo da dinâmica de sistemas. Pontos fixos são estados em que o sistema permanece constante ao longo do tempo e a estabilidade desses pontos é crucial para entender como o sistema responde a perturbações.

Alguns exemplos são conhecidos para compreensão da dinâmica de sistemas não-lineares, dentre eles está o mapa logístico, conceito apresentado a seguir.

3.2 Mapa logístico

Um exemplo popular para explorar as características dos sistemas não lineares é o mapa logístico, desenvolvido por Robert May (1976) para análise do crescimento populacional de insetos

$$x_{n+1} = rx_n(1 - x_n) \quad (1)$$

com os valores do parâmetro $r \in [0, 4]$ e com $x_n \in [0, 1]$.

Essa função tem comportamento estável para determinados valores de r . Com órbitas limitadas e periódicas ou convergentes. Conforme apresentado por Strogatz (2015), para $1 < r$

$r < 3$, a população cresce e eventualmente atinge um estado estacionário diferente de zero, conforme ilustrado na Figura 9 quadro superior.

Ainda de acordo com Strogatz, para r maiores, por exemplo $r = 3,3$, a população aumenta novamente, mas agora oscila em torno de um estado estacionário, alternando entre uma grande população numa geração e uma população menor na seguinte, conforme apresentado na Figura 9 quadro inferior.

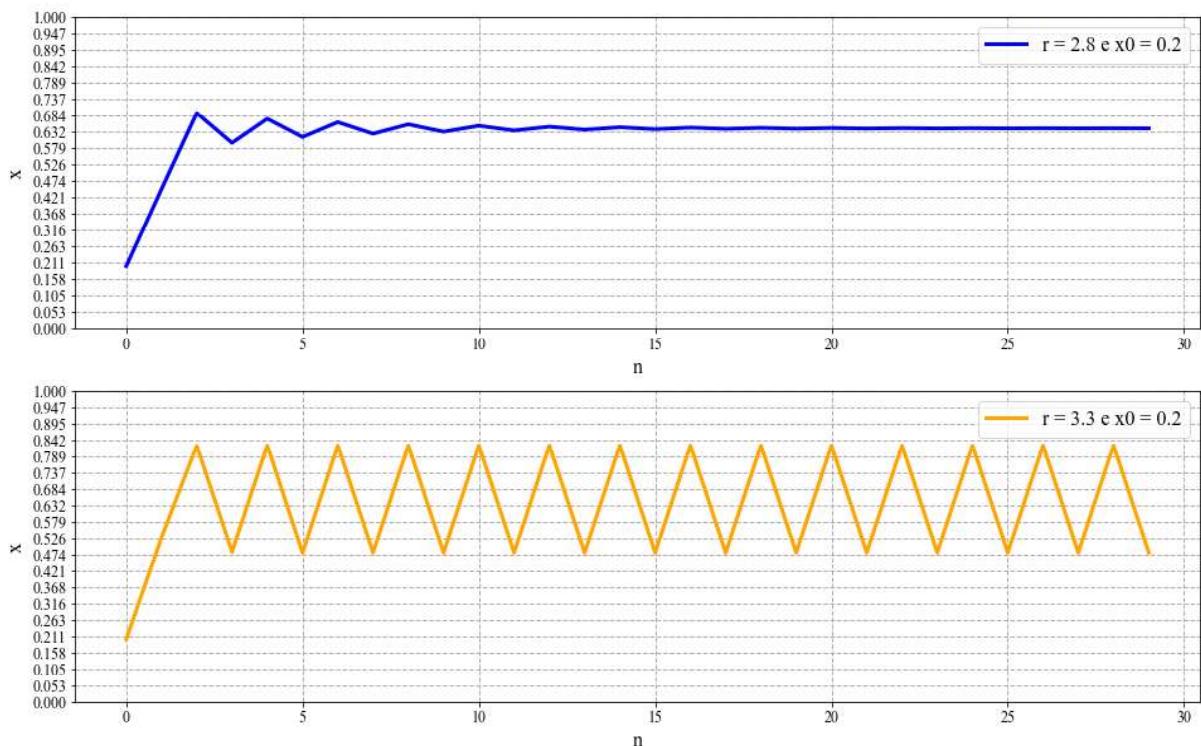


Figura 9 Mapa logístico para $r = 2,8$ e $r = 3,3$, com $x_0 = 0,2$

Mas é para valores de r mais próximos de 4 que a função logística se mostra interessante no estudo de sistemas dinâmicos não-lineares, pelo fato do comportamento do sistema mudar radicalmente. A Figura 10 apresenta a evolução de duas trajetórias do mapa logístico com $r = 4$ mas com condições iniciais diferentes em 0,001.

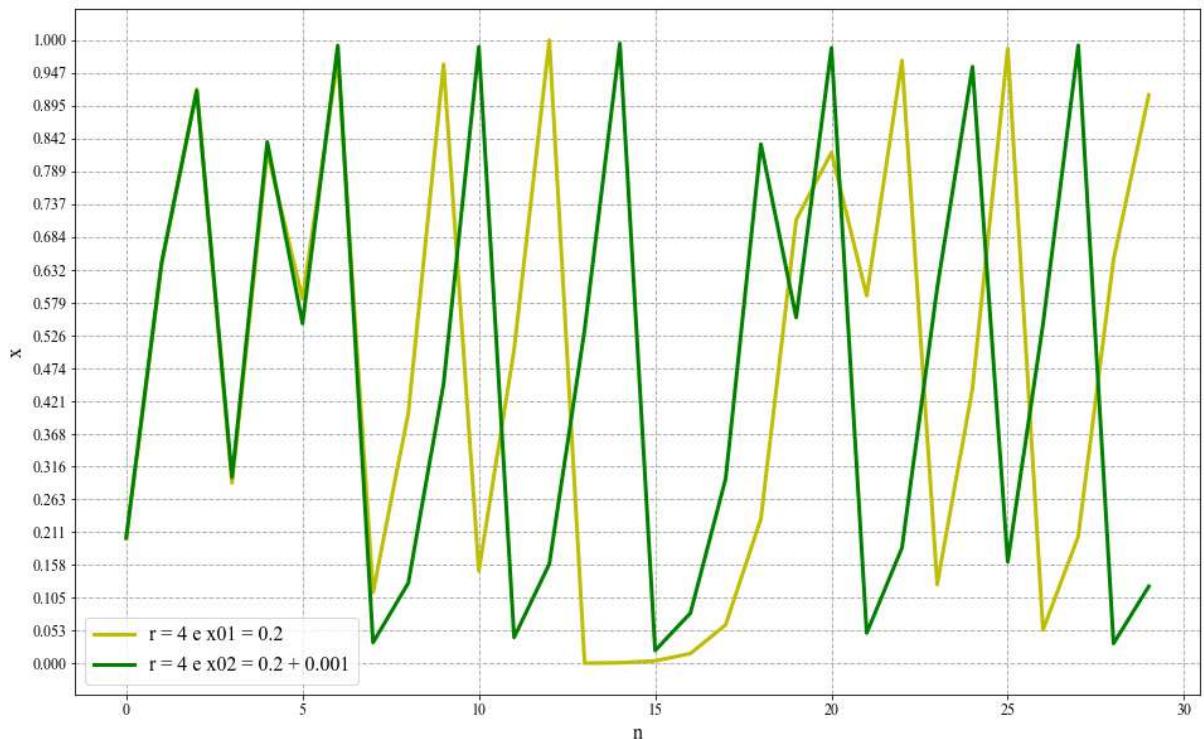


Figura 10 Mapas logísticos para $r = 4$ e com $x_0 = 0,2$ e $x_0 = 0,2 + 0,001$

Com $r = 4$, a equação logística apresenta uma órbita aperiódica, ou seja, iterando a equação o resultado não se repete. Isso significa que o sistema apresenta imprevisibilidade e demonstra o chamado “efeito borboleta” ou dependência sensível às condições iniciais (Richardson e Mitchell).

De forma mais generalizada e expandindo a análise, a alteração na dinâmica da função se apresenta de forma mais clara quando se simula, para diferentes valores de r , o valor de x_{n+1} depois de 1000 iterações, conforme a Figura 11.

