

Faculdade de Engenharia da Universidade do Porto



**Análise de Séries Temporais Através de
Representações do Espaço de Fases**

Richard Moisés Alves Pinto

Dissertação no âmbito do
Mestrado Integrado em Engenharia Electrotécnica e de Computadores
Major em Telecomunicações

Orientador: Prof. Francisco José de Oliveira Restivo (DEI)
Co-orientador: Prof. Francisco José Lage Campelo Calheiros (DEC)

Junho 2009

© Richard Pinto, 2009

Dedicatória

Gostaria de dedicar este trabalho a memória do meu pai Fernando, e a minha mãe Alice que com todo o seu amor, abnegação e sacrifício foi e continua a ser o pilar central da nossa família...

Resumo

O estudo de comportamentos caóticos em diversos fenómenos naturais e sociais abriu uma nova porta para o entendimento de comportamentos que até pouco tempo atrás eram catalogados como aleatórios ou imprevisíveis. Em particular o estudo do chamado caos determinístico visa encontrar modelos matemáticos que consigam descrever fenómenos cujo comportamento é dominado pelo caos. Neste trabalho é feito um estudo sobre alguns dos principais métodos para a descrição e modelagem de séries temporais com especial destaque para a representação no espaço de fases. É desenvolvida uma ferramenta informática com o objectivo de fornecer um suporte para o estudo de séries temporais a partir de representações no espaço de fases, sendo também incluídas nesta ferramenta, algumas outras funcionalidades concebidas para dar um apoio na detecção de padrões de comportamento caótico nos dados analisados. Finalmente é feito um estudo baseado em observações reais de caudais de rios e de cotações de títulos na bolsa de valores, com o objectivo de tentar descobrir algum padrão de comportamento caótico nestes dados. Os resultados obtidos para os caudais de rios mostram ausência de tais padrões de comportamento caótico. Já os resultados para cotações de títulos mostram comportamentos mais próximos do caos, sendo contudo impossível por agora, dar um veredicto definitivo acerca destes comportamentos.

Abstract

The study of the chaotic behaviour of several natural and social phenomena has opened a new door to understanding behaviours that until recently were classified as random or unpredictable. In particular, the study of the so called deterministic chaos aims to the discovery of mathematical models capable of describing phenomena whose behaviour is dominated by chaos.

In this dissertation a study is made of the main methods for the description and modelling of time séries, with special emphasis for its representation in the phase space. A computer based tool is developed with the objective of supporting the study of time séries in the phase space. Other functions conceived exactly to support the detection of patterns of chaotic behaviour in the data under analysis, have also been included in this tool. Finally, it is presented a study based on real observations of river flow measurements and of the value of company stocks in the stock exchange, in an attempt to find some pattern of chaotic behaviour in the analysed time séries.

Results for the river flow measurements show the absence of patterns of chaotic behaviour, while the results for the stock exchange values of the studied company exhibit some behaviour possibly chaotic. It is however impossible to give a definitive verdict at this stage of the study.

Agradecimentos

Gostaria de agradecer primeiramente a Deus Jeová por ter-me dado tranquilidade nas horas mais difíceis, ao meu orientador, Prof. Francisco Restivo, pela motivação inculcada e pelo tempo e apoio disponibilizados ao longo deste trabalho e ao meu co-orientador, Prof. Francisco José Calheiros, pelos seus comentários e sugestões experientes. Finalmente uma última nota de agradecimento ao Prof. José Campos Neves pela sua colaboração e pelo acesso aos seus programas de análise de sinais no espaço de fases, e a todas aquelas pessoas que deram o seu contributo de maneira directa ou indirecta na realização deste trabalho.

Índice

1 - Introdução	1
1.1 – Objectivos da dissertação	1
1.2 – Estrutura da dissertação.....	2
2 - Séries temporais e métodos lineares.....	3
2.1 - Objectivos do estudo das séries temporais.....	4
2.2 - Processos estocásticos estacionários.....	5
2.3 - Processos Estacionários Lineares	6
2.3.1 - Modelos Autoregressivos [AR(p)]	6
2.3.2 - Modelos Médias Moveis [MA(q)].....	7
2.3.3 - Modelos Mistos Autoregressivos e Médias Moveis [ARMA(p,q)]	7
2.4 - Correlação em séries temporais	8
2.4.1 - O Correlograma.....	8
2.4.2 - Correlograma em séries aleatórias.....	8
2.4.3 - Correlograma de curto prazo	9
2.4.4 - Correlograma em séries alternadas.....	9
2.4.5 - Correlograma em séries não estacionarias	9
2.5 - Espectro em séries temporais.....	10
2.5.1 - Série e Transformada de Fourier Discreta (DFT).....	10
2.5.2 - Análise espectral em processos estocásticos estacionários	10
2.6 – Vantagens e limitações dos modelos lineares estacionários	11
3 - Caos em Sistemas dinâmicos.....	13
3.1 - Caos	14
3.2 - Representações no espaço de fases.....	15
3.2.1 - Pontos fixos	15
3.2.2 - Atractores	16

3.3 - Mapa logístico.....	16
3.4 - Reconstrução do espaço de fases	19
3.4.1 - Escolha do desfasamento temporal	20
3.4.2 - Escolha da dimensão de imersão do espaço de reconstrução	21
3.5 - Expoentes de Lyapunov	22
4 - Programa de análise de séries temporais.....	24
4.1 – Implementação	25
4.1.1 – Instalação e requisitos	26
4.2 – Descrição.....	26
4.2.1 – Elementos do interface gráfico	26
4.2.2 – Menus	27
4.2.2.1 – ImportData.....	27
4.2.2.2 – Generate.....	29
4.2.2.3 – Noise.....	30
4.2.2.4 – Calculate.....	30
4.3 – Testes	33
5 - Casos de estudo	37
5.1 - Análise de caudais de rios.....	37
5.2 - Análise de cotações de títulos na bolsa de valores nacional	44
6 - Conclusões	53
Anexo.....	55
Refêrencias.....	57

Lista de figuras

Figura 2.1 - Precipitação média global entre os anos 1991 – 2000 [3].	4
Figura 2.2 - Segmento de sinal sinusoidal e respectiva função autocorrelação.	9
Figura 3.1 - pêndulo forçado amortecido [13; p7].	14
Figura 3.2 - Mapa logístico para $A = 2.9$.	16
Figura 3.3 - Mapa logístico para $A = 3.2$.	17
Figura 3.4 - Mapa logístico para $A = 3.5$.	17
Figura 3.5 - Mapa logístico para $A = 3.57$.	18
Figura 3.6 - Mapa logístico para $A = 4$.	18
Figura 3.7 - Diagrama de Bifurcação para o mapa logístico [12; p28].	19
Figura 3.8 - Mapa de Hénon e Mapa de Hénon reconstruído.	20
Figura 3.9 - Cálculo do maior expoente de Lyapunov [11].	23
Figura 4.1 - Interface gráfico.	26
Figura 4.2 - Janela de selecção do menu <i>ImportData</i> -> <i>text_file</i> .	28
Figura 4.3 - Janela de diálogo do menu <i>ImportData</i> -> <i>excel_file</i> .	28
Figura 4.4 - Janela de entrada de dados do menu <i>Generate</i> -> <i>function_1_variable</i> .	29
Figura 4.5 - Janela de entrada de dados para o comando <i>Generate</i> -> <i>function_recursive</i> .	29
Figura 4.6 - Janela de entrada de dados para o comando <i>Noise</i> -> <i>additive/multiplicative</i> .	30
Figura 4.7 - Informação mútua média em função do desfasamento temporal L , para o mapa logístico com os parâmetros usuais para 200 amostras.	31
Figura 4.8 - Dimensão de reconstrução óptima para o mapa logístico com os parâmetros usuais para 200 amostras.	31
Figura 4.9 - Janela de entrada de dados para o comando <i>Calculate</i> -> <i>Maximum_Lyapunov_Exponent</i> .	32
Figura 4.10 - Mapa logístico, com os parâmetros usuais, para 1000 amostras.	33
Figura 4.11 - Modelo populacional de Ricker's, com os parâmetros usuais, 1000 amostras.	34
Figura 4.12 - Mapa de Hénon, com os parâmetros usuais, para 1000 amostras.	34
Figura 4.13 - Mapa de Lozi, com os parâmetros usuais, para 1000 amostras.	35
Figura 4.14 - Mapa logístico atrasado, com os parâmetros usuais, para 1000 amostras.	35
Figura 4.15 - Segmento de sinal sinusoidal com a sua correspondente autocorrelação e informação mútua média em função do desfasamento temporal L .	36

Figura 5.1 - Cronograma, autocorrelação e espectro do caudal médio diário do rio Guadiana no posto de medição de Albernoa, desde 31/10/1982 até 31/10/1997.....	37
Figura 5.2 - Espaço de fases para o desfasamento temporal: 1 dia.....	38
Figura 5.3 - Espaço de fases para o desfasamento temporal: 1º zero da FAC= 58 dias.....	39
Figura 5.4 - Espaço de fases para o desfasamento temporal: 1º mínimo da IMM= 204 dias. ...	39
Figura 5.5 - Cronograma, autocorrelação e espectro do caudal médio diário (m3/seg) do rio Lima no posto de medição de Forjães, desde 01/10/1982 até 30/09/1989.....	40
Figura 5.6 - Espaço de fases para o desfasamento temporal: 1 dia.....	40
Figura 5.7 - Espaço de fases para o desfasamento temporal: 1º zero da FAC= 96 dias.....	41
Figura 5.8 - Espaço de fases para o desfasamento temporal: 1º mínimo da IMM= 83 dias.	41
Figura 5.9 - Cronograma, autocorrelação e espectro do caudal médio diário (m3/seg) do rio Mondego no posto de medição de Ponte Casével, desde 01/10/1975 até 30/09/1990.....	42
Figura 5.10 - Espaço de fases para o desfasamento temporal: 1 dia.....	43
Figura 5.11 - Espaço de fases para o desfasamento temporal: 1º zero da FAC= 106 dias.....	43
Figura 5.12 - Espaço de fases para o desfasamento temporal: 1º mínimo da IMM= 130 dias. .	44
Figura 5.13 - Cronograma, autocorrelação e espectro para o valor final diário (Euros) das acções do BCP no PSI20, desde 01/01/2003 até 11/05/2009.....	45
Figura 5.14 - Espaço de fases para o desfasamento temporal: 1 dia.....	46
Figura 5.15 - Espaço de fases para o desfasamento temporal: 1º zero da FAC= 241 dias.....	46
Figura 5.16 - Espaço de fases para o desfasamento temporal: 1º mínimo da IMM= 214 dias...47	47
Figura 5.17 - Cronograma, autocorrelação e espectro do valor final diário (Euros) das acções da Brisa no PSI20, desde 01/01/2003 até 11/05/2009.	47
Figura 5.18 - Espaço de fases para o desfasamento temporal: 1 dia.....	48
Figura 5.19 - Espaço de fases para o desfasamento temporal: 1º zero da FAC= 396 dias.....	49
Figura 5.20 - Espaço de fases para o desfasamento temporal: 1º mínimo da IMM= 104 dias. .	49
Figura 5.21 - Cronograma, autocorrelação e espectro do valor final diário (Euros) das acções da Sonae no PSI20, desde 30/06/2000 até 11/05/2009.	50
Figura 5.22 - Espaço de fases para o desfasamento temporal: 1 dia.....	50
Figura 5.23 - Espaço de fases para o desfasamento temporal: 1º zero da FAC= 366 dias.....	51
Figura 5.24 - Espaço de fases para o desfasamento temporal: 1º mínimo da IMM= 211 dias...52	52

Abreviaturas

AR	Auto-Regressivo.
ARMA	Auto-Regressivo e Médias Móveis.
DFT	Transformada de Fourier Discreta.
FAC	Função Autocorrelação.
IMM	Informação Mútua Média.
MA	Médias Móveis.
PSI20	Portuguese Stock Index 20.

Capítulo 1

Introdução

Desde os primórdios da civilização o homem tem tentado entender os fenómenos que o rodeiam, por vezes com finalidades práticas, científicas ou meramente por curiosidade. E basta olhar para a nossa história para compreendermos a importância das descobertas científicas na nossa vida corrente. De Galileo a Leonardo da Vinci a Newton e Einstein, para só referirmos estes quatro nomes, toda a nossa história tem sido pautada por descobertas científicas.

Comportamentos complexos estão presentes em muitos dos fenómenos naturais e sociais. A teoria do caos [32] resulta da procura de novos modelos para tentar explicar estes comportamentos que até pouco tempo atrás eram considerados totalmente imprevisíveis, sendo uma forte sensibilidade as condições iniciais a principal característica dos sistemas caóticos. Em particular o chamado caos-determinístico [16] baseia-se na ideia de que modelos matemáticos simples podem gerar padrões complexos, do mesmo modo que fenómenos complicados podem ser modelados por regras simples.

Cabe agora a pergunta, Por que razões cientistas, engenheiros e matemáticos estão interessados pelo caos? A primeira razão [14-p3] é que o estudo do caos tem dado origem a novas ferramentas conceptuais e teóricas que nos permitem categorizar e entender comportamentos complexos, o que era impossível com as teorias anteriores. A segunda razão é que a natureza está cheia de comportamentos caóticos, sendo alguns exemplos, osciladores mecânicos, sistemas ópticos não lineares, reacções químicas, células nervosas, fluidos aquecidos entre muitos outros.

Alguns contributos no desenvolvimento deste campo foram o estudo da estabilidade do sistema solar feito por Henri Poincaré por volta do ano 1900 e o estudo da turbulência feito por Andrey Kolmogorov por volta do ano 1941.

1.1 – Objectivos da dissertação

O primeiro objectivo deste trabalho é o estudo de métodos para a descrição de séries temporais, com especial destaque para a representação no espaço de fases.

O segundo objectivo passa pela criação duma aplicação informática que sirva como ferramenta para a aplicação dos métodos estudados na descrição e modelagem de séries temporais. As principais características pretendidas nesta ferramenta são:

- Interface gráfico de fácil compreensão.
- Robustez, que permita a análise duma ampla gama de séries temporais originadas por diferentes tipos de sistemas.
- Eficiência, especialmente no aspecto do tempo de processamento, visto ser necessária a implementação de algoritmos iterativos.
- Possibilidade de importar dados a partir de ficheiros externos.
- Possibilidade de gerar uma série de dados a partir de equações recursivas e não recursivas.
- Boa capacidade para a análise de séries de dados a partir de representações no espaço de fases.

Por último pretende-se fazer um estudo envolvendo séries temporais que resultem de observações de fenómenos naturais e sociais, tais como caudais de rios e cotações de títulos na bolsa de valores, isto com a esperança de conseguir encontrar algum indício de comportamentos caótico-determinísticos nestas séries.

1.2 – Estrutura da dissertação

Esta dissertação divide-se em seis capítulos, sendo o primeiro esta introdução.

Os dois capítulos seguintes contêm uma revisão do estado da arte correspondente ao estudo de séries temporais a partir de métodos lineares e representações no espaço de fases.

No capítulo dois é feito um estudo sobre séries temporais e das características que distinguem a este tipo dados de outros dados estatísticos. De seguida são analisados alguns modelos lineares para processos estocásticos estacionários e, por último, é feita uma descrição da função autocorrelação e do espectro, como valiosas ferramentas no estudo dos sistemas dinâmicos lineares.

No capítulo três, ainda correspondente ao estado da arte, é feito um estudo sobre sistemas dinâmicos caóticos, sendo o principal elemento deste estudo, a reconstrução no espaço de fases a partir de uma única série temporal, com base no teorema de Takens. Ainda são referenciados parâmetros fundamentais nesta reconstrução tais como o desfasamento temporal e a dimensão óptima de reconstrução. Por último é feita uma descrição dos chamados expoentes de Lyapunov, como medida para o nível de caos evidenciado por um sistema na sua representação no espaço de fases, com especial ênfase no maior expoente de Lyapunov.

O capítulo quatro nasce da necessidade de criação de uma ferramenta informática que permita aplicar os conhecimentos adquiridos até agora, no estudo de séries temporais a partir de representações no espaço de fases, para um caso concreto. Neste capítulo é feita uma descrição dos principais componentes e funcionalidades da ferramenta informática desenvolvida. Finalmente são feitos testes de validação para tal ferramenta.

No capítulo cinco, é feita uma análise concreta para séries temporais observadas no domínio das ciências naturais (caudal média diário de rios) e no domínio das ciências sociais (cotações de títulos na bolsa de valores), sendo um dos objectivos desta análise, a procura de padrões indicadores da presença de caos determinístico dentro das séries temporais.

Finalmente, no capítulo seis apresentam-se as principais conclusões deste trabalho.

Capítulo 2

Séries temporais e métodos lineares.

Uma sucessão cronológica ou série temporal [1] define-se como um conjunto de observações feitas de maneira sequencial ao longo do tempo. O que distingue as séries temporais de outros tipos de dados estatísticos é que neste caso a ordem das observações é crucial, uma vez que neste tipo de dados as observações vizinhas serão dependentes e estamos interessados em analisar e modelar esta dependência.

Algumas características particulares deste tipo de dados são:

- A ordem temporal das observações deve ser considerada.
- Observações correlacionadas são mais difíceis de analisar do que observações não correlacionadas e dada a sua natureza requerem técnicas específicas para o seu estudo.
- O processo de selecção de modelos adequados pode ser bastante complicado.
- É mais difícil lidar com observações perdidas e dados discrepantes devido à natureza sequencial dos dados.

Outra característica importante das séries temporais é a presença de componentes [2; p6] tais como tendências e variações sazonais ou cíclicas. A tendência pode ser vista como a inércia da sucessão, a sua variação “em média” ao longo do tempo. A sazonalidade descreve variações em relação a tendência que ocorrem geralmente em períodos de um ano. Por último as variações cíclicas se referem a movimentos oscilatórios, que não apresentam qualquer periodicidade definida, sendo que os ciclos muito longos são dificilmente separáveis da tendência.

Podemos encontrar exemplos de séries temporais nos mais diversos domínios. Em economia podemos falar de taxas de juro, produto interno bruto, etc.; em meteorologia, registo de temperaturas, registo de precipitação, etc.; em medicina, leitura de electrocardiogramas, leitura de electroencefalogramas, etc.

A modo de ilustração, pode-se considerar o seguinte gráfico meteorológico que representa a média de precipitação mundial entre os anos de 1991 e 2000.

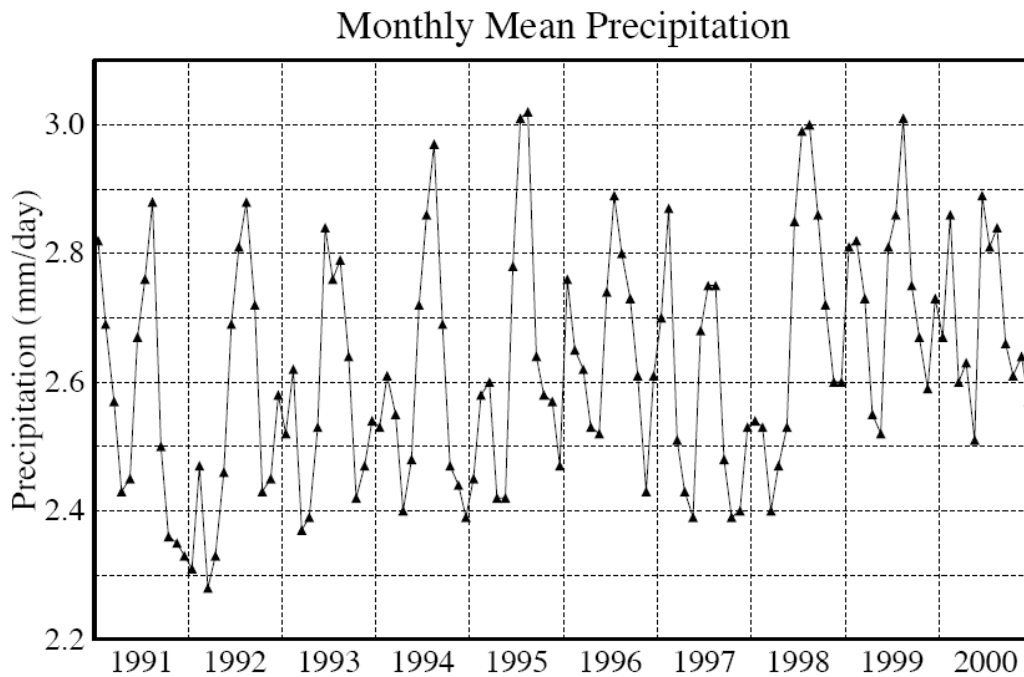


Figura 2.1 - Precipitação média global entre os anos 1991 – 2000 [3].

É fácil distinguir, no gráfico anterior, algumas das componentes antes mencionadas. Há uma ligeira componente de tendência global para o aumento da precipitação e há uma componente sazonal que dita valores de máximo para os meses correspondentes ao verão do hemisfério norte.

2.1 - Objectivos do estudo das séries temporais.

Em [2; p5] são mencionados os quatro objectivos gerais no estudo de séries temporais:

1. Descrição: É a tarefa primária a efectuar e consiste, numa primeira fase, da análise gráfica da série com respeito ao tempo e a conseqüente procura de valores como: máximos, mínimos, pontos discrepantes, etc. Numa segunda fase pode-se fazer o cálculo de indicadores como: média aritmética, variância de todos os termos ou dos termos correspondentes a troços sucessivos, distância média entre máximos ou mínimos locais, etc.
2. Explicação: Consiste na construção de modelos que permitam tentar explicar a evolução de uma série temporal a partir da variação de uma outra série, sendo um caso importante o das funções de transferência, que permitem explicar uma “série de saída (*output*)” como função de uma ou mais “séries de entrada(*inputs*)”.
3. Previsão: Como o termo sugere, consiste em prever valores futuros a partir de valores passados. Podem-se destacar duas famílias de métodos de previsão. Os primeiros são os métodos multivariáveis que procuram relacionar a série com outras séries a partir de

modelos. Os segundos são os métodos univariáveis, onde a previsão tem unicamente como base os valores actuais e passados da própria série.

4. Controlo: Se vemos uma série como o resultado (*output*) de um processo qualquer, a série pode dar informações úteis sobre o próprio processo com vista a poder manter o controlo sobre ele. Pode-se considerar a modo de exemplo, uma linha de produção onde são definidos valores limites para algumas características dos artigos produzidos. Caso os valores se mantenham nos limites estabelecidos, admite-se uma situação de controlo. Caso contrario o processo é parado para detecção e correcção dos factores que originaram o comportamento anómalo.

2.2 - Processos estocásticos estacionários

Uma série temporal [4; p24] gerada por um processo determinístico é aquela onde é possível prever exactamente valores futuros a partir de valores passados. Contudo muitos fenómenos dinâmicos correspondem ao conjunto dos chamados processos estocásticos ou aleatórios onde valores futuros são apenas parcialmente determinados a partir de valores passados. Os modelos criados para séries estocásticas são também denominados processos estocásticos.

Em geral um processo estocástico define-se como uma família de variáveis aleatórias $\{X_n\} = (X_1, X_2, \dots, X_n)$ indexadas no tempo. Isto quer dizer que se uma série temporal $\{x(n)\} = \{x(1), x(2), \dots, x(n)\}$ for gerada a traves do processo estocástico $\{X_n\}$, o valor $x(1)$ é um resultado aleatório que segue a distribuição de probabilidades de X_1 , o valor $x(2)$ é um resultado aleatório que segue a distribuição de probabilidades de X_2 , e assim sucessivamente.

Uma classe importante de processos estocásticos é os chamados processos estacionários. Seguindo [2; p23] e [5; p51] um processo estacionário é aquele onde o sistema se apresenta num estado de equilíbrio estatístico em torno de um nível médio fixo, ou seja, não apresenta tendência. Isto quer dizer que o sistema tem propriedades probabilísticas invariantes ao longo do tempo.

De maneira mais rigorosa, um processo estocástico $\{X_n\}$, é estritamente estacionário se a função densidade de probabilidade conjunta da família de variáveis aleatórias do processo, depende unicamente da localização relativa das variáveis. Isto quer dizer que se $f\{X_1, X_2, \dots, X_n\}$ denota a função densidade de probabilidade conjunta para o processo $\{X_n\}$, então para que o processo seja considerado estritamente estacionário deve verificar-se que,

$$f\{X_{1+k}, X_{2+k}, \dots, X_{n+k}\} = f\{X_1, X_2, \dots, X_n\}, \quad (2.1)$$

para qualquer k inteiro.

Contudo a estacionaridade no sentido estrito é uma propriedade muito exigente e de difícil verificação. Uma definição menos exigente de estacionaridade, corresponde a os processos fracamente estacionários até 2ª ordem. Neste caso é apenas necessário que as funções densidade de probabilidade das variáveis aleatórias possuam características comuns quando expressas nos termos até 2ª ordem.