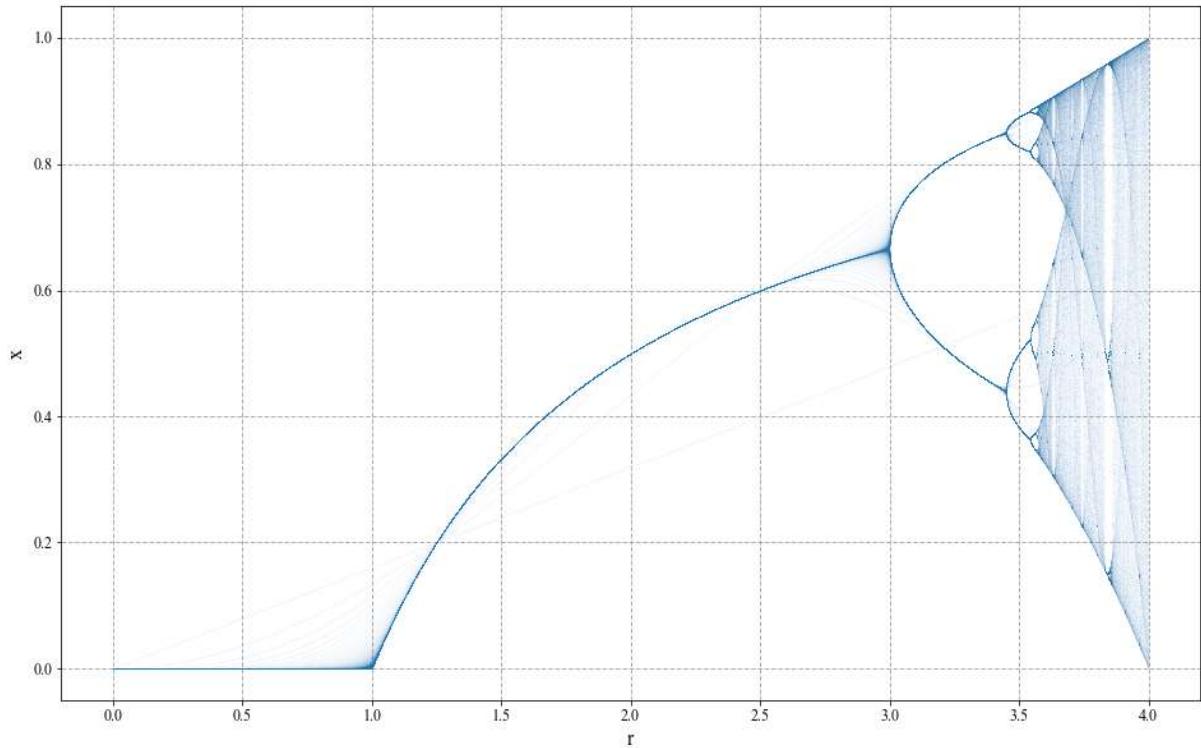


Figura 11 Diagrama de bifurcação do mapa logístico

De acordo com Strogatz (2015), em $r = 3,4$, o atrator é um ciclo de período 2, conforme indicado pelos dois ramos. À medida que r aumenta, ambos os ramos se dividem simultaneamente, produzindo um ciclo de período 4. Uma cascata de duplicações adicionais de período ocorre à medida que r aumenta, produzindo o período 8, o período 16 e assim por diante, até que em $r \approx 3,57$, onde o mapa se torna caótico e o atrator muda de um conjunto finito para um conjunto infinito de pontos.

O mapa logístico, simples em sua forma, revela comportamentos surpreendentemente complexos quando certos parâmetros são ajustados. O sistema transita de um comportamento previsível para um caótico, ilustrando como pequenas alterações podem levar a grandes mudanças na dinâmica do sistema.

Esta transição do previsível para o imprevisível no mapa logístico serve como um modelo fundamental para visualizar o caos em sistemas, conceito esse que será trazido em seguida.

3.3 Caos em sistemas dinâmicos

Sistemas caóticos são sistemas complexos que pertencem a classe dos sistemas dinâmicos determinísticos (Guegan, 2009). Esses sistemas são caracterizados pela forte não-linearidade e de acordo com Alligood (1996) uma órbita caótica é aquela que continua a experimentar para sempre o comportamento instável que uma órbita exibe perto de uma fonte, mas que não é fixa ou periódica. Em qualquer ponto de tal órbita, existem pontos arbitrariamente próximos que se afastarão do ponto durante a iteração adicional.

Essa característica pode ser observada na figura abaixo, onde se apresentam as diferenças de duas trajetórias com condições iniciais diferentes em 0,001. Na parte superior da Figura 12 observa-se que, para $r = 2,8$, à medida que as iterações ocorrem a diferença das trajetórias tende a zero. Enquanto no quadro inferior da figura, com as trajetórias tendo $r = 4$, à medida que as iterações ocorrem a diferença entre as trajetórias não apresenta um padrão.

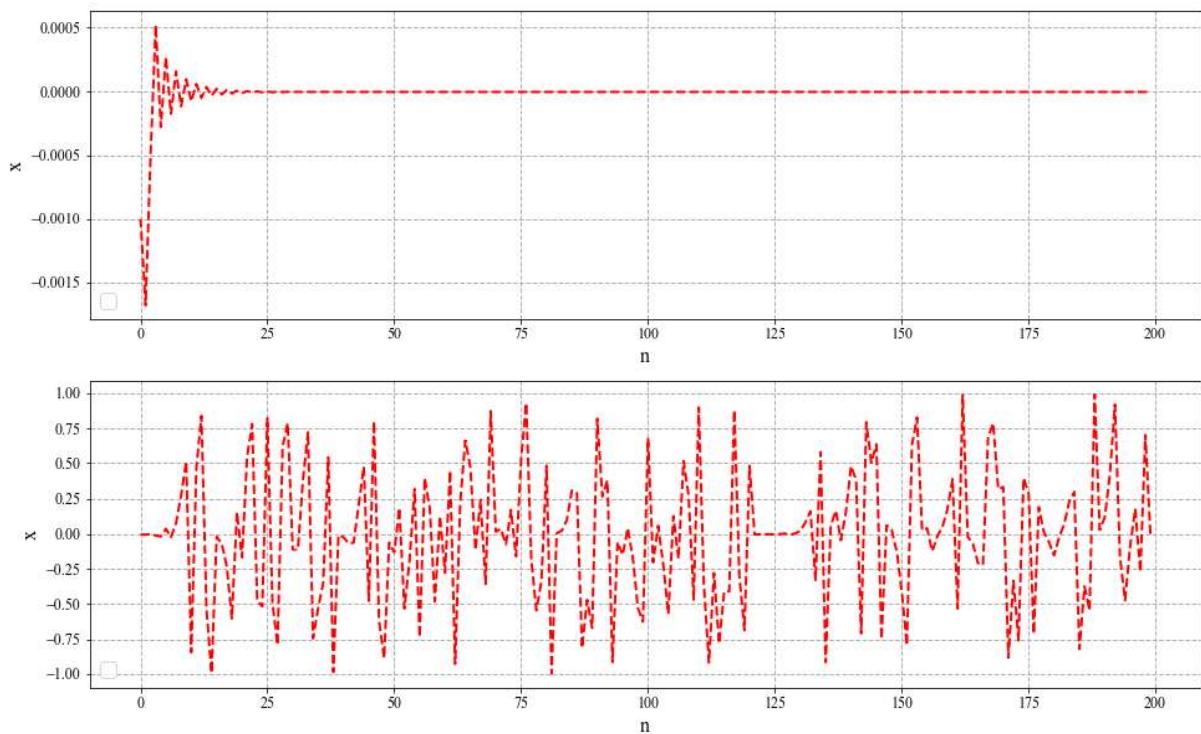


Figura 12 Diferença de duas trajetórias com condições iniciais diferentes em 0,001, para $r = 2,8$ (acima) e $r = 4$ (abaixo)

Segundo apresentado por Alligood et al (1996) um sinal é interpretado como tendo comportamento caótico se ele é limitado em suas órbitas, determinístico, as órbitas são aperiódicas e apresenta dependência sensível às condições iniciais.

O link da noção do caos com o comportamento estocástico é a base para aprofundar os estudos nesse campo, permitindo conectar a teoria a análises reais e extrair sinais de caos de conjuntos de dados.

Saber se um sistema ou um conjunto de dados apresenta comportamento caótico é uma maneira de melhorar as previsões, pois indica se os dados são sensíveis às condições iniciais e determina que métodos de modelagem mais avançados sejam empregados nas projeções, além de maior cuidado na tomada de decisão. E, para quantificar o caos em um sistema, utiliza-se o expoente de Lyapunov que mede a divergência exponencial das trajetórias em um espaço de fase.

3.4 Expoente de Lyapunov

Uma órbita caótica é aquela que sempre continua a experimentar um comportamento instável perto de uma fonte, mas essa órbita não é fixa ou periódica. ... Em qualquer ponto da órbita, existem pontos arbitrariamente próximos que se afastarão do ponto durante as próximas iterações. Essa irregularidade contínua é quantificada pelos números de Lyapunov e expoentes de Lyapunov (Alligod, 1996).

Ainda para o autor o número de Lyapunov é, portanto, a taxa média de divergência de duas órbitas ao longo das iterações e o expoente de Lyapunov é o logaritmo natural desse número. Uma função pode ser dita caótica se o expoente de Lyapunov for maior que zero.

De acordo com Alligood (1996) o expoente de Lyapunov é definido como o limite, quando o tempo tende ao infinito, do logaritmo natural da razão entre a distância final e a distância inicial entre duas trajetórias, dividido pelo tempo. Seja então f um mapa do domínio real \mathbb{R} . O Expoente de Lyapunov $h(x_1)$ da órbita $\{x_1, x_2, x_3, \dots\}$ é:

$$h(x_1) = \lim_{n \rightarrow \infty} \left(\frac{1}{n} \right) [\ln|f'(x_1)| + \dots + \ln|f'(x_n)|] \quad (3)$$

Exemplificando o expoente de Lyapunov através do mapa logístico, é possível calcular o expoente para diferentes valores de r , considerando 2000 iterações. A Figura 14 apresenta o resultado, paralelamente ao diagrama de bifurcação.