Isto quer dizer que um processo estocástico $\{X_n\}$ é fracamente estacionário até 2ª ordem se a sua família de variáveis aleatórias verifica as seguintes propriedades para a média, variância e covariância,

$$\begin{aligned} E(X_k) &= \mu \text{ (constante),} \\ \text{Var}(X_k) &= \sigma^2 \text{ (constante),} \\ \text{Cov}(X_k, X_{k+1}) &= \text{Cov}(X_1, X_2), \end{aligned} \tag{2.2}$$

para qualquer k inteiro, sendo necessário que μ e σ^2 sejam valores finitos. Note-se que a covariância entre duas variáveis consecutivas do processo devem sempre ser iguais a covariância entre as duas primeiras variáveis do processo.

Um processo estocástico fracamente estacionário até 2ª ordem de particular importância é o chamado ruído branco $\{\epsilon_n\} = \{\epsilon_1, \epsilon_2, \dots, \epsilon_n\}$, que se caracteriza por ter correlação nula para variáveis correspondentes a instantes de tempo diferentes. De agora em diante uma série temporal gerada a partir de um processo de ruído branco será denotada por $\{\epsilon(n)\} = \{\epsilon(1), \epsilon(2), \dots, \epsilon(n)\}$.

2.3 - Processos Estacionários Lineares

Esta secção tem como finalidade a descrição de alguns modelos cujo objectivo é tentar modelar comportamentos de sistemas estacionários lineares. Os modelos que vão ser abordados [2; p 41] são baseados na seguinte ideia de Yule(1927): se existe dependência entre variáveis aleatórias referentes a sucessivos pontos do tempo pode conceber-se que o processo estocástico é gerado por consecutivos choques aleatórios (também designados inovações) causados por um processo de ruído branco.

2.3.1 - Modelos Autoregressivos [AR(p)]

Se diz que uma série $\{x(n)\} = \{x(1), x(2), \dots, x(n)\}$, corresponde a um processo AR(p) de ordem p quando verifica-se que,

$$x(n) = \phi_1 \cdot x(n-1) + \phi_2 \cdot x(n-2) + \dots + \phi_p \cdot x(n-p) + \epsilon(n), \tag{2.3}$$

onde $\epsilon(t)$ constitui ruído branco. Utilizando a transformação $z^{n_0} = x(n - n_0)$, O polinómio autoregressivo de ordem p e variável z , correspondente ao modelo AR(p) anterior, vem dado por,

$$\phi_p(z) = 1 - \phi_1 \cdot z - \phi_2 \cdot z^2 - \dots - \phi_p \cdot z^p. \tag{2.4}$$

Pode-se então rescrever a relação (2.3), como sendo $\phi_p(z) \cdot x(n) = \epsilon(n)$, onde o processo é sempre invertível. No que diz respeito a estacionaridade, um processo AR(p) será estacionário se o polinómio $\phi_p(z)$ não tem raízes no círculo de raio unitário.

2.3.2 - Modelos Médias Moveis [MA(q)]

Se diz que uma série $\{x(n)\} = \{x(1), x(2), \dots, x(n)\}$, corresponde a um processo MA(q) de ordem q quando verifica-se que,

$$x(n) = \epsilon(n) - \theta_1 \cdot \epsilon(n-1) - \theta_2 \cdot \epsilon(n-2) - \dots - \theta_q \cdot \epsilon(n-q), \quad (2.5)$$

onde $\epsilon(n)$ constitui ruído branco. De forma análoga aos processos AR(p), o polinómio médias moveis de ordem q e variável z , vem dado por,

$$\theta_q(z) = 1 - \theta_1 \cdot z - \theta_2 \cdot z^2 - \dots - \theta_q \cdot z^q. \quad (2.6)$$

Pode-se então rescrever a relação (2.5), como sendo $x(n) = \theta_q(z) \cdot \epsilon(n)$. Nos modelos MA(q) a ideia principal é, mais uma vez, tentar exprimir o processo $x(n)$ em termos de um processo mais simples e puramente aleatório como é o ruído branco.

Em termos de estacionaridade, os processos MA(q) são sempre estacionários desde que os coeficientes $\{\theta_i\} = \{\theta_1, \theta_2, \dots\}$, sejam números reais finitos. Por último um processo MA(q) será invertível se o polinómio $\theta_q(z)$ não tiver raízes no círculo de raio unitário.

2.3.3 - Modelos Mistos Autoregressivos e Médias Moveis [ARMA(p,q)]

Como se viu até agora, um processo estacionário e invertível pode representar-se na forma autoregressiva ou na forma médias móveis. Contudo há situações onde o sistema gerador da série temporal observada, tem uma estrutura que resulta da combinação das duas formas de modelos anteriores.

Sendo assim, diz-se que uma série $\{x(n)\} = \{x(1), x(2), \dots, x(n)\}$, corresponde a um processo ARMA(p,q) quando se verifica que,

$$x(n) = \phi_1 \cdot x(n-1) + \phi_2 \cdot x(n-2) + \dots + \phi_p \cdot x(n-p) + \epsilon(n) - \theta_1 \cdot \epsilon(n-1) - \theta_2 \cdot \epsilon(n-2) - \dots - \theta_q \cdot \epsilon(n-q), \quad (2.7)$$

ou de maneira equivalente $\phi_p(z) \cdot x(n) = \theta_q(z) \cdot \epsilon(n)$, onde $\epsilon(n)$ constitui ruído branco e onde $\phi_p(z)$ e $\theta_q(z)$ correspondem, respectivamente, a os polinómios autoregressivo e médias moveis, mencionados anteriormente.

Como seria de esperar, um processo ARMA(p,q) será estacionário se as raízes do polinómio $\phi_p(z)$ estão fora do círculo de raio unitário e será invertível se as raízes do polinómio $\theta_q(z)$ estão fora do círculo de raio unitário.

2.4 - Correlação em séries temporais

O coeficiente de correlação amostral mede o grau de correlação entre os valores de duas séries temporais. Seguindo [1] e [6; p23], dadas duas séries temporais $\{x(n)\} = \{x(1), x(2), \dots, x(n)\}$ e $\{y(n)\} = \{y(1), y(2), \dots, y(n)\}$, o coeficiente de correlação amostral r entre $\{x(n)\}$ e $\{y(n)\}$ vem dado por,

$$r = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{i=1}^n (y_i - \bar{y})^2}}, \quad (2.8)$$

onde $\bar{x} = \sum_{i=1}^n x_i / n$ e $\bar{y} = \sum_{i=1}^n y_i / n$ representam respectivamente as médias de $\{x(n)\}$ e $\{y(n)\}$.

Na prática, em muitas situações vamos a estar interessados em medir o grau de correlação, não entre duas séries diferentes, mas sim entre uma série e a mesma série desfasada no tempo. Neste caso não falamos de um coeficiente de correlação amostral, mas sim de vários coeficientes de autocorrelação, cada um correspondente a um desfasamento temporal específico. Sendo assim, para uma dada série temporal $\{x(n)\}$ de média \bar{x} , o coeficiente de autocorrelação r_k correspondente ao desfasamento k vem dado por,

$$r_k = \frac{\sum_{i=1}^{n-k} (x_i - \bar{x})(x_{i+k} - \bar{x})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2}}. \quad (2.9)$$

2.4.1 - O Correlograma

Uma poderosa ferramenta usada para extrair características duma série temporal é o chamado correlograma, que não é mais do que o gráfico que faz corresponder a cada coeficiente de autocorrelação r_k o seu respectivo desfasamento temporal k . Contudo a interpretação adequada do correlograma nem sempre é uma tarefa simples. De seguida são mencionadas algumas indicações a respeito deste conceito.

2.4.2 - Correlograma em séries aleatórias

Uma das primeiras questões a tentar descobrir através do correlograma é se a série é aleatória ou não. Quando uma série é totalmente aleatória, os valores sucessivos não guardam correlação entre si, pelo que seria de esperar que todos os coeficientes r_k fossem iguais a zero ou muito próximos desse valor. De facto pode-se demonstrar que os coeficientes de autocorrelação vão ter uma distribuição com média ≈ 0 e variância $\approx 1/n$, o qual implica limites de confiança de aproximadamente 95%. Isto ilustra a dificuldade de interpretação do correlograma, já que mesmo para o caso de séries completamente aleatórias um limite de confiança de 95% implica que e de esperar que um de cada vinte coeficientes esteja fora dos limites esperados.

2.4.3 - Correlograma de curto prazo

Uma série temporal tem correlação de curto prazo quando um valor acima da média tende a ser seguido por um ou mais valores acima da média, da mesma maneira um valor abaixo da média tende a ser seguido por um ou mais valores abaixo da média. Neste caso é de esperar que valores altos do correlograma sejam seguidos por valores com tendência a ficar cada vez menores até se aproximarem de zero e, de maneira análoga, é de esperar que valores baixos do correlograma sejam seguidos por valores com tendência a ficar cada vez maiores até se aproximarem de zero.

2.4.4 - Correlograma em séries alternadas

Se uma série temporal tem tendência a alternar em redor de um valor médio, então o correlograma também vai ter este comportamento. A título de exemplo, pode-se ver na seguinte ilustração, as características antes mencionadas para um segmento de um sinal sinusoidal e a sua respectiva autocorrelação.

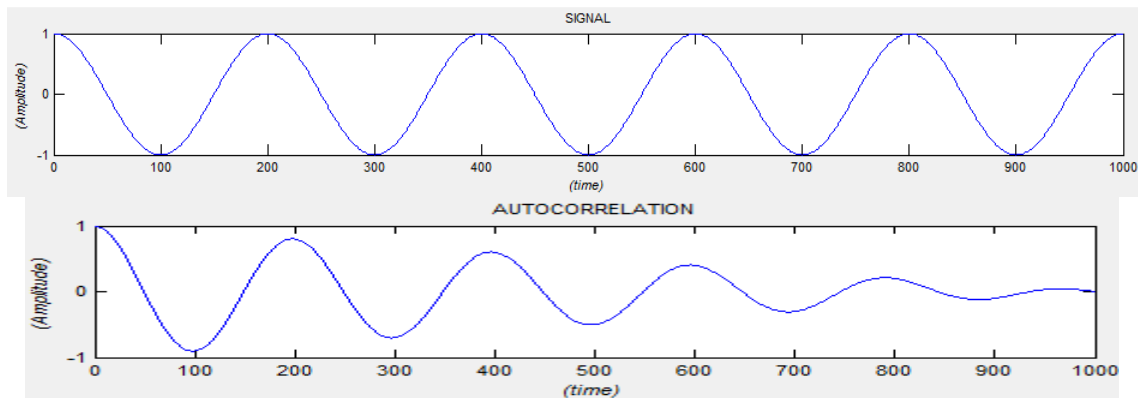


Figura 2.2 - Segmento de sinal sinusoidal e respectiva função autocorrelação.

Cabe ainda sublinhar que na ilustração anterior, o decaimento da função autocorrelação, deve-se ao facto de estar a considerar apenas um segmento de sinal, não um sinal infinito.

2.4.5 - Correlograma em séries não estacionárias

Para uma série temporal com tendência, os valores de r_k não decairão até zero a não ser para desfasamentos grandes. Isto deve-se ao facto de que quando uma série apresenta tendência, vai haver muitos valores consecutivos agrupados em cima ou em baixo da média geral. Neste caso pouca ou nenhuma informação pode ser extraída porque a que a tendência dominará as outras características, pelo que a autocorrelação só terá algum significado em séries estacionárias.

2.5 - Espectro em séries temporais

Até agora tem-se estudado as propriedades das séries temporais no domínio do tempo. A análise no domínio das frequências, embora relacionada com a análise no domínio do tempo, tem uma perspectiva diferente. Em vez de se concentrar [2; p169] na relação de dependência entre valores correspondentes a diferentes instantes no tempo, procura investigar quais são as diferentes frequências ou componentes periódicas que contribuem para a variabilidade da série temporal.

2.5.1 - Série e Transformada de Fourier Discreta (DFT)

Uma série temporal $\{x(n)\}$ é periódica [7; p211] com período N quando se verifica a condição $x(n) = x(n+N)$. O período fundamental corresponde ao menor inteiro positivo N para o qual se verifica a relação anterior e $w_0 = 2\pi/N$ é definida como a frequência fundamental. O conjunto de sinais exponenciais complexos discretos periódicos, de período N , é dado por $\Phi_k[n] = e^{jkw_0n} = e^{jk(2\pi/N)n}$, para $k = 0, \pm 1, \pm 2, \dots, \pm(N-1) = \langle N \rangle$; onde todos estes sinais possuem frequências fundamentais que são múltiplos de $2\pi/N$.