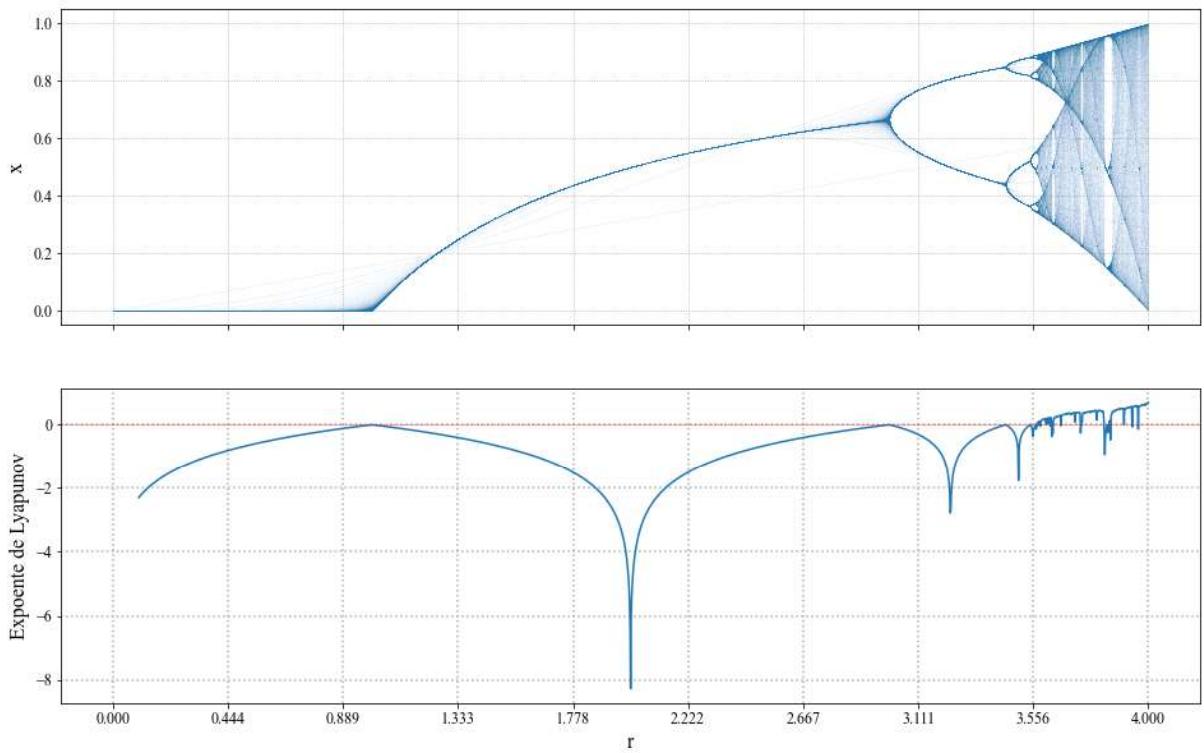


Figura 13 Diagrama de bifurcação e Expoente de Lyapunov vs r

Strogatz (2015) nos diz que, olhando o gráfico do expoente de Lyapunov vs r paralelamente ao diagrama de bifurcação, nota-se que λ permanece negativo para $r < 3,57$, e se aproxima de zero nas bifurcações de duplicação de período. Os picos negativos correspondem aos 2^n -ciclos. O início do caos é visível perto de $r \approx 3,57$, onde λ apresenta valores positivos. Ainda segundo o autor, para $r > 3,57$ o expoente de Lyapunov geralmente aumenta, exceto pelas quedas causadas pelas janelas de comportamento periódico.

O gráfico mostra em quais regiões a função exibe comportamento caótico com sensibilidade às condições iniciais (expoente acima de 0), regiões com comportamento periódico ou convergente (expoente abaixo de zero) e as regiões de transição do comportamento periódico para caótico (expoente igual a zero).

Outro ponto relevante para determinação do expoente de Lyapunov, conforme a própria fórmula 5 apresenta, é a quantidade de períodos para determinação de λ . A Figura 15, abaixo, mostra a evolução do expoente de Lyapunov a medida que mais iterações são adicionadas no cálculo.

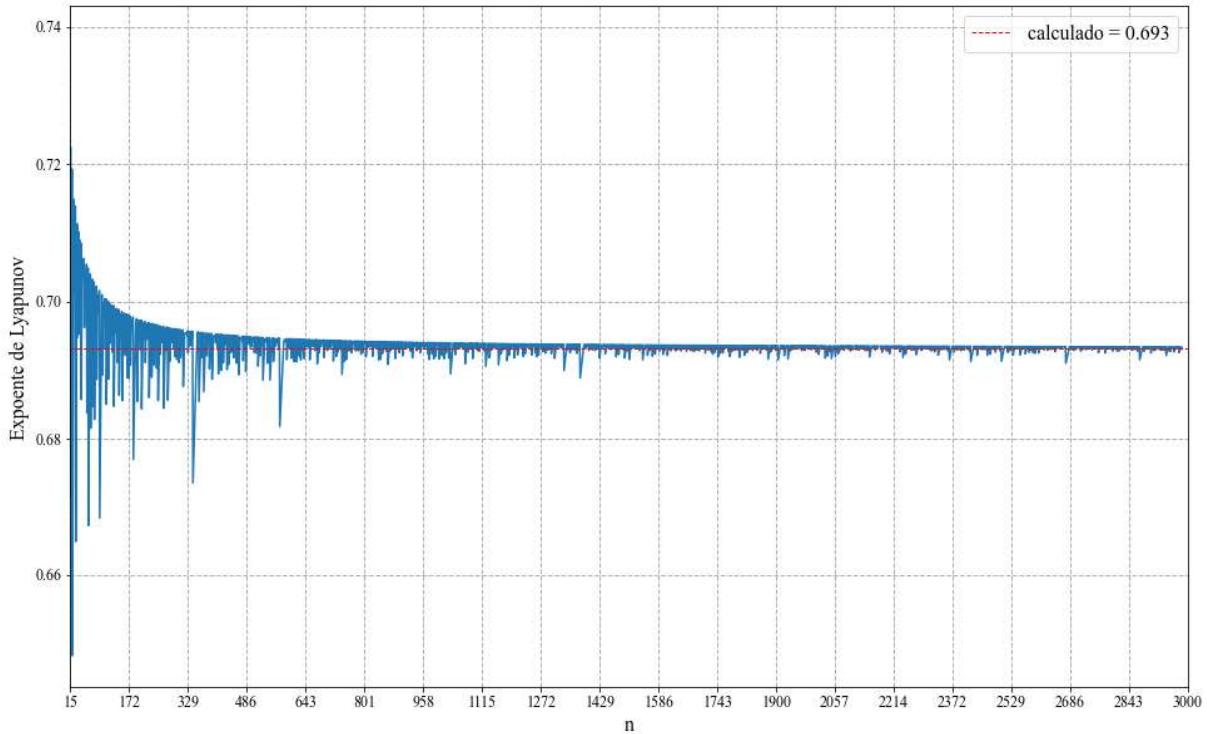


Figura 14 Expoente de Lyapunov vs número de iterações n

Para cada dimensão do espaço de fase de um sistema, existe um expoente de Lyapunov, formando o que se chama de espectro de Lyapunov. Frequentemente, o foco recai sobre o maior expoente de Lyapunov, pois este determina o comportamento geral do sistema. Se o maior expoente de Lyapunov é positivo, isso indica que o sistema é divergente, característico de um comportamento caótico. Se é negativo, isso corresponde a uma contração do espaço de fase, indicando que o sistema age de maneira estacionária ou periodicamente estável. O espectro ainda pode ser misto, indicando que algumas variáveis ou direções no sistema exibem sensibilidade às condições iniciais (característica do caos) e outras variáveis são estabilizadoras.

Para séries temporais produzidas por sistemas dinâmicos, a presença de um expoente positivo indica caos. Além disso, em muitas aplicações é suficiente calcular apenas o maior expoente de Lyapunov (Rosenstein, Collins e De Luca, 1993).

Para isso, será utilizado o algoritmo proposto por Rosenstein, Collins e De Luca. E a ideia do algoritmo é calcular o maior expoente de Lyapunov pela reconstrução da dinâmica da série temporal usando o método de incorporação de atraso (*delay embedding method*), que escolhe um atraso τ e uma dimensão de imersão m . A partir de uma série temporal $x(t)$, construa vetores de estado $Y(i)$ no espaço de fase m-dimensional:

$$Y(i) = [x(i), x(i + \tau), x(i + 2 * \tau), \dots, x(i + (m - 1) * \tau)] \quad (4)$$

para $i = 1, 2, \dots, N - (m - 1)\tau$, onde N é o número total de pontos de dados. Em seguida se calcula a distância Euclidiana entre cada par de pontos no espaço de fases reconstruído e se identifica o vizinho mais próximo $Y(j)$ de cada ponto $Y(i)$.

Calcula-se então a média logarítmica das distâncias entre os pares de pontos vizinhos, ao longo do tempo. Para cada i , calcule a média do logaritmo das distâncias $d(i, k)$ entre o ponto $Y(i)$ e seu vizinho mais próximo ao longo do tempo k , ou seja:

$$D(i) = \frac{1}{K} \sum_{k=0}^{K-1} \log(d(i, i+k)) \quad (5)$$

onde K é o número de passos no tempo ao longo do qual as trajetórias são seguidas. O expoente de Lyapunov é a inclinação da curva ajustada aos pontos $D(i)$ em função do tempo. Ou seja, é dado pela inclinação da linha de melhor ajuste (geralmente obtida por regressão linear) dos pontos $(k, D(i))$.

A Figura 16 compara os valores esperados e calculados pelo algoritmo de Rosenstein, para o expoente de Lyapunov com diferentes valores de r na equação logística.

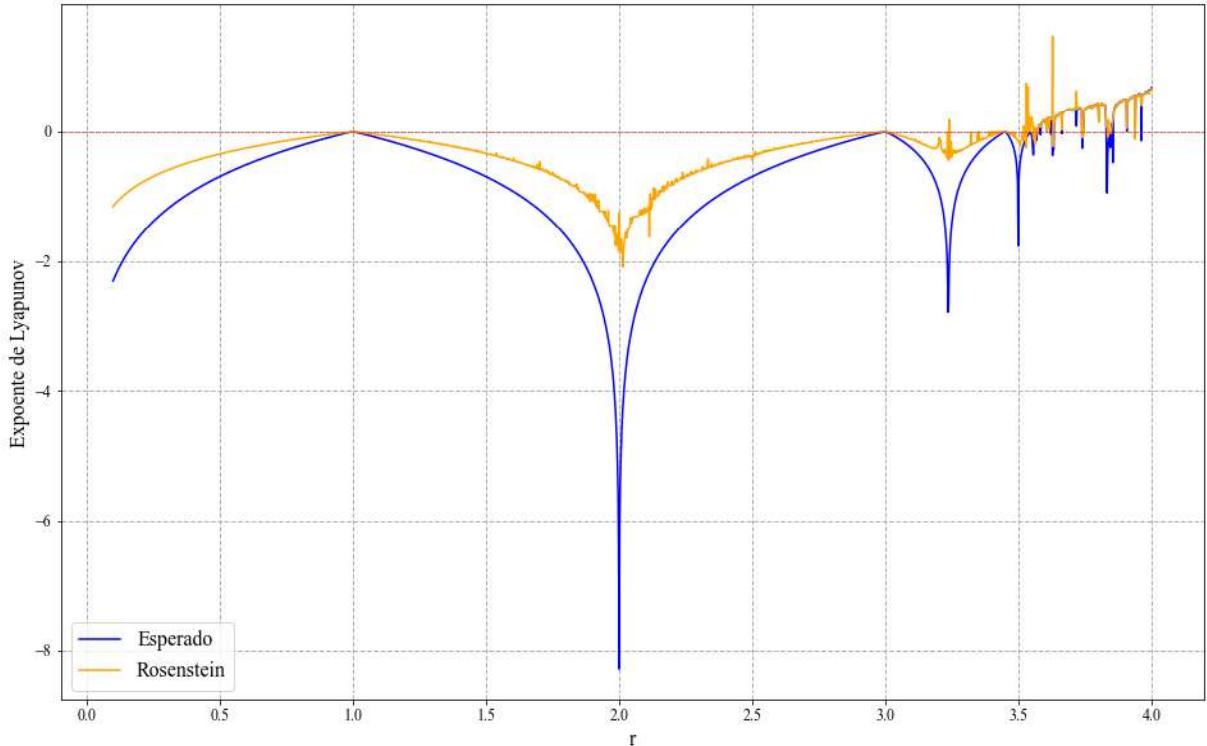


Figura 15 Expoente de Lyapunov Esperado e calculado pelo método de Rosenstein

Comparando-se ainda o algoritmo de Rosenstein com o valor esperado do expoente de Lyapunov para diferentes valores de iterações n , vemos que o algoritmo se aproxima do valor calculado, conforme Figura 17.

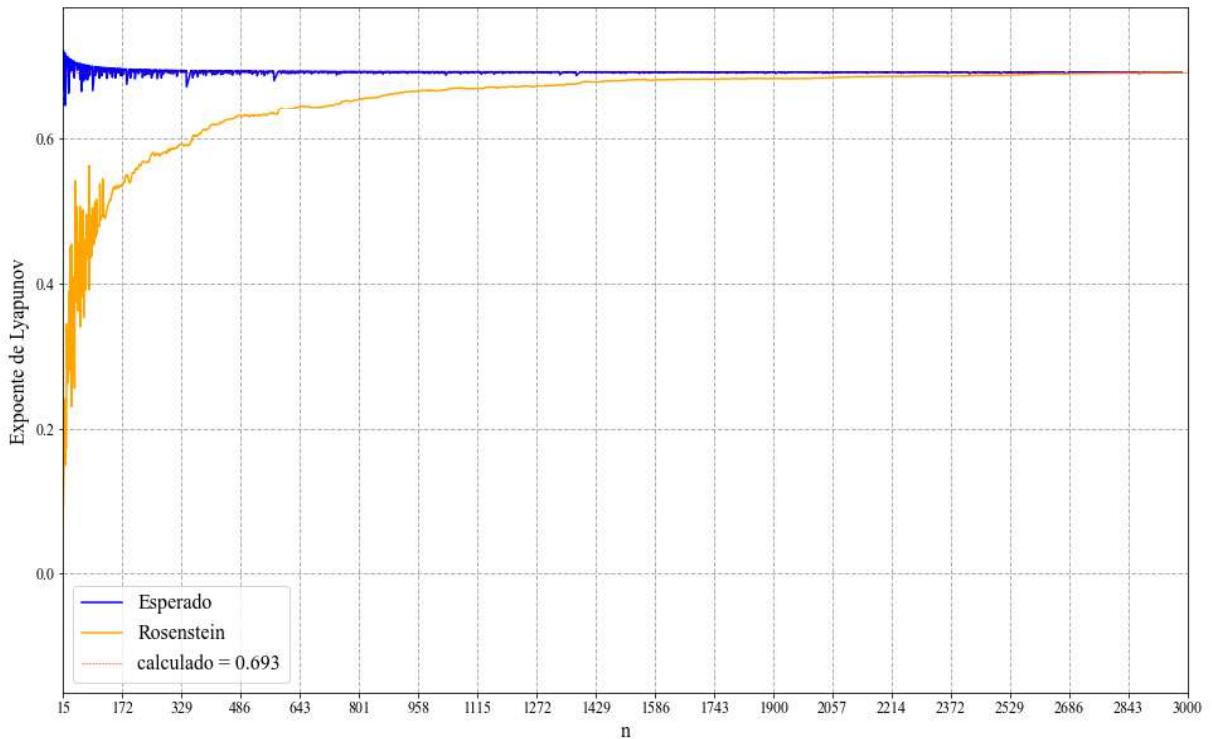


Figura 16 Exponente de Lyapunov Esperado e pelo método de Rosenstein, para diferentes valores de n

O algoritmo de Rosenstein é vantajoso porque é relativamente simples de implementar e não requer conhecimento prévio das equações do sistema dinâmico. No entanto, é sensível à escolha dos parâmetros, como a dimensão de incorporação, o atraso na incorporação e o número de pontos usados para o ajuste exponencial.

Outros indicadores são auxiliares para análise do comportamento de uma série de dados, o expoente de Hurst e a entropia amostral são dois desses exemplos que veremos adiante.

3.5 Expoente de Hurst

O Expoente de Hurst, denotado por H, é uma medida estatística utilizada para analisar séries temporais, especialmente para identificar a presença de autocorrelação de longo prazo ou memória nos dados. Este expoente foi originalmente desenvolvido por Harold Edwin Hurst (Hurst, 1951) ao estudar as variações anuais do nível do rio Nilo e o controle de vazões dos reservatórios.

Hurst utilizou em seus estudos uma razão adimensional R/S, onde R é a medida entre níveis máximo e mínimo do reservatório e S é o desvio padrão dos níveis medidos. O valor R/S permite comparar o reescalonamento de séries temporais das mais diferentes. Hurst concluiu que o reescalonamento, para muitas séries temporais, pode ser descrito por uma regra, conforme Hurst (1965).

$$\frac{R}{S} = kN^H \quad (6)$$

onde R/S = reescalonamento da série de dados registrados

k = uma constante

N = número de observações

H = expoente de Hurst

Para este trabalho utilizaremos o método descrito por Weron (2002), em que muito se assemelha a Mandelbrot e Wallis (1969), para estimar o expoente de Hurst, H. Esta metodologia é uma abordagem estatística que se baseia na análise de séries temporais. Esse método é particularmente útil para identificar a presença de dependência de longo alcance ou "memória" em dados financeiros, geofísicos, hidrológicos, entre outros.

Seja $X(t)$ uma série temporal com N observações. São então calculados os desvios em relação à média, isto é, $Y(t) = X(t) - \bar{X}$, onde \bar{X} é a média da série temporal. Em seguida é calculada a soma cumulativa $Z(t) = \sum_{i=1}^t Y(i)$ para $t = 1, 2, \dots, N$. Então, para cada intervalo selecionado n calcula-se o intervalo reescalonado R/S: sendo $R(n)$ a amplitude da série cumulativa $Z(t)$ em cada segmento de tamanho n , ou seja, $R(n) = \max(Z_{1,n}, Z_{2,n}, \dots, Z_{n,n}) - \min(Z_{1,n}, Z_{2,n}, \dots, Z_{n,n})$ e seja ainda $S(n)$ o desvio padrão das observações originais $X(t)$, em cada segmento de tamanho n .

Posteriormente é calculada a razão $R(n)/S(n)$ para diferentes tamanhos de segmento n . O expoente de Hurst é então calculado como a inclinação da reta de melhor ajuste em um gráfico $\log(R(n)/S(n))$ versus $\log(n)$.

Um valor de $H > 0,5$ indica uma tendência à persistência, onde um aumento é provável de ser seguido por outro aumento e uma queda por outra queda. Um valor de $H < 0,5$ indica antipersistência, onde um aumento é provável de ser seguido por uma queda e vice-versa. Um valor de $H = 0,5$ sugere um comportamento semelhante a um passeio aleatório, sem memória.

Na Figura 18 é apresentada uma comparação da evolução do expoente de Hurst para o mapa logístico, variando os valores de r e comparando com o diagrama de bifurcação. Observa-se um expoente próximo de 1 enquanto $1 < r < 3$, ou seja, o mapa logístico tem comportamento periódico ou convergentes. Para $3 > r > 3,57$, o expoente de Hurst fica abaixo de 0,5 motivado pelo comportamento oscilatório do mapa logístico. E para $r > 3,57$ o expoente de Hurst tende a 0,5 indicando o comportamento aleatório da série.