Em [7] demonstra-se, que uma série periódica $\{x(n)\}$ de período fundamental N pode ser expressa como uma soma de sinais exponenciais complexos, obtendo-se assim o par de relações que correspondem, respectivamente, a série de Fourier para o caso discreto e a sua função inversa,

$$\begin{aligned} a_k &= \frac{1}{N} \sum_{n=\langle N \rangle} x(n) \cdot e^{-jkw_0n} = \frac{1}{N} \sum_{n=\langle N \rangle} x(n) \cdot e^{-jk(2\pi/N)n} , \\ x(n) &= \sum_{k=\langle N \rangle} a_k \cdot e^{jkw_0n} = \sum_{k=\langle N \rangle} a_k \cdot e^{jk(2\pi/N)n} . \end{aligned} \quad (2.10)$$

Contudo na natureza, muitas séries apresentam comportamentos aperiódicos, o qual impossibilita a sua representação em série de Fourier. Nestes casos, tal e como é demonstrado em [7], pode-se expressar uma série aperiódica $\{x(n)\}$ em função de sinais exponenciais complexos, obtendo-se assim o par de relações que correspondem, respectivamente, a transformada de Fourier discreta e a sua função inversa,

$$\begin{aligned} X(e^{jw}) &= \sum_{n=-\infty}^{+\infty} x(n) \cdot e^{-jwn} , \\ x(n) &= \frac{1}{2\pi} \int_{2\pi} X(e^{jw}) \cdot e^{jwn} dw . \end{aligned} \quad (2.11)$$

2.5.2 - Análise espectral em processos estocásticos estacionários

Para o estudo que se segue vale a pena relembrar os conceitos de autocovariancia e autocorrelação para processos estocásticos estacionários. Dado [2; p27] um processo estocástico estacionário $\{X_n\} = (X_1, X_2, \dots)$ de média μ e variancia σ^2 , a função autocovariancia γ_k mede a

intensidade com que covariam pares de valores do processo, correspondentes a instantes de tempo com um desfasamento k . Sendo assim a função autocovariância γ_k vem dada por,

$$\gamma_k = E\{(X_n - \mu)(X_{n+k} - \mu)\}, \quad (2.12)$$

onde $E\{\}$ representa o valor esperado. A partir da definição anterior, pode-se ainda obter a função autocorrelação ρ_k para o mesmo desfasamento k como sendo $\rho_k = \gamma_k / \gamma_0$, onde $\gamma_0 = \sigma^2$.

Estamos agora em condições de definir [2; p183] a densidade espectral $f(w)$, para uma série de valores obtidos a partir de um processo estocástico estacionário como sendo,

$$f(w) = \frac{1}{2\pi} \sum_{k=-\infty}^{\infty} \gamma_k \cdot e^{-jwk}, \quad (2.13)$$

Onde k é qualquer valor inteiro, $w \in (-\pi, +\pi)$ e γ_k é a respectiva função autocovariância do processo, sendo necessário que $\sum_{k=-\infty}^{\infty} |\gamma_k| < \infty$. Por último fazendo $f(w)/\sigma^2$ obtemos a densidade espectral normalizada $f_N(w)$ como sendo,

$$f_N(w) = \frac{1}{2\pi} \sum_{k=-\infty}^{\infty} \rho_k \cdot e^{-jwk}. \quad (2.14)$$

2.6 – Vantagens e limitações dos modelos lineares estacionários

Em [22; p6] e [20] são referenciadas vantagens e limitações dos modelos lineares estacionários ARMA. Algumas das vantagens deste tipo de modelos são:

- As equações diferenças lineares, base dos modelos ARMA, são as equações diferenças mais simples que existem.
- O tempo de processamento computacional necessário para a geração deste tipo de modelos é relativamente baixo.
- Ao longo de muitos anos estes modelos têm permitido resultados razoavelmente satisfatórios, quando utilizados como ferramentas para análise, previsão e controlo de dados.

Por outro lado algumas das desvantagens dos modelos lineares estacionários ARMA são:

- Possuindo uma distribuição conjunta simétrica, os modelos ARMA não são os mais indicados para dados fortemente assimétricos.
- Características nos dados como valores muito afastados da média, valores de pico súbitos ou variações de grande amplitude fazem este tipo de dados pouco interessantes para a modelação linear.

A detecção de uma componente não estocástica ou determinística pode ser levada a cabo usando uma versão discreta da ferramenta introduzida por Henri Poincaré, denominada diagramas de fase. A esta última ferramenta adiciona-se um conjunto de outras ferramentas, que surgiram no contexto da análise das séries não lineares e que serão abordadas no capítulo seguinte.

Capítulo 3

Caos em Sistemas dinâmicos

Um sistema dinâmico [13; p6] é aquele que pode ser definido através de um modelo matemático que envolve a evolução do estado do sistema ao longo do tempo. Isto quer dizer que num sistema dinâmico os vários estados do sistema ao longo do tempo estão relacionados entre si.

Um sistema dinâmico unidimensional em tempo contínuo é aquele que pode ser representado através da equação diferencial $dx(t)/dt = F[x(t)]$, onde t é uma variável contínua. De maneira análoga a representação de um sistema dinâmico unidimensional em tempo discreto pode ser representado através da equação as diferenças $x_{n+1} = F(x_n)$, onde n é uma variável discreta.

Pode-se ainda generalizar as representações anteriores para o caso de sistemas N -dimensionais, sendo neste caso os sistemas definidos, não por uma equação, mas sim por um sistema de N equações, que no caso de tempo contínuo adquirem a forma,

$$\begin{aligned}\frac{dx_1}{dt} &= F_1(x_1, x_2, \dots, x_N), \\ \frac{dx_2}{dt} &= F_2(x_1, x_2, \dots, x_N), \\ &\dots \\ \frac{dx_N}{dt} &= F_N(x_1, x_2, \dots, x_N),\end{aligned}\tag{3.1}$$

e que no caso de tempo discreto adquirem a forma,

$$\begin{aligned}x_{1(n+1)} &= F_1(x_{1(n)}, x_{2(n)}, \dots, x_{N(n)}), \\ x_{2(n+1)} &= F_2(x_{1(n)}, x_{2(n)}, \dots, x_{N(n)}), \\ &\dots \\ x_{N(n+1)} &= F_N(x_{1(n)}, x_{2(n)}, \dots, x_{N(n)}),\end{aligned}\tag{3.2}$$

sendo as variáveis x_1, x_2, \dots, x_n denotadas de agora em adiante como as variáveis de estado do sistema. A título de exemplo, considere-se o sistema formado por um pêndulo forçado amortecido,

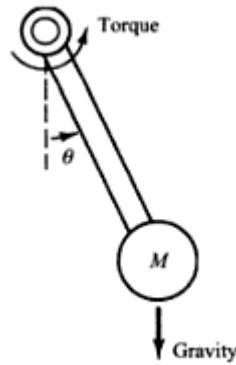


Figura 3.1 - pêndulo forçado amortecido [13; p7].

cujo comportamento é dado pela equação diferencial,

$$\frac{d^2\theta}{dt^2} + v \frac{d\theta}{dt} + \sin(\theta) = T \sin(2\pi ft). \quad (3.3)$$

Utilizando a mudança de variáveis $x_1 = \frac{d\theta}{dt}$, $x_2 = \theta$, $x_3 = 2\pi ft$, podemos rescrever a equação diferencial (2.3) como sendo equivalente a um sistema de 3 equações diferenciais de 1ª ordem, que vem dado por,

$$\begin{aligned} \frac{dx_1}{dt} &= T \sin(x_3) - \sin(x_2) - vx_1, \\ \frac{dx_2}{dt} &= x_1, \\ \frac{dx_3}{dt} &= 2\pi f, \end{aligned} \quad (3.4)$$

onde, a partir da representação anterior, pode-se concluir que o sistema tem três variáveis de estado, correspondendo então a um sistema tridimensional.

3.1 - Caos

Até recentemente, os sistemas dinâmicos [16] eram classificados em três categorias, segundo o padrão de variação no tempo das grandezas que caracterizam os seus estados. Essas categorias eram:

- Sistemas estáveis, convergindo para um valor fixo.
- Sistemas periódicos, estabelecendo-se em oscilações periódicas.
- Sistemas ruidosos, imprevisíveis, caracterizados por flutuações irregulares.

Porém, em 1963, Edward Lorenz fez uma descoberta que surpreendeu o mundo, enquanto estudava um modelo de previsão do tempo. Seu modelo seguia um padrão que não se encaixava em nenhum dos anteriores, exibindo um comportamento bastante complexo, embora fosse definido apenas por poucas e simples equações diferenciais. Sistemas como o de Lorenz são denominados “caótico-determinísticos” ou seja, embora apresentem um comportamento aperiódico e imprevisível, a sua dinâmica é governada por equações diferenciais ou equações diferenciais determinísticas simples.

A principal característica [15; p15] no comportamento de sistemas caóticos é uma elevada sensibilidade nas condições iniciais, o que implica que a evolução do sistema pode ser alterada por pequenas perturbações. Esta propriedade faz com que o comportamento caótico possua uma grande flexibilidade, sendo interessante para sistemas que necessitam uma reacção rápida a determinadas perturbações. Outra característica fundamental deste tipo de sistemas é que a sua representação no espaço de fases corresponde a um atrator estranho, assunto que será abordado na continuação deste capítulo.

3.2 - Representações no espaço de fases

Define-se [16] por espaço de fases a um sistema de coordenadas associado às variáveis de estado que descrevem a dinâmica dum sistema. Sendo assim as coordenadas num ponto do espaço de fases, para um determinado instante, correspondem aos valores das variáveis de estado do sistema para esse instante. Tais coordenadas são comumente designadas por vectores de estado.

3.2.1 - Pontos fixos

Um conceito importante na análise do espaço de fases é o chamado ponto fixo [14;p79]. No caso de sistemas unidimensionais contínuos os pontos fixos são aqueles onde $dx/dt = 0$, condição que equivale a $x(n+1) = x(n)$ para o caso de sistemas unidimensionais discretos.

Pode-se identificar três tipos possíveis de pontos fixos:

1. Pontos fixos estáveis: São aqueles que atraem as trajectórias próximas a eles.
2. Pontos fixos instáveis: São aqueles que repelem as trajectórias próximas a eles.
3. Pontos de sela: São aqueles que atraem as trajectórias provenientes numa direcção mas repelem as trajectórias provenientes de outras direcções.

3.2.2 - Atratores

O atrator é a representação [16] da dinâmica de um sistema no espaço de fases. Alguns exemplos de atratores são:

- Atrator Ponto fixo: Correspondem a sistemas estáveis, cuja representação no espaço de fases converge ao longo do tempo para um estado fixo.
- Atrator Ciclo limite: Correspondem a sistemas periódicos, cuja representação no espaço de fases esta definida por uma órbita fechada, composta por dois ou mais pontos.
- Atrator estranho: Correspondem a sistemas caóticos, cuja representação no espaço de fases tem órbitas que nunca repetem o mesmo caminho; contudo, as órbitas estão confinadas (atraídas) a uma região limitada do espaço de fases. Muitas vezes é-lhes atribuída uma dimensão fractal.

3.3 - Mapa logístico

Um bom exemplo de um sistema que pode gerar padrões complexos apesar de ter uma representação simples é o chamado mapa logístico, que vem referenciado em [12; p22] e [17], e que não é mais do que um sistema unidimensional definido por,

$$x(n + 1) = A \cdot x(n) \cdot [1 - x(n)] , \quad (3.5)$$

onde à medida que se vai mudando o valor da constante A , também se vai mudando o comportamento do sistema. A seguir vamos analisar alguns casos, para diferentes valores de A , dada a condição inicial $X_0 = 0.55$, sendo os espaços de fase apresentados, aqueles que fazem corresponder $x(n+1)$ a $x(n)$.