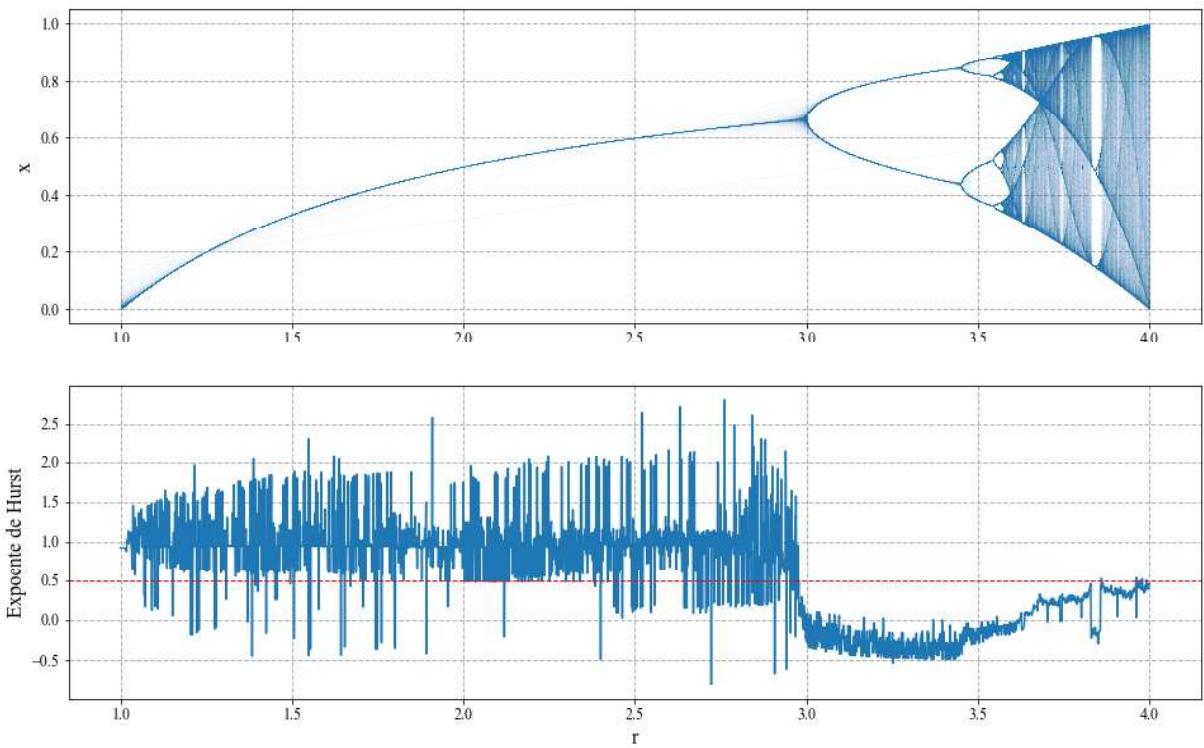


Figura 17 Expoente de Hurst Esperado e Diagrama de Bifurcação, para o mapa logístico

Mandelbrot e Wallis (1969) exploram a robustez do método de intervalo reescalonado (R/S) como uma ferramenta para detectar dependência de longo prazo em séries temporais e fornecer uma medida quantitativa da autocorrelação de longo prazo. O método R/S é então eficaz em identificar a presença de persistência ou memória em séries temporais.

3.6 Entropia Amostral

Richman e Moorman (2000) introduzem a *Sample Entropy (SampEn)* como uma medida menos dependente do comprimento da série em comparação com a *Approximate Entropy (ApEn)*. A *SampEn* é projetada para avaliar a complexidade de séries temporais, fornecendo uma medida da quantidade de previsibilidade ou regularidade em dados fisiológicos e outros tipos de séries temporais.

Dada uma série temporal $\{x_1, x_2, \dots, x_N\}$ com N pontos, é preciso definir dois parâmetros: o comprimento da sequência m e o limiar de similaridade r . Então é preciso formar sequências de comprimento m para cada i de 1 a $N - m + 1$, isto é, $X_m(i) = \{x_i, x_{i+1}, \dots, x_{i+m-1}\}$.

Para cada par de vetores $X(i)$ e $X(j)$ com $j \neq i$, calcule a distância d entre eles, distância máxima $d_{\max}(X(i), X(j)) = \max(|x_{i+k} - x_{j+k}|)$ para $k = 0, \dots, m - 1$, é frequentemente usada. É feita então a contagem de sequências similares para cada i , conta-se o número de vetores $X(j)$ ($j \neq i, j = 1, \dots, N - m$) tal que $d_{\max}(X(i), X(j)) \leq r$. Denote esta quantidade como $B(i)$ para vetores de comprimento m e $A(i)$ para vetores de comprimento $m+1$.

Calcula-se a probabilidade $B = \frac{1}{(N-m)(N-m-1)} \sum_{i=1}^{N-m} B(i)$ e $A = \frac{1}{(N-m-1)(N-m-2)} \sum_{i=1}^{N-m-1} A(i)$. A SampEn é definida como

$$\text{SampEn}(m, r, N) = -\ln \left(\frac{A}{B} \right). \quad (7)$$

Essencialmente, segundo os autores, a entropia amostral (SampEn) compara a probabilidade de que sequências de dados que são próximas (similares) para um dado comprimento m permaneçam próximas para o próximo comprimento de sequência $m+1$. Uma SE menor indica uma maior regularidade ou previsibilidade na série temporal, enquanto uma SE maior indica maior complexidade ou irregularidade.

Na Figura 19 vemos o comportamento da entropia para o mapa logístico, variando o valor de r . A entropia segue zerada ou muito pequena até $r \approx 3,57$, depois disso ela apresenta valores crescentes até $r = 4$, tendo momentos de queda nos 2^n -ciclos.

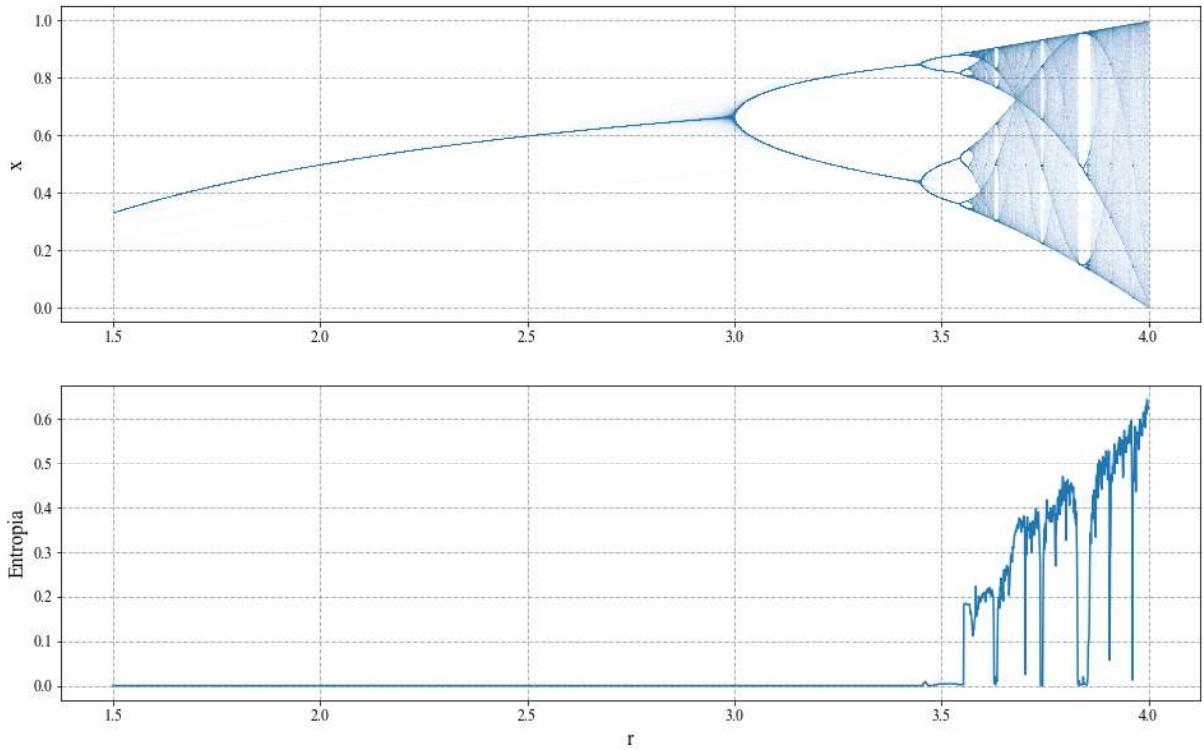


Figura 18 Entropia Amostral (SampEn) e Diagrama de Bifurcação, para o mapa logístico

Independentemente do sistema analisado, prever seu comportamento ao longo do tempo é algo de interesse. Mas essa tarefa não é simples, principalmente para sistemas complexos que possuem não-linearidade, ou seja, sistemas altamente sensíveis às condições iniciais.

Isso sugere que, embora um sistema não linear evolua no tempo com um comportamento instável e aperiódico, tal comportamento pode ser determinístico, pois seu estado futuro depende de seu estado atual.

Neste capítulo, foram explorados os conceitos e metodologias associadas ao expoente de Lyapunov, expoente de Hurst e entropia amostral, ferramentas auxiliares para a análise de séries temporais. O expoente de Lyapunov permite identificar a presença de comportamentos caóticos em séries temporais. O expoente de Hurst permite identificar os comportamentos persistentes, antipersistentes e aleatórios. Já a entropia amostral avalia a regularidade e a previsibilidade dos dados.

Juntos, esses indicadores abrem caminho para uma compreensão da dinâmica de séries temporais, permitindo uma análise não-linear do comportamento dos dados. No capítulo seguinte será analisada a aplicação desses métodos nas séries de preços de energia elétrica e quais *insights* podem ser feitos a partir dos resultados.

4 RESULTADOS E DISCUSSÃO

Os conceitos apresentados no Capítulo 3 tem papel importante na compreensão da complexidade, persistência e dinâmica temporal de séries temporais. A particular aplicação desses conceitos na série de preços de energia elétrica apresentada no Capítulo 2 tem seus resultados discutidos nesta seção do trabalho.

Conforme consta na introdução, a motivação deste estudo foi trazer um primeiro entendimento do comportamento não-linear da série histórica do preço de energia elétrica e como isso poder ter uma correlação com a dinâmica de operação, planejamento e formação de preço.

Inicialmente, o expoente de Lyapunov avalia a estabilidade e sensibilidade do sistema a pequenas variações nas condições iniciais. Isso auxilia na compreensão da resiliência do mercado de energia elétrica diante de perturbações e eventos extremos. Aplicando então o algoritmo de Rosenstein as séries semanal e horária de preço, foram obtidos os seguintes valores

- Semanal: 0,16
- Horário: 0,32

Uma segunda análise é avaliar as subséries. A Figura 20 apresenta o comportamento para 5000 subséries de tamanho 2190 para os preços horários (2190 horas equivale a três meses) e de tamanho 52 para os preços semanais (52 semanas equivale a um ano).

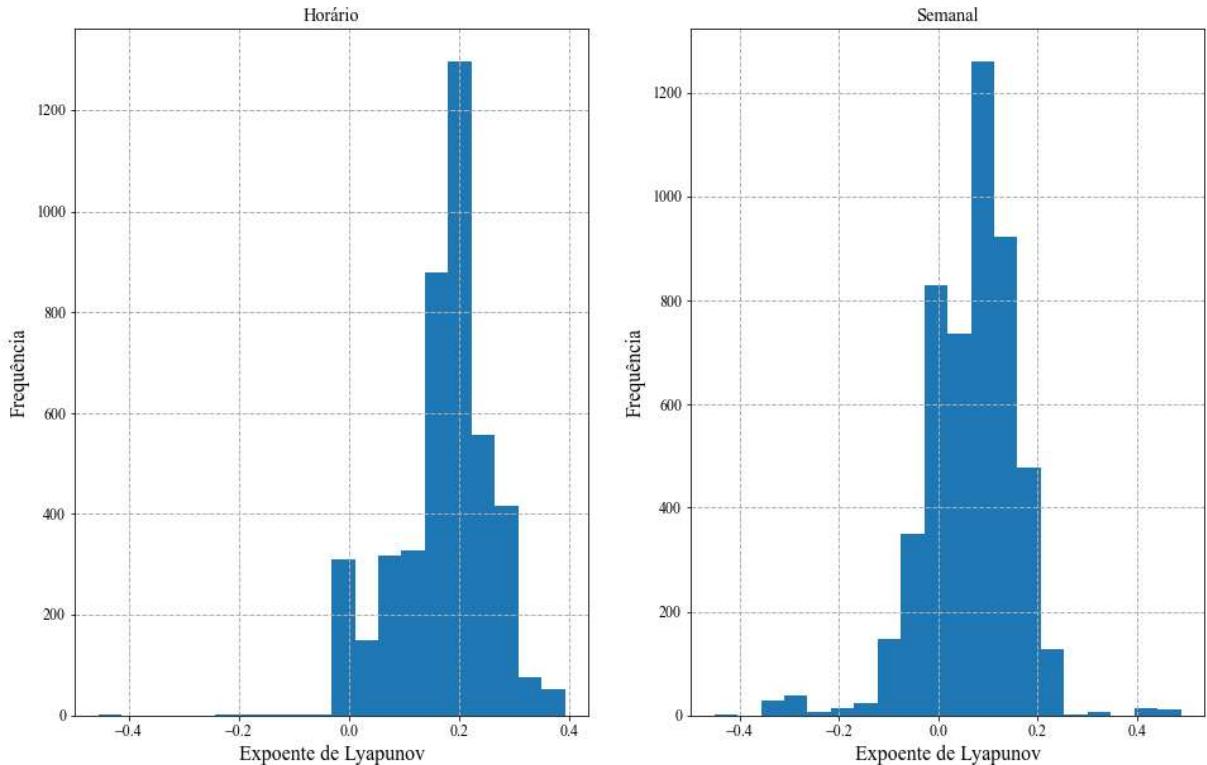


Figura 19 Histogramas do expoente de Lyapunov para 5000 subséries, de tamanho 2190 para os preços horários e tamanho 52 para os preços semanais

Os resultados obtidos das 5000 subséries são os seguintes:

- Semanal: $0,09 \pm 0,06$
- Horário: $0,17 \pm 0,08$

Os expoentes positivos obtidos, tanto para as séries inteiras quanto para as subséries, indicam que o preço de energia é sensível as condições iniciais e isso sugere maior dificuldade na previsão dos preços. Esse comportamento pode ser associado as diferentes fontes energéticas que compõe a matriz elétrica e trazem incertezas na formação do preço.

A segunda parte envolveu o cálculo do expoente de Hurst, para indicar a persistência ou reversão de tendências nos preços. Como visto no Capítulo 3, essa métrica oferece uma perspectiva para modelagem de longo prazo, fornecendo informações sobre a natureza das tendências. Foram calculados os expoentes de Hurst para as séries semanal e horária, com os resultados

- Semanal: 0,85
- Horário: 0,95

Uma segunda análise, conforme a Figura 21, foi o comportamento do expoente à medida que a série aumenta, ou seja, para visualizar o comportamento do expoente conforme aumentam as quantidades de registros das séries.

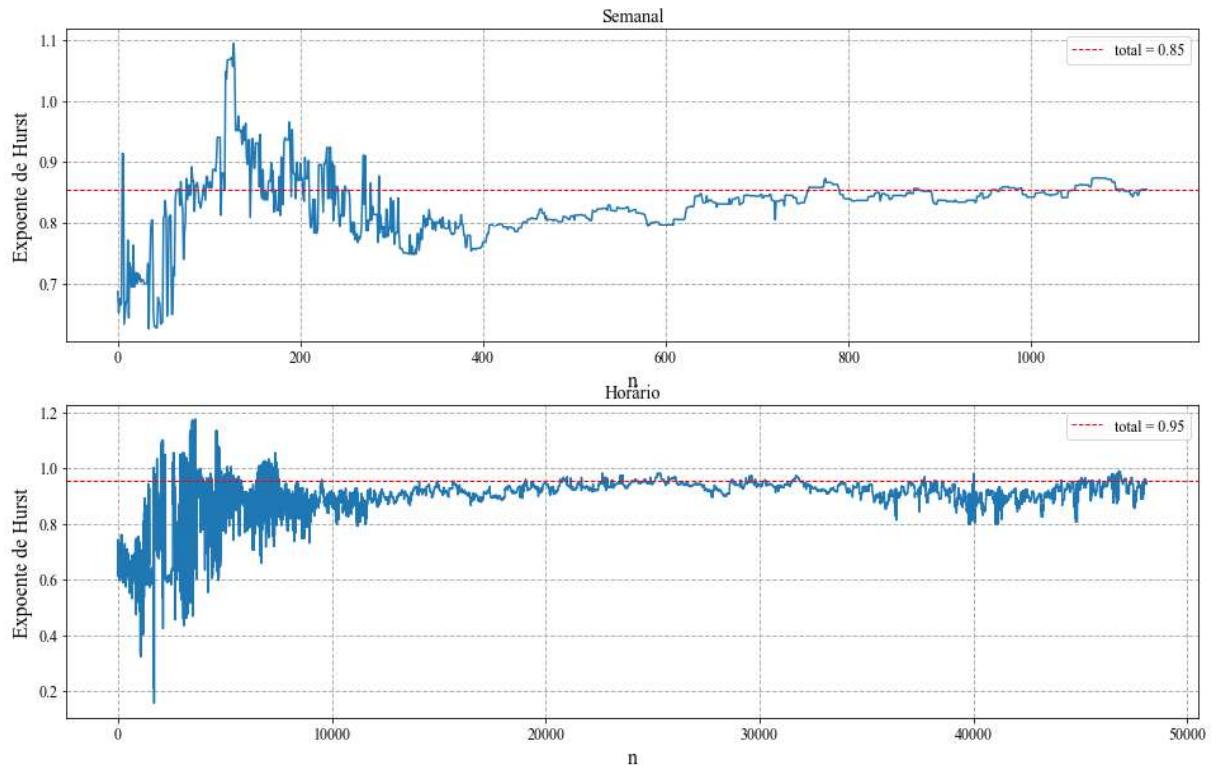


Figura 20 Expoente de Hurst vs número de registros n

Por fim, a Figura 22 apresente um estudo do comportamento do expoente de Hurst para 5000 subséries de tamanho 2190 para os preços horários (2190 horas equivale a três meses) e de tamanho 52 para os preços semanais (25 semanas equivale a um ano). Onde foi possível notar, para os preços horários, uma aproximação dos valores de expoente de Hurst das subséries com a série principal.