1. Cronograma e espaço de fases do mapa logístico para $A = 2.9$.

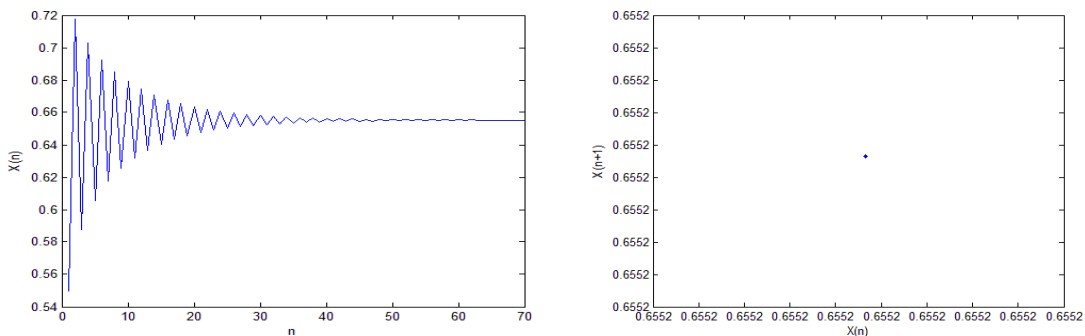


Figura 3.2 - Mapa logístico para $A = 2.9$.

Para $A = 2.9$, o sistema é estável e $x(n)$ converge para um ponto de equilíbrio quando $n \rightarrow \infty$. A representação no espaço de fases corresponde a um atrator do tipo ponto fixo, onde só aparece um ponto fixo estável, quando tomados apenas valores de $x(n)$ para $n \rightarrow \infty$.

2. Cronograma e espaço de fases do mapa logístico para $A = 3.2$.

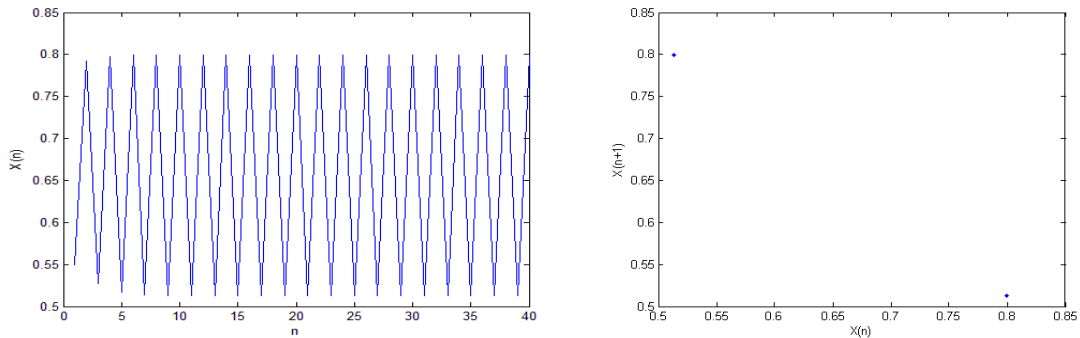


Figura 3.3 - Mapa logístico para $A = 3.2$.

Para $A = 3.2$, o sistema torna-se instável periódico e $x(n)$ oscila entre dois valores fixos. A representação no espaço de fases corresponde a um atrator do tipo ciclo limite, onde o sistema oscila entre dois estados, quando tomados apenas valores de $x(n)$ para $n \rightarrow \infty$.

3. Cronograma e espaço de fases do mapa logístico para $A = 3.5$.

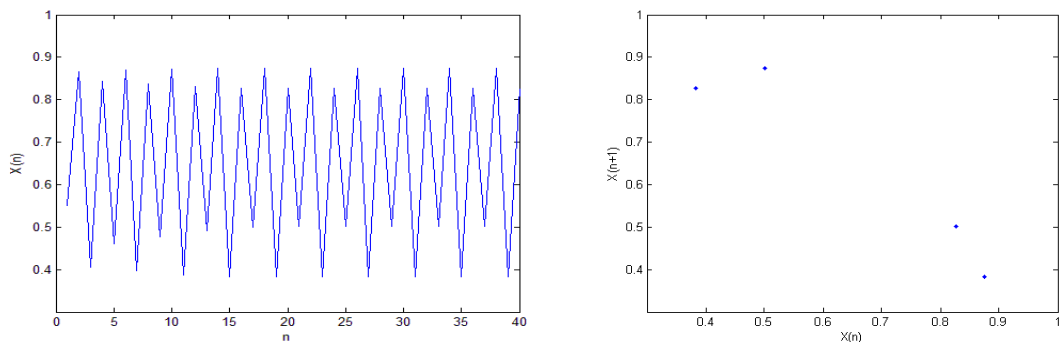


Figura 3.4 - Mapa logístico para $A = 3.5$.

Para $A = 3.5$, o sistema continua instável periódico e $x(n)$ oscila agora, não entre dois valores fixos, mais sim entre quatro valores fixos. A semelhança do caso anterior, a representação no espaço de fases corresponde a um atrator do tipo ciclo limite, onde o sistema oscila entre quatro estados, quando tomados apenas valores de $x(n)$ para $n \rightarrow \infty$.

Chegada esta altura, a medida que vamos aumentando o valor de A , começa também a aumentar o período de $x(n)$ e a aumentar o número de estados possíveis do sistema.

4. Cronograma e espaço de fases do mapa logístico para $A = 3.57$.

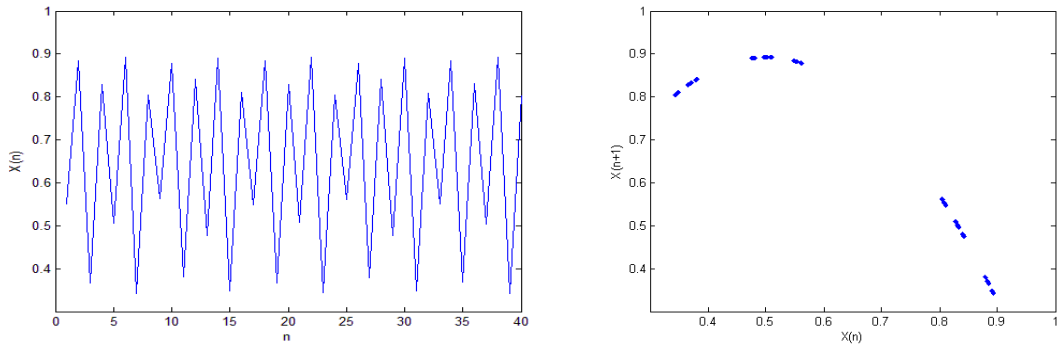


Figura 3.5 - Mapa logístico para $A = 3.57$.

Para $A = 3.57$, o sistema começa a dar os primeiros indícios de caos, começando a oscilar entre cada vez mais estados possíveis.

5. Cronograma e espaço de fases do mapa logístico para $A = 4$.

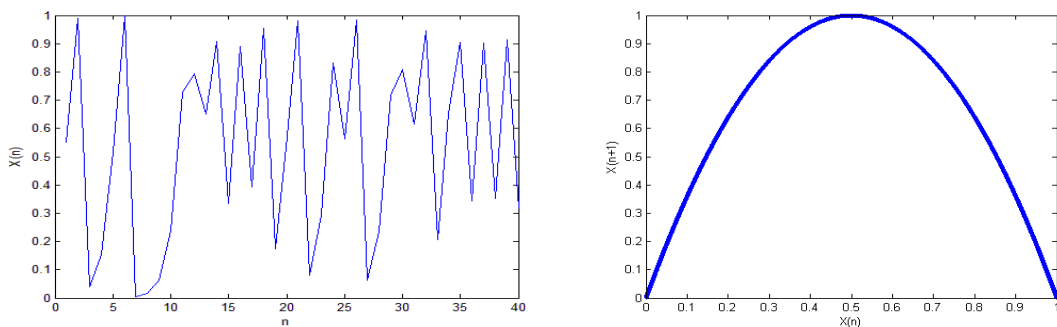


Figura 3.6 - Mapa logístico para $A = 4$.

Para $A = 4$, o sistema já não é periódico, revela-se caótico, passando por infinitos estados possíveis entre 0 e 1 onde. A representação no espaço de fases corresponde a um atrator do tipo caótico onde todas as órbitas, apesar de nunca percorrer o mesmo caminho, estão confinadas a uma região do espaço com forma parabólica.

O termo bifurcação [15; p145] está associado a uma mudança qualitativa na evolução dum sistema como consequência da variação dos seus parâmetros. O fenómeno da bifurcação está estreitamente relacionado com a existência de caos uma vez que um sistema caótico sempre apresenta algum tipo de bifurcação.

À seguinte ilustração corresponde ao diagrama de Bifurcação para o mapa logístico, onde consegue-se ver por quantos estados possíveis passa $x(n)$ quando $n \rightarrow \infty$, para um dado valor de A . Por cada bifurcação ocorre uma duplicação do período e do número de estados possíveis até o sistema entrar em regime caótico.

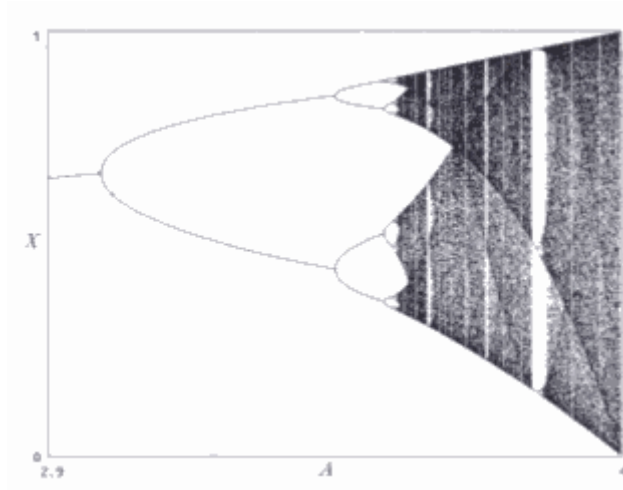


Figura 3.7 - Diagrama de Bifurcação para o mapa logístico [12; p28].

3.4 - Reconstrução do espaço de fases

Num sistema cujas equações dinâmicas são conhecidas, a caracterização do seu atrator associado é relativamente simples. Contudo em muitos casos desconhecemos as equações dinâmicas que governam um determinado sistema, limitando-nos apenas a conhecer uma série de valores, correspondentes à evolução temporal de uma das variáveis do sistema.

Cabe agora perguntar-se, será que podemos extrair características do sistema ou inclusive reconstruir o seu atrator a partir de uma única série temporal, correspondente a uma das suas variáveis? A resposta a esta pergunta prende-se com o teorema de Takens, que vem referenciado em [16] e segundo o qual pode-se reconstruir a trajectória completa de um sistema num espaço de fases, a partir da medida de uma única variável do sistema. O método é baseado na obtenção de vectores de estado atrasados da série temporal original $\{x(n)\}$, de modo que o espaço de fases passe a ser definido pelo conjunto de vectores $y(k)$ definidos por,

$$y(k) = \{x(k), x(k + L), \dots, x(k + [d - 1]L)\}, \quad (3.6)$$

onde L é o tempo de atraso de Takens e d a dimensão de imersão do espaço de fases, sendo ambos parâmetros estudados em detalhe mais a frente.

O mapa de Hénon, referenciado em [12; p421] é um bom exemplo de como é possível reconstruir a geometria de um atrator, partindo apenas dos valores observados para uma das suas variáveis. Tal sistema é dado pelas equações,

$$\begin{aligned} x(n + 1) &= 1 - a \cdot x^2(n) + b \cdot y(n), \\ y(n + 1) &= x(n), \end{aligned} \quad (3.7)$$

sendo $x(n)$ e $y(n)$ as suas variáveis de estado. Na seguinte ilustração é possível ver primeiramente, o atrator do mapa de Hénon definido pelos valores das variáveis de estado $\{x(n), y(n)\}$ e logo a seguir, o atrator reconstruído apenas com os valores da variável $x(n)$ utilizando o teorema de Takens com $L=1$ e $d=2$. Foram usados os parâmetros $a=1.4$, $b=0.3$ e as condições iniciais $x(1)=0$, $y(1)=0.9$.

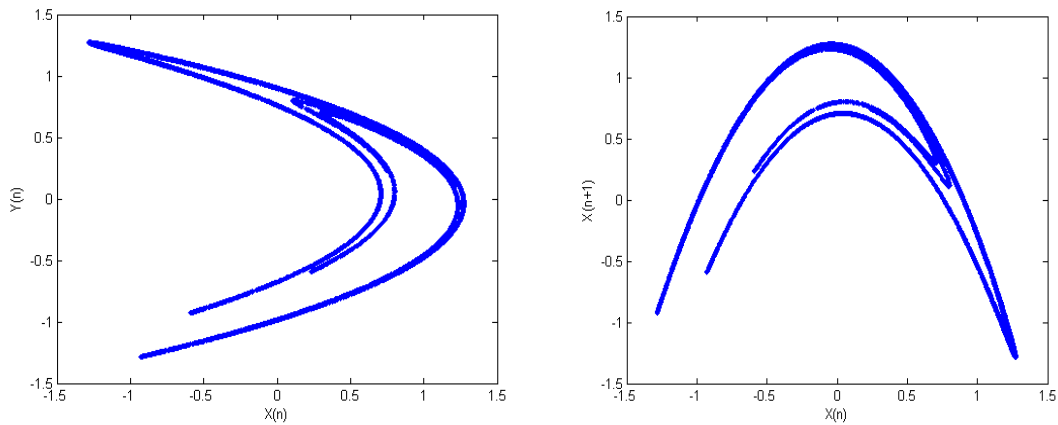


Figura 3.8 - Mapa de Hénon e Mapa de Hénon reconstruído.

Tal como se pode ver, a reconstrução do atrator a partir do teorema de Takens, para valores específicos de L e d , preserva a geometria do atrator original. Cabe destacar que a representação $\{x(n), y(n)\}$ equivalente a $\{x(n+1), y(n+1)\}$ neste caso vem a ser igual a $\{x(n+1), x(n)\}$, motivo pelo qual o atrator original corresponde exactamente ao atrator reconstruído com os “eixos trocados”.

3.4.1 - Escolha do desfaseamento temporal

Nos primeiros trabalhos de reconstrução do espaço de fases [15; p193] para uma série temporal $\{x(n)\}$ era traçado o gráfico que fazia corresponder a $x(n)$ o valor $x(n+1)$, mas um problema dessa abordagem era a alta correlação que existe entre valores muito próximos da série temporal, resultando em um espaço de fases fortemente alinhado na direcção da bissectriz do plano, causando resultados distorcidos. Por outro lado se o valor do desfaseamento for muito elevado, os valores poderiam ficar muito descorrelacionados.

Dois métodos se destacam na determinação do desfaseamento L , um baseia-se na função autocorrelação (estudada anteriormente) e o outro baseia-se no método da informação mútua média. O primeiro método consiste em atribuir como valor para o desfaseamento, o valor correspondente ao primeiro zero da função autocorrelação, onde verifica-se uma não correlação linear entre $x(n)$ e $x(n+L)$. Contudo para sistemas que apresentam comportamentos não lineares, uma medida de dependência não linear, como a informação mútua média (IMM), será mais efectiva.

O método da informação mútua média é descrito em [18], sendo também explicado em [19] e [21]. A IMM indica em que grau, uma série temporal contém informação acerca de outra série temporal. Dada uma série temporal $\{x(n)\}$ de comprimento N e uma outra série temporal $\{y(m)\}$ de comprimento M , a IMM entre $\{x(n)\}$ e $\{y(m)\}$ vem dada por,

$$I(x; y) = \sum_{i=1}^N \sum_{j=1}^M P[x(i), y(j)] \cdot \log_2 \left\{ \frac{P[x(i), y(j)]}{P[x(i)] \cdot P[y(j)]} \right\}, \quad (3.8)$$

onde $P[x(i)]$ é a probabilidade da medida $x(i)$, $P[y(j)]$ é a probabilidade da medida $y(j)$ e finalmente, $P[x(i), y(j)]$, é a probabilidade conjunta das medidas $[x(i), y(j)]$. No nosso caso estamos interessados em medir o grau de informação que há entre uma série temporal $\{x(n)\}$ e a mesma série temporal desfasada L posições. Ao substituir $y(j) = x(j+L)$ na expressão (3.8), obtêm-se informação mútua média de $\{x(n)\}$, em função do desfasamento L .

Em [18], o valor recomendado para o desfasamento, na reconstrução do espaço de fases, é o valor de L para o qual a função informação mútua média tem o seu primeiro valor de mínimo.

3.4.2 - Escolha da dimensão de imersão do espaço de reconstrução

A dimensão de imersão do espaço de reconstrução corresponde ao número mínimo de coordenadas do espaço, necessárias para representar correctamente a dinâmica do sistema. Pretende-se então determinar [20], a dimensão necessária para que as órbitas geradas pela série temporal não se autointersectem.

Em [8; p40] e [9], é descrito o algoritmo para o cálculo da dimensão ideal do espaço de reconstrução, a partir do método dos falsos vizinhos próximos. Seguindo a representação (3.6) e calculando o valor de L com base no método da informação média mútua, vamos definir o ponto $y^{NN}(k) = \{x^{NN}(k), x^{NN}(k+L), \dots, x^{NN}(k+[d-1]L)\}$ como sendo o ponto com menor distancia euclidiana em relação a $y(k)$ e, por tanto, o vizinho mais próximo de $y(k)$ no espaço de fases. Sendo assim o quadrado da distancia entre $y(k)$ e $y^{NN}(k)$ vem dado por,

$$R_d^2(k) = \{x(k) - x^{NN}(k)\}^2 + \{x(k+L) - x^{NN}(k+L)\}^2 + \dots + \{x(k+[d-1]L) - x^{NN}(k+[d-1]L)\}^2. \quad (3.9)$$

Cabe agora a pergunta, qual seria o acréscimo da distancia entre $y(k)$ e $y^{NN}(k)$ se aumentarmos a dimensão do espaço de d para $d+1$?. Na verdade é esta a questão chave para determinar a dimensão óptima do espaço de fases a partir deste método. O que acontece é que se houver um acréscimo considerável na distancia entre $y(k)$ e $y^{NN}(k)$ ao aumentar a dimensão, isto vai querer dizer que na realidade os pontos são o resultado de projecções em planos de dimensão inferior à ideal, sendo então na verdade falsos vizinhos.

Continuando o raciocínio, o quadrado da distancia entre $y(k)$ e $y^{NN}(k)$ ao passar de d para $d+1$ vem dado por $R_{d+1}^2(k) = R_d^2(k) + \{x(k+dL) - x^{NN}(k+dL)\}^2$, isto quer dizer que o acréscimo no quadrado da distancia ao aumentar a dimensão será,

$$A^2_{d+1} = \{x(k+dL) - x^{NN}(k+dL)\}^2. \quad (3.10)$$

Ficamos agora finalmente em condições de escrever o critério que define se dois pontos são falsos vizinhos. Sendo assim dois pontos $y(k)$ e $y^{NN}(k)$ são falsos vizinhos, se ao aumentar a dimensão do plano de d para $d+1$ se verifica que $A^2_{d+1}/R_d^2(k) > R_t$, o que equivale a,

$$\frac{|x(k+dL)-x^{NN}(k+dL)|}{|R_d(k)|} > R_t, \quad (3.11)$$

onde $R_t = 15$, na prática para uma larga variedade de sistemas, resulta ser um bom valor de comparação.

Uma vez avaliados todos os pontos do espaço de fases e os seus vizinhos próximos, ao passar de d para $d+1$, a dimensão do espaço de reconstrução óptima será aquela para a qual o número de falsos vizinhos tende para 0.

Uma última nota recai sobre o caso de séries geradas por processos aleatórios. Nestes casos a condição anterior pode levar-nos a escolha duma dimensão menor a dimensão ideal, pelo que nestas situações também é adoptado um outro critério que compara a distância com o tamanho do atractor no espaço de fases. Este outro critério vem dado por,

$$\frac{R_{d+1}}{R_A} \geq 2, \quad (3.12)$$

onde comumente é usada a expressão $R^2_A = \frac{1}{N} \sum_{k=1}^N [s(k) - \bar{s}]^2$ para achar o valor de R_A , sendo \bar{s} a média aritmética dos valores da série.

3.5 - Expoentes de Lyapunov

O expoente de Lyapunov [11] é um parâmetro que caracteriza a dinâmica de um atractor. Ele mede a taxa de divergência de órbitas vizinhas dentro do atractor e, assim, quantifica a dependência ou sensibilidade do sistema em relação as condições iniciais. A existência de pelo menos um expoente de Lyapunov positivo, é um forte indício da presença de caos no sistema.

Existem vários métodos [16] para o cálculo de expoentes de Lyapunov (Wolf, Eckmann e Ruelle, Brown e Bryant), sendo o método de wolf o mais amplamente considerado na literatura e aquele que vai ser descrito a seguir com base em [10; p255] e [11]. Vamos também concentrar o nosso estudo apenas no cálculo do maior expoente de Lyapunov, visto ser este o mais importante em termos de informação.

Considere-se a série temporal $\{x(k)\}$ e a sua representação no espaço de fases $\{y(k)\}$, onde os pontos do espaço são dados pela expressão (3.6).

Sendo assim define-se uma trajectória fiducial $\{y(0), y(1), \dots\}$ que, como se pode ver em [12, p116], corresponde à primeira órbita do espaço de fases. Seja o ponto $z(0)$ o vizinho mais próximo de $y(0)$ e L_0 a distancia entre eles de maneira a verificar $L_0 = |y(0) - z(0)| < \varepsilon$, onde ε é um valor arbitrário. De seguida acompanha-se a evolução temporal de $y(0)$ e $z(0)$ nas suas respectivas trajectórias, até que num instante t_1 a distancia L_{t_1} entre $y(t_1)$ e $z(t_1)$ exceda ε . Neste

momento substitui-se $z(t)$ por um novo ponto $z'(t)$ que seja o vizinho mais próximo de $y(t)$ na direcção do segmento L_0' , de maneira a verificar novamente $L_i = |y(t) - z'(t)| < \varepsilon$. O processo prossegue até que todos os pontos da trajectória fiducial sejam percorridos.

Uma vez acabado o processo, o maior expoente de Lyapunov λ_1 é obtido como a média de $\ln(L_{i+1}/L_i)$ ao longo da trajectória fiducial, ou seja,

$$\lambda_1 = \frac{1}{t_M - t_0} \sum_{i=0}^{M-1} \log_2 (L_{i+1}/L_i), \quad (3.13)$$

onde M é o número total de vezes que se escolheu um novo vizinho e $t_M - t_0$ representa o tempo transcorrido ao percorrer a trajectória fiducial desde $y(0)$ até $y(t_M)$.

Na seguinte ilustração pode-se visualizar como o maior expoente de Lyapunov é estimado a partir da taxa de crescimento dos segmentos L_i . Quando o comprimento do segmento L_i' entre as duas trajectórias excede ε , um novo ponto vizinho é escolhido de forma a minimizar o ângulo θ_i .

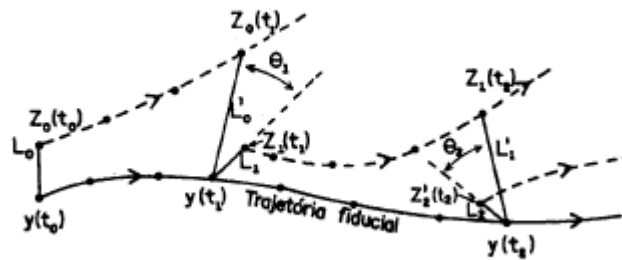


Figura 3.9 - Cálculo do maior expoente de Lyapunov [11].

As operações necessárias para o cálculo do maior expoente de Lyapunov e para alguns dos parâmetros mencionados anteriormente são de natureza iterativa. Isto faz com que seja indispensável o uso de aplicações informáticas como ferramentas de apoio na realização dos cálculos, sendo este tópico abordado no capítulo seguinte.

Capítulo 4

Programa de análise de séries temporais

A análise de séries temporais e em geral qualquer operação relacionada com processamento de sinal requer um elevado número de operações e cálculos a efectuar sobre os dados em análise, sendo que tais dados podem também ser muito numerosos. É quase indispensável neste tipo de análises o uso de computadores e programas informáticos como uma ferramenta de suporte que permita a realização dos cálculos necessários.

Hoje em dia existem algumas ferramentas informáticas desenvolvidas especialmente para o estudo de séries temporais a partir de representações no espaço de fases e que permitem realizar alguns cálculos importantes como a dimensão de reconstrução óptima e expoentes de Lyapunov.