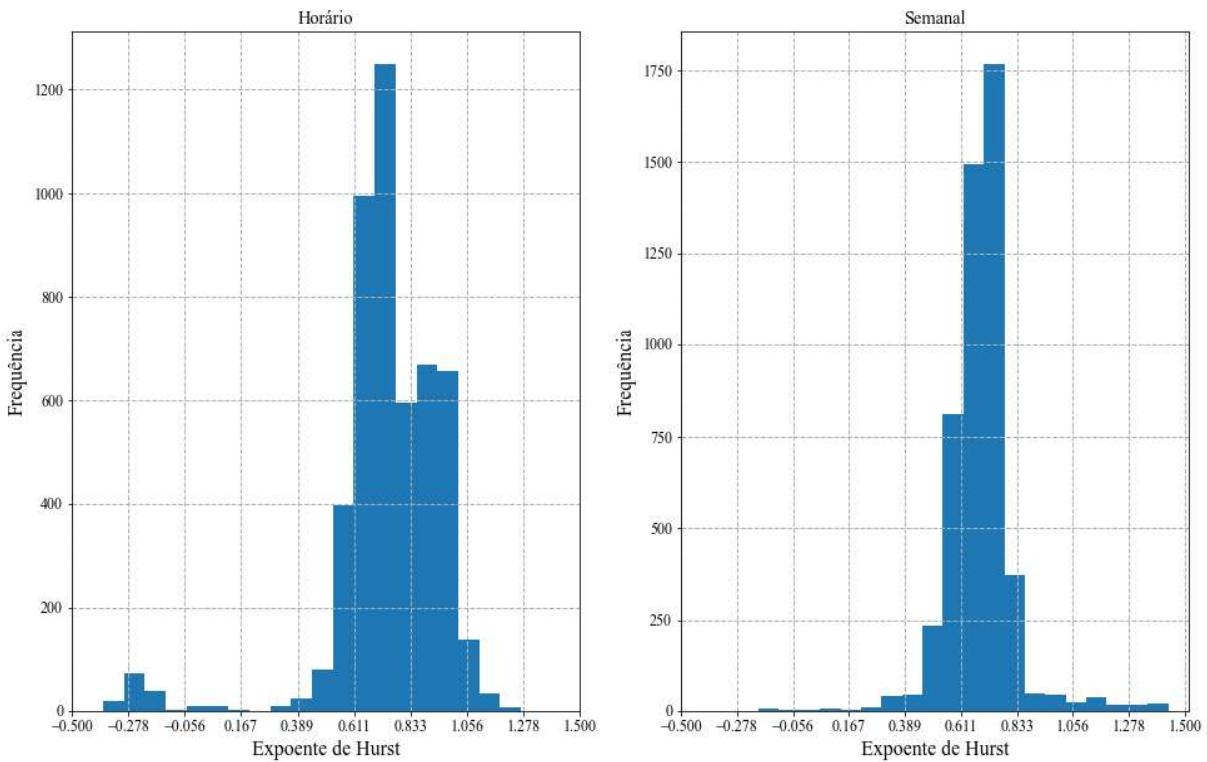


Figura 21 Histogramas do expoente de Hurst para 5000 subséries, de tamanho 2190 para os preços horários e tamanho 52 para os preços semanais

Os resultados obtidos das 5000 subséries são os seguintes:

- Semanal: $0,69 \pm 0,13$
- Horário: $0,74 \pm 0,22$

Os valores obtidos maiores que 0,5 para o expoente de Hurst, para a série total e subséries, sugerem uma tendência de persistência, indicando que aumentos ou diminuições nos preços têm maior probabilidade de serem seguidos por movimentos similares. Isso implica uma memória longa na série temporal, onde os preços futuros são influenciados por comportamentos passados de maneira significativa, podendo indicar menos volatilidade e mais previsibilidade nos movimentos de preços a longo prazo. Isso pode estar associado ao fato de cenários meteorológicos não se alterarem rapidamente e influenciarem a formação do preço durante longos períodos.

O terceiro indicador avaliado foi a entropia amostral, que é uma medida de desordem e da imprevisibilidade nos padrões temporais dos preços. Essa ferramenta é auxiliar para entendimento da volatilidade, identificar mudanças de regime e avaliar a presença de padrões sazonais, oriundos das fontes energéticas que compõe a matriz.

Para todo o histórico analisado, as entropias obtidas foram:

- Semanal: 0,22
- Horário: 0,03

Uma outra análise feita, conforme a Figura 23, foi o estudo das 5000 subséries de tamanho 2190 para os preços horários (2190 horas equivale a três meses) e de tamanho 52 para os preços semanais (52 semanas equivale a um ano).

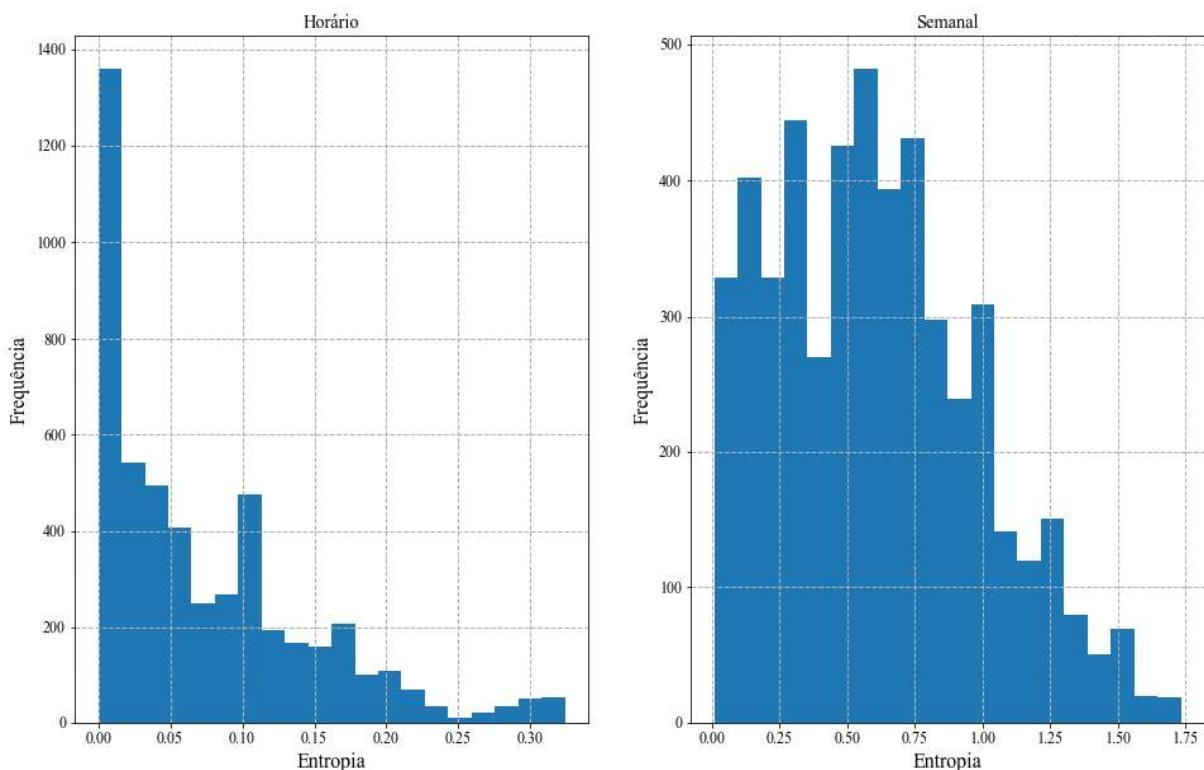


Figura 22 Histogramas da Entropia para 5000 subséries, de tamanho 2190 para os preços horários e tamanho 52 para os preços semanais

Os resultados obtidos das 5000 subséries são os seguintes:

- Semanal: $0,61 \pm 0,37$
- Horário: $0,07 \pm 0,05$

Para uma série de preços de energia elétrica, uma SampEn acima de 0 significa que existem padrões de comportamento que não são completamente previsíveis ou regulares. Entretanto, existe uma diferença do comportamento da entropia para as séries horária e semanal. A série semanal apresenta maior entropia, este valor relativamente mais alto que o horário sugere que a série de preços semanais de energia elétrica tem uma maior complexidade e menor previsibilidade em comparação com a série horária. Isso pode ser devido a variações mais

significativas nos preços ao longo das semanas. Para a série horária este valor mais baixo indica que a série de preços horários de energia elétrica é mais regular e previsível. Variações nos preços da energia elétrica dentro de um único dia tendem a seguir um padrão mais consistente.

A presença de complexidade nos preços pode ser reflexo de volatilidade no mercado, onde fatores como variações na oferta e demanda, mudanças regulatórias, flutuações nos custos de produção de energia e eventos externos, influenciam os preços de maneira não linear e imprevisível.

A SampEn acima de 0, mesmo que levemente acima, indica que os preços possuem um certo grau de complexidade e imprevisibilidade, refletindo a influência de múltiplos fatores internos e externos que afetam o mercado. O expoente de Lyapunov positivo reforça essa interpretação, apontando para uma sensibilidade às condições iniciais e uma tendência ao comportamento caótico, onde pequenas mudanças podem levar a variações significativas nos preços. Isso sugere que o mercado de energia é dinâmico, com a potencialidade de mudanças abruptas e imprevisíveis nos preços decorrentes de eventos pequenos.

Por outro lado, um expoente de Hurst próximo de 1 sugere uma forte tendência de persistência nos preços da energia elétrica, indicando que as tendências atuais têm uma alta probabilidade de continuar no futuro próximo. Essa persistência, quando vista em conjunto com a complexidade e a sensibilidade às condições iniciais, sugere que, embora os preços possam seguir uma tendência estável a longo prazo, eles estão sujeitos a volatilidades de curto prazo e a desvios imprevisíveis. Para os participantes do mercado, isso implica na necessidade de estratégias de gestão de risco planejadas e a utilização de ferramentas analíticas para entender a dinâmica do mercado de energia. Os operadores e investidores devem estar preparados para responder as mudanças no mercado, aproveitando as tendências persistentes, mas também protegendo-se contra a volatilidade e os riscos associados à natureza imprevisível e potencialmente caótica dos preços da energia elétrica.

A identificação do comportamento caótico nos preços de energia impulsiona a inovação em modelagem e análise. Métodos mais avançados podem ser desenvolvidos para capturar melhor a complexidade subjacente e melhorar as estratégias de previsão.

5 CONCLUSÕES

A utilização das métricas descritas no Capítulo 3 deste trabalho, na análise da série de preços de energia elétrica do Brasil, serve de ponto de partida para um aprofundamento do estudo do setor elétrico e como a operação e planejamento do sistema podem influenciar na formação do preço, trazendo reflexos para a economia do país.