Algumas destas ferramentas com especial destaque são:

- **TSTOOL**: Desenvolvida no *Drittes Physikalisches Institut* [27], é uma ferramenta feita em linguagem *Matlab*, *C* e *C++*, pelo que para correr esta aplicação é necessário ter, a parte do programa *Matlab*, compiladores para os ficheiros feitos nas outras linguagens de programação. Esta aplicação consegue efectuar, a partir duma série temporal, a reconstrução no espaço de fases, cálculo de falsos vizinhos próximos para uma dada dimensão, cálculo da informação mútua média e ainda expoentes de Lyapunov, entre outras coisas. Outra característica importante desta aplicação é o facto de ser gratuita.
- **TISEAN**: Desenvolvida por Rainer Hegger, Holger Kantz e Thomas Schreiber [28], é uma ferramenta feita em linguagem *C* e *FORTRAN*. Em termos de potencialidades de cálculos é similar a *TSTOOL* sendo igualmente disponibilizada de maneira gratuita.
- **DATAPLORE**: Desenvolvida pela *Ixellence GmbH* [29] em colaboração com outras instituições, permite obter, entre outras coisas, representações no espaço de fases, testes estatísticos, análise espectral, modelos lineares para os dados, etc. A principal diferença entre esta aplicação e as anteriores é que esta ferramenta tem um custo associado, logo não é disponibilizada de maneira gratuita.

Apesar de que as aplicações anteriores e outras não mencionadas aqui resultam um bom suporte na análise de séries temporais, o desenvolvimento de uma aplicação informática própria que consiga assegurar os cálculos mais importantes será mais uma valia para conseguir entender melhor os problemas associados a análise de séries temporais. Ao anterior soma-se o facto de que a implementação de uma aplicação informática própria permite adaptar a ferramenta as nossas necessidades particulares e ainda permite a possibilidade de futura de expansão da aplicação.

Uma vez tomada a decisão de desenvolver uma aplicação informática própria, algumas características importantes a garantir nesta aplicação são:

- Dada uma série temporal, permitir a visualização do espaço de fases correspondente, para diferentes valores de desfasamento temporal.
- Permitir também a visualização do cronograma, autocorrelação e componentes em frequência da série.
- Possuir meios para o cálculo de parâmetros importantes no estudo de sistemas caóticos tais como: informação mútua média, dimensão de reconstrução óptima e expoentes de Lyapunov.
- Capacidade para a importação de dados a partir de ficheiros externos.

4.1 – Implementação

A linguagem de programação escolhida foi *Matlab*, visto ser uma linguagem de alto nível mundialmente difundida, onde o programador tem ao seu dispor um elevado número de bibliotecas com funções especialmente concebidas para o processamento de sinal. Em [30] é feita uma descrição pormenorizada sobre a linguagem de programação *Matlab*, sendo em particular abordado o tópico que trata do desenvolvimento de interfaces gráficas com recurso a este software.

O algoritmo usado no cálculo da dimensão de reconstrução ideal baseia-se no critério dos falsos vizinhos próximos, explicado no capítulo três. Já o algoritmo responsável pelo cálculo da informação mútua média foi desenvolvido por Alexandros Leontitsis. Enquanto ao maior expoente de Lyapunov, o algoritmo responsável por este cálculo foi desenvolvido por Max Logunov e baseia-se no método de Wolf descrito também no capítulo três. Estes dois últimos algoritmos são incluídos na *toolbox lab432*, disponível em [31].

Na implementação da aplicação foram ultrapassadas algumas dificuldades como à obtenção de tempos de execução aceitáveis em algoritmos altamente iterativos e a implementação de funções como *zoom* e *pan* num interface que integra sistemas de coordenadas de diferentes dimensões.

4.1.1 – Instalação e requisitos

Para o uso da aplicação é necessário ter instalado o programa *Matlab* e as *toolboxes*: *Symbolic Math* e *Signal Processing*. Todos os ficheiros constituintes da aplicação informática encontram-se numa pasta de nome “*PhaseSpaceTool*”. Para executar a aplicação basta seleccionar no *Matlab* como *Current Directory* a pasta “*PhaseSpaceTool*” e executar na linha de comandos o comando “*interface*”.

4.2 – Descrição

Nesta secção é feita uma descrição das principais características e funcionalidades da ferramenta desenvolvida para a análise de séries temporais. A seguinte ilustração mostra o aspecto geral do interface gráfico desenvolvido, sendo mais à frente feita uma descrição em detalhe dos seus elementos constituintes.

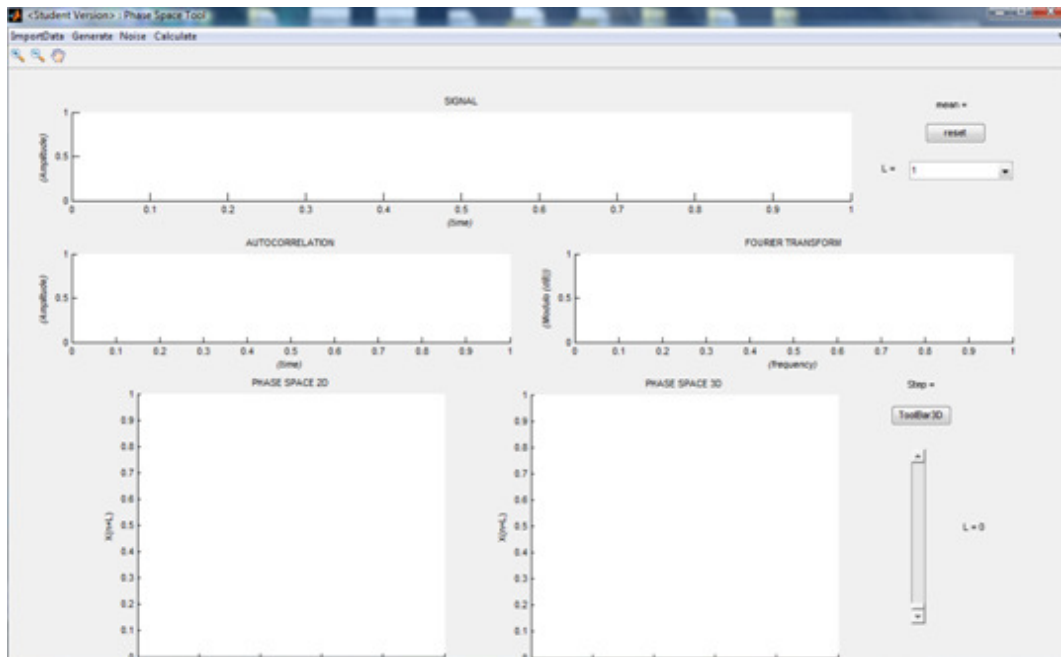


Figura 4.1 - Interface gráfico.

4.2.1 – Elementos do interface gráfico

A seguir é feita uma descrição detalhada dos elementos constituintes do interface gráfico:

1. Gráficos: O interface permite visualizar cinco gráficos correspondentes ao cronograma, função autocorrelação, espectro, espaço de fases em duas dimensões e espaço de fases em três dimensões.

2. Barra de funções *Zoom* e *Pan*: Localizada na parte superior esquerda, permite efectuar operações de *Zoom* e *Pan* sobre todos os eixos em duas dimensões do interface gráfico.
3. Botão “ToolBar3D”: Abre uma janela emergente com o espaço de fases em três dimensões. Nesta janela é possível efectuar as operações de *Zoom* e *Pan* sobre o espaço de fases em três dimensões.
4. Botão “reset”: Repõe em todos os eixos as definições originais, ou seja, aquelas que foram geradas ao importar ou gerar dados por última vez.
5. *Slider*: Localizado na parte inferior direita, permite alterar os valores para o desfasamento temporal utilizado no espaço de fases. Deslocando o valor do *slider* é possível ver, de maneira iterativa, como muda o espaço de fases em função do desfasamento temporal.
6. Popup menu “L”: Na literatura são descritos alguns valores de referência óptimos para o desfasamento temporal. Este menu, localizado na parte superior direita, permite atribuir de maneira automática um dos valores de referência ao desfasamento temporal. Contudo no caso do desfasamento temporal correspondente ao 1º mínimo da IMM, com o fim de ter maiores garantias no resultado obtido, é recomendado fazer uma verificação do resultado gráfico para a IMM (esta gráfico será explicado mais a frente) antes de fazer a atribuição deste valor.
7. Texto estático “Step”: Mostra o período de amostragem utilizado nas observações da série temporal. Por defeito este valor é 1.
8. Texto estático “mean”: Mostra o resultado para o valor médio do cronograma.

4.2.2 – Menus

Segue-se uma descrição dos menus disponibilizados no interface gráfico.

4.2.2.1 – ImportData

A partir deste menu é possível carregar uma série temporal a partir de um ficheiro externo de texto (.txt) ou Excel 2003 (.xls). Contudo é necessário ter algum cuidado na formatação dos dados no ficheiro fonte para a correcta execução do programa. Tais cuidados são:

1. Ficheiro de texto: Dados devem estar separados por espaço, tabulação, vírgula ou ponto e vírgula. Casas decimais devem ser marcadas com ponto. No ficheiro só devem aparecer os valores numéricos a analisar, não deve aparecer texto. Tal e como é ilustrado na seguinte figura, uma vez executado o comando *ImportData->text_file*, aparecerá a seguinte janela de selecção,

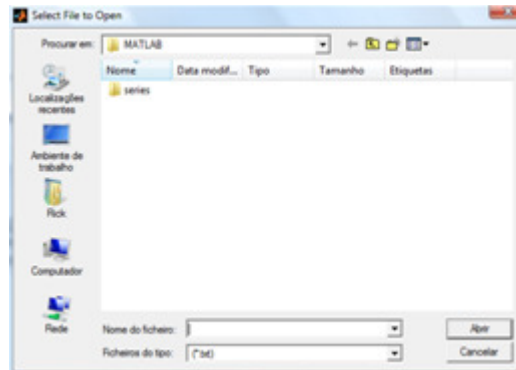


Figura 4.2 - Janela de selecção do menu *ImportData -> text_file*.

a qual permite abrir um ficheiro de texto a partir de qualquer localização no computador.

2. Ficheiro Excel 2003: Dados devem estar dispostos em colunas, não em linhas. Casas decimais devem estar separadas por vírgula. Neste caso é ainda possível seleccionar dados específicos da folha de cálculo, uma vez executado o comando *ImportData->Excel_file*, aparecerá uma janela de selecção similar à do caso anterior. Após a selecção do ficheiro Excel 2003 a importar, aparecerá a seguinte caixa de diálogo,

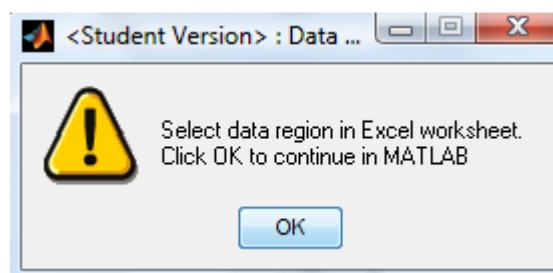


Figura 4.3 - Janela de diálogo do menu *ImportData -> excel_file*.

à qual deve ser confirmada logo após feita a selecção de dados na folha de cálculo aberta. Por último é importante referir que os dados seleccionados na folha de cálculo devem estar todos na mesma coluna.

4.2.2.2 – Generate

A partir deste menu é possível criar uma série de valores a partir de uma função de uma variável ou a partir de uma função recursiva. Também é possível gerar valores correspondentes a funções conhecidas e outras com especial importância no estudo dos sistemas caóticos. De forma mais detalhada, as funcionalidades oferecidas no menu *Generate* são:

1. Funções de uma variável: Ao executar o comando *Generate -> function_1_variable* aparecerá a seguinte janela para entrada de dados,

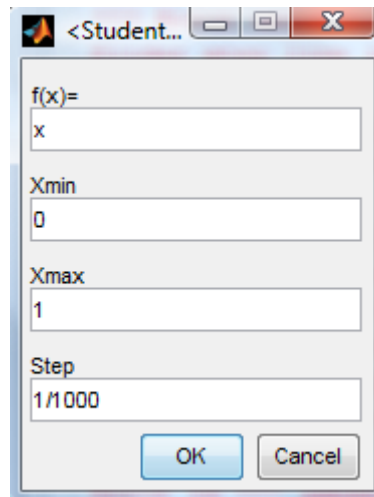


Figura 4.4 - Janela de entrada de dados do menu *Generate -> function_1_variable*.

na qual deve ser especificada uma função de uma variável (não necessariamente x), o intervalo onde será avaliada a função e o período de amostragem.