Os resultados compilados, conforme figura 23, indicam que pequenas variações nos preços podem levar a divergências significativas no longo prazo, dificultando previsões. Inéncias no mercado se refletem na série, mantendo os preços em uma trajetória consistente ao longo do tempo. E existem elementos de aleatoriedade e complexidade que podem introduzir incertezas na previsão de preços.

		Expoente de Lyapunov	Expoente de Hurst	Entropia Amostral
Semanal	Série total	0,16	0,85	0,22
	Subséries	0,09	0,69	0,61
Horário	Série total	0,32	0,95	0,03
	Subséries	0,17	0,74	0,07

Figura 23 Tabela resumo dos valores médios dos indicadores analisados, para a série total e 5000 subséries

A partir deste estudo inicial, pesquisas futuras podem expandir a aplicação desses indicadores para outras séries temporais econômicas e financeiras, explorando sua utilidade em diferentes contextos e mercados. A continuação deste trabalho poderá abordar limitações identificadas, testar a aplicabilidade dos métodos em previsões de longo prazo e desenvolver modelos preditivos que incorporem essas métricas, abrindo espaço para a análise econômica e a gestão de riscos no mercado de energia.

REFERÊNCIAS

- [1] Operador Nacional do Sistema; Câmara de Comercialização de Energia Elétrica. **Treinamento Newave e Decomp – O SIN e os modelos para o Planejamento da Operação Energética.** Disponível em: sintegre.ons.org.br
- [2] Operador Nacional do Sistema. **O PLANEJAMENTO DA OPERAÇÃO ENERGÉTICA NO SISTEMA INTERLIGADO NACIONAL - CONCEITOS, MODELAGEM MATEMÁTICA, PREVISÃO DE GERAÇÃO E CARGA.** 2020, Artliber Editora Ltda.
Disponível em: <https://www.ons.org.br/AcervoDigitalDocumentosEPublicacoes/Livros-da-Diretoria-de-Planejamento-na-biblioteca-digital-ONS/LIVRO-O-PLANEJAMENTO-DA-OPERA%C3%87%C3%83O-ENERG%C3%89TICA-NO-SISTEMA%20INTERLIGADO-NACIONAL.pdf>
- [3] ALLIGOOD, K. T.; SAUER, T. D.; YORKE, J. A. **CHAOS: An introduction to Dynamical Systems.** Nova York: Springer-Verlag New York, 1996.
- [4] MANDELBROT, B. B.; WALLIS, J. R. **Robustness of the rescaled range R/S in the measurement of noncycle log run statistical dependence.** Water Resources Research, 1969.
- [5] PINCUS, S. M. **Approximate entropy as a measure of system complexity.** Proceedings of the National Academy of Sciences, v. 88, n. 6, p. 2297-2301, 1990.
- [6] CASTRO, L. I; BORGES, G. G.; SIMONE, L. F. C. **O FUTURO DA FORMAÇÃO DE PREÇOS DE ENERGIA NO BRASIL.** Revista de Economia Contemporânea, 2023.
- [7] STROGATZ, S. H. **Nonlinear Dynamics and Chaos: With ap applications to Physics, Biology, Chemstry and Engineering.** 2015, 2 ed., CRC Press.
- [8] BOYCE, W. E.; DiPrima, R. C. **Equações Diferenciais Elementares e Problemas de Valores de Contorno.** 2009, 9^a Edição, Editora LTC.
- [9] MAY, R. M, **Simple mathematical models with very complicated dynamics.** Nature Journal, 1976.
- [10] RICHARDSON, G.; MITCHELL, M. **complexityexplorer.org.** Santa Fe Institute.
Disponível em: <https://www.complexityexplorer.org/about>
- [11] GUEGAN, D. **Chaos in Economics and Finance.** HAL Open Science, 2009. Annual Reviews in Control, pp.89-93.
Disponível em: <https://shs.hal.science/halshs-00375713v2>
- [12] ROSENSTEIN, M. T.; COLLINS, J. J.; DE LUCA, C. J. **A practical method for calculating largest Lyapunov exponents from small data sets.** NeuroMuscular Research Center and Department of Biomedical Engineering, 1992. Boston University, Boston, MA, USA.
- [19] HURST, H. E. **LONG-TERM STORAGE CAPACITY OF RESERVOIRS.** Transactions of the American Society of Civil Engineers, 1951.

- [20] HURST, H. E.; BLACK, R. P.; SIMAIKA, Y. M. **Long-term storage: an experimental study**. London: Constable London, 1965.
- [21] WERON, R. **Estimating long-range dependence: finite sample properties and confidence intervals**. *Physica A: Statistical Mechanics and its Applications*, 2002.
- [22] RICHMAN, J. S.; MOORMAN, J. R. **Physiological time-series analysis using approximate entropy and sample entropy**. *American Journal of Physiology-Heart and Circulatory Physiology*, 2000

ANEXO A – CÓDIGOS COMPUTACIONAIS

Código utilizado para cálculo dos resultados no Capítulo 4 – Resultados e Discussão.

```

###BIBLIOTECAS UTILIZADAS
import pandas as pd
import nolds
import numpy as np

### Carrega dados do Excel em um DataFrame do pandas

#Preço Semanal
path_semanal = 'C:/caminho_do_arquivo/nome_do_arquivo_semanal.xlsx'
sheet_name_semanal = 'Planilha1' # Substitua pelo nome da sua planilha
dados_preco_semanal = pd.read_excel(path_semanal, sheet_name=sheet_name_semanal)

#Preço Horário
path_hora = 'C:/caminho_do_arquivo/nome_do_arquivo_hora.xlsx'
sheet_name_hora = 'Planilha1' # Substitua pelo nome da sua planilha
dados_preco_hora = pd.read_excel(path_hora, sheet_name=sheet_name_hora)

### Escolhe a coluna desejada e a converte para um array numpy, depois limpas os dados

col_preco_semanal = 'PLD' # Substitua pelo nome da sua coluna
preco_semanal = dados_preco_semanal[col_preco_semanal].values
# Remove NaNs e Infs da lista
dados_preco_semanal_limpo = [x for x in preco_semanal if not (np.isnan(x) or np.isinf(x))]

col_preco_hora = 'Valor' # Substitua pelo nome da sua coluna
preco_hora = dados_preco_hora[col_preco_hora].values
# Remove NaNs e Infs da lista
dados_preco_hora_limpo = [x for x in preco_hora if not (np.isnan(x) or np.isinf(x))]

###Calcula o expoente de Lyapunov total da série
lyapunov_total_semanal = nolds.lyap_r(dados_preco_semanal_limpo, emb_dim = 4, lag = 1
                                         , min_tsep = None, tau = 1, min_neighbors = 20, trajectory_len = 10
                                         , fit=u'RANSAC', debug_plot=False, debug_data=False, plot_file=None, fit_offset=0)
print(round(lyapunov_total_semanal,5))

lyapunov_total_hora = nolds.lyap_r(dados_preco_hora_limpo, emb_dim = 4, lag = 1
                                     , min_tsep = 250, tau = 1, min_neighbors = 20, trajectory_len = 10
                                     , fit=u'poly', debug_plot=False, debug_data=False, plot_file=None, fit_offset=0)
print(round(lyapunov_total_hora,5))

###Calcula o expoente de Hurst total da série
hurst_total_semanal = nolds.hurst_rs(dados_preco_semanal_limpo, nvals=None, fit=u'RANSAC',
                                       debug_plot=True, debug_data=False, plot_file=None, corrected=True, unbiased=False)
print(round(hurst_total_semanal,5))

hurst_total_hora = nolds.hurst_rs(dados_preco_hora_limpo, nvals=None, fit=u'RANSAC',
                                   debug_plot=True, debug_data=False, plot_file=None, corrected=True, unbiased=False)
print(round(hurst_total_hora,5))

###Calcula a Entropia total da séric
samplen_total_semanal = nolds.samplen(dados_preco_semanal_limpo, emb_dim=4, tolerance=None,
                                         debug_plot=False, debug_data=False, plot_file=None)
print(round(samplen_total_semanal,5))

samplen_total_hora = nolds.samplen(dados_preco_hora_limpo, emb_dim=4, tolerance=None,
                                    debug_plot=False, debug_data=False, plot_file=None)
print(round(samplen_total_hora,5))

```

```
###Calculo das subséries
cont = 0
cont_histograma_lyap = []
cont_histograma_hurst = []
cont_histograma_sampen = []
while cont<5000:

    # Escolher aleatoriamente o ponto de início para obter uma amostra contínua de 5000 dados
    ponto_inicio = np.random.randint(0, len(dados_preco_hora_limpo) - 730)

    # Obter uma amostra aleatória de 1000 dados consecutivos
    amostra_aleatoria = dados_preco_hora_limpo[ponto_inicio:ponto_inicio + 730]

    # Calculo dos indicadores
    amostra_aleat_lyap = nolds.lyap_r(amostra_aleatoria, emb_dim = 4, lag = 1
        , min_tsep = 250, tau = 1, min_neighbours = 20, trajectory_len = 10
        , fit=u'poly', debug_plot=False, debug_data=False, plot_file=None, fit_offset=0)
    amostra_aleat_hurst = nolds.hurst_rs(amostra_aleatoria, nvals=None, fit=u'RANSAC'
        , debug_plot=False, debug_data=False, plot_file=None, corrected=True, unbiased=False)
    amostra_aleat_sampen = nolds.sampen(amostra_aleatoria, emb_dim=1, tolerance=None,
        debug_plot=False, debug_data=False, plot_file=None)

    # Armazena os dados
    cont_histograma_lyap.append(amostra_aleat_lyap)
    cont_histograma_hurst.append(amostra_aleat_hurst)
    cont_histograma_sampen.append(amostra_aleat_sampen)
    cont +=1
```