2. Funções Recursivas: As funções recursivas são de especial interesse no estudo dos sistemas dinâmicos. Ao executar o comando *Generate -> function_recursive*, aparecerá a seguinte janela para entrada de dados,

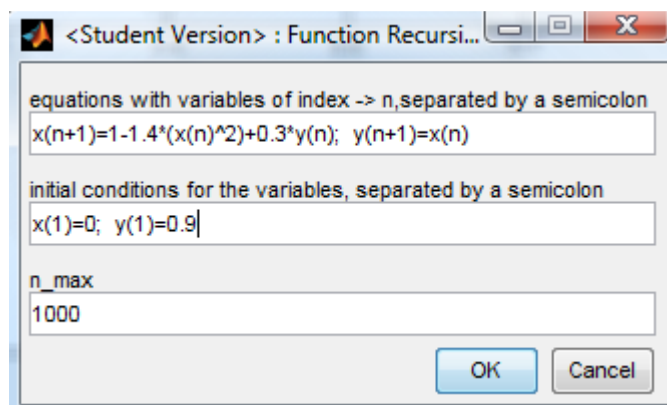


Figura 4.5 - Janela de entrada de dados para o comando *Generate -> function_recursive*.

na qual é possível inserir uma ou varias equações recursivas relacionadas entre si. Na primeira linha da janela devem ser introduzidas as equações recursivas separadas por ponto e vírgula. O índice utilizado para as variáveis nas equações deve ser n . Na segunda linha devem ser introduzidas as condições iniciais para as variáveis, em concordância com o número de variáveis e a ordem de cada uma. Cada condição inicial deve ir separada por ponto e virgula e a introdução das condições iniciais deve seguir o seguinte formato: $X(1)=condição1; X(2)=condição2;...Y(1)=condição3...$; finalmente na terceira linha da janela deve ser introduzido o número máximo de iterações a usar na avaliação das expressões recursivas.

3. Funções conhecidas: O resto das opções do menu *Generate* estão ligadas à geração de valores a partir de funções conhecidas. Em qualquer dos casos, ao executar o comando correspondente, aparecerá uma janela para entrada de dados similar as mencionadas anteriormente, onde o utilizador deve introduzir os parâmetros para as funções.

4.2.2.3 – Noise

A partir deste menu é possível adicionar ruído a os valores previamente seleccionados, de maneira aditiva ou multiplicativa. Em geral ao executar o comando *Noise -> additive* ou *Noise -> multiplicative*, aparecerá a seguinte janela para entrada de dados,

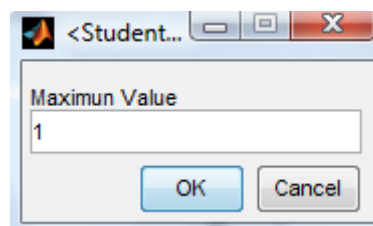


Figura 4.6 - Janela de entrada de dados para o comando *Noise -> additive/multiplicative*.

onde deve ser especificado o valor máximo para o factor aditivo ou multiplicativo do ruído.

4.2.2.4 – Calculate

A partir deste menu é possível realizar diversos cálculos de especial interesse na análise de comportamentos caóticos em séries temporais. De forma mais detalhada, as funcionalidades oferecidas no menu *Calculate* são:

1. Cálculo da informação mútua média: Tal e como foi referido no capítulo três, o primeiro mínimo da função IMM corresponde a um dos valores óptimos possíveis para o desfasamento temporal na reconstrução do espaço de fases. Ao executar o comando *Calculate -> Mútual_Information*, é feito o cálculo da IMM, para os valores previamente seleccionados, em função do desfasamento temporal. O resultado é

apresentado de forma gráfica. A modo de exemplo ilustra-se o resultado obtido no cálculo da MAI para o mapa logístico com os parâmetros usuais para 200 amostras,

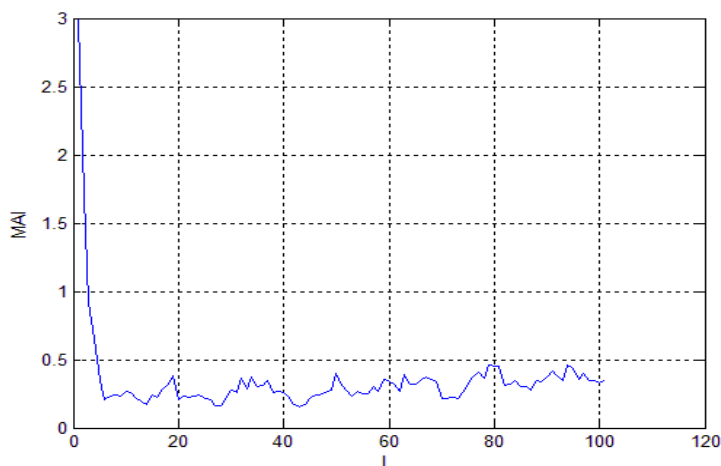


Figura 4.7 - Informação mútua média em função do desfasamento temporal L , para o mapa logístico com os parâmetros usuais para 200 amostras.

2. Cálculo da dimensão de reconstrução: Também referenciada no capítulo três, a escolha acertada da dimensão de reconstrução é um passo fundamental para uma correcta representação no espaço de fases. O algoritmo para o cálculo da dimensão de reconstrução óptima baseia-se no critério dos falsos vizinhos próximos. Ao executar o comando *Calculate -> Embedding_Dimension*, é feito o cálculo da dimensão de reconstrução óptima e é apresentada, de forma gráfica, a percentagem de falsos vizinhos próximos para cada dimensão analisada. A modo de exemplo ilustra-se o resultado obtido no cálculo da dimensão de reconstrução óptima para o mapa logístico com os parâmetros usuais para 200 amostras.

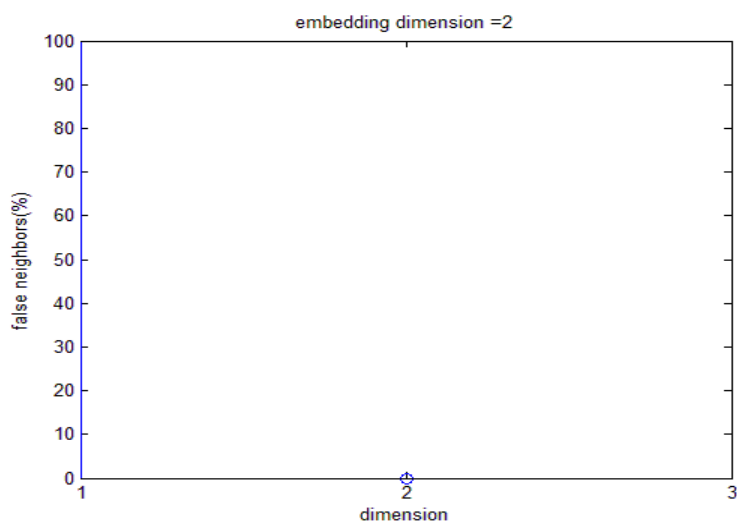


Figura 4.8 - Dimensão de reconstrução óptima para o mapa logístico com os parâmetros usuais para 200 amostras.

3. Cálculo do maior expoente de Lyapunov: No capítulo três é descrito o maior expoente de Lyapunov como um parâmetro que caracteriza a dinâmica de um atrator. Ele mede a taxa de divergência entre órbitas vizinhas dentro do atrator. A divergência entre órbitas vizinhas do espaço de fases é um indicador importante da presença de caos. No cálculo do maior expoente de Lyapunov existem dois parâmetros, relativamente arbitrários, sendo estes E_{min} e T_{min} . E_{min} corresponde ao limite de divergência entre órbitas vizinhas, ultrapassado este limite é feita a escolha dum novo ponto como vizinho próximo, tal como é explicado no capítulo três. T_{min} é o desfasamento temporal mínimo obrigatório, que deve ser respeitado no momento da escolha dum novo ponto como vizinho próximo. A função do parâmetro T_{min} é garantir que na escolha de um novo ponto como vizinho próximo, são apenas avaliados pontos de diferentes órbitas e não pontos consecutivos.

Ao executar o comando *Calculate* -> *Maximum_Lyapunov_Exponent* deve ser ainda especificada a dimensão para a qual pretende-se fazer o cálculo. Após a execução do comando aparece a seguinte janela,

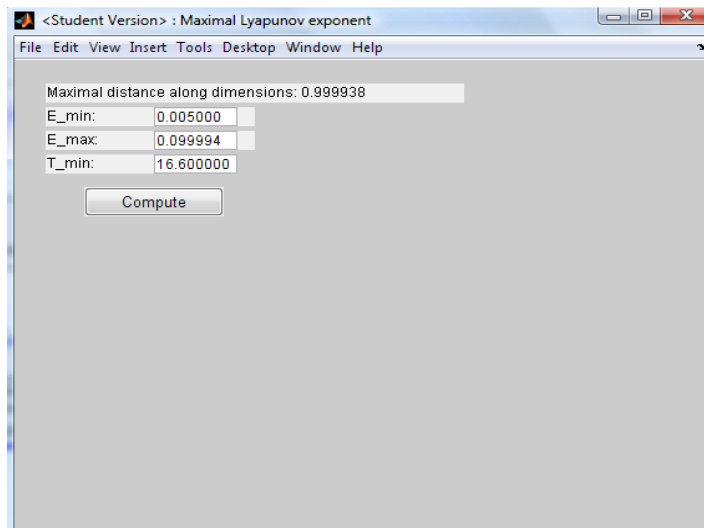


Figura 4.9 - Janela de entrada de dados para o comando *Calculate* -> *Maximum_Lyapunov_Exponent*.

onde devem ser introduzidos valores para os parâmetros E_{min} e T_{min} e logo a seguir deve ser pressionado o botão “*Compute*”. O resultado é apresentado na mesma janela, sendo possível ver a evolução temporal de tal valor. É importante referir que nem sempre é possível efectuar o cálculo do maior expoente de Lyapunov, para valores aceitáveis de E_{min} e T_{min} , visto este cálculo estar dependente do nível de divergência em órbitas vizinhas no espaço de fases seleccionado.

Finalmente é importante referir, em modo geral, que todos os dados introduzidos nas diversas janelas de entrada de dados devem estar em conformidade com a sintaxe do *Matlab*.

4.3 – Testes

Com a finalidade de validar resultados obtidos a partir do programa de análise desenvolvido, foram realizados testes para os diversos algoritmos, a partir de funções conhecidas e com especial relevância no estudo dos sistemas caóticos. Em [12; p417] e no anexo desta dissertação são mencionados valores de referência de alguns sistemas caóticos comuns. A comparação entre estes valores e os resultados obtidos a partir do programa desenvolvido constituem a base dos testes.

1. Mapa logístico, com os parâmetros usuais, para 1000 amostras.

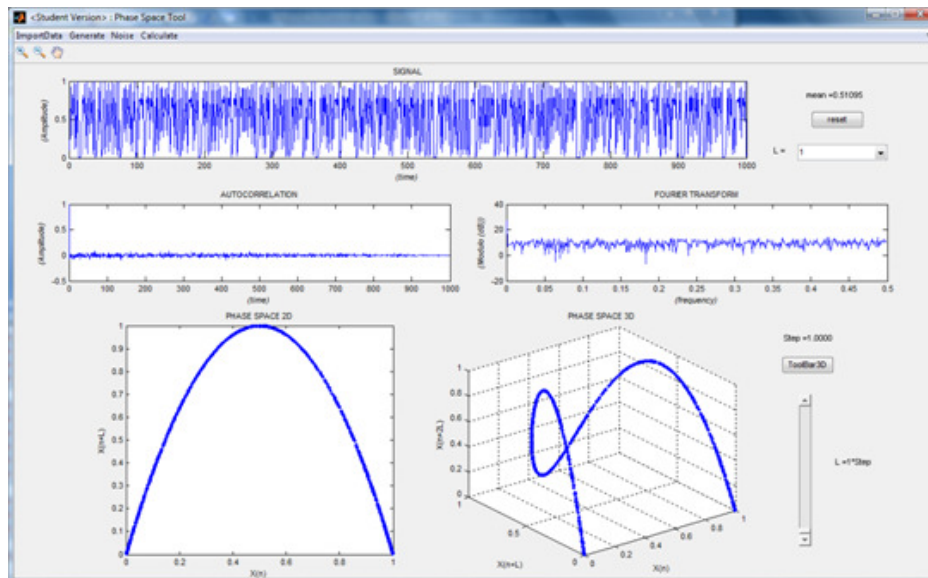


Figura 4.10 - Mapa logístico, com os parâmetros usuais, para 1000 amostras.

- Resultado para a dimensão de reconstrução óptima: 0% de falsos vizinhos próximos para dimensão de reconstrução 2.
- Resultado para o maior expoente de Lyapunov: 0.69, para $E_{\min} = 0.014$ e $T_{\min} = 33$.

2. Modelo populacional de Ricker's, com os parâmetros usuais, para 1000 amostras.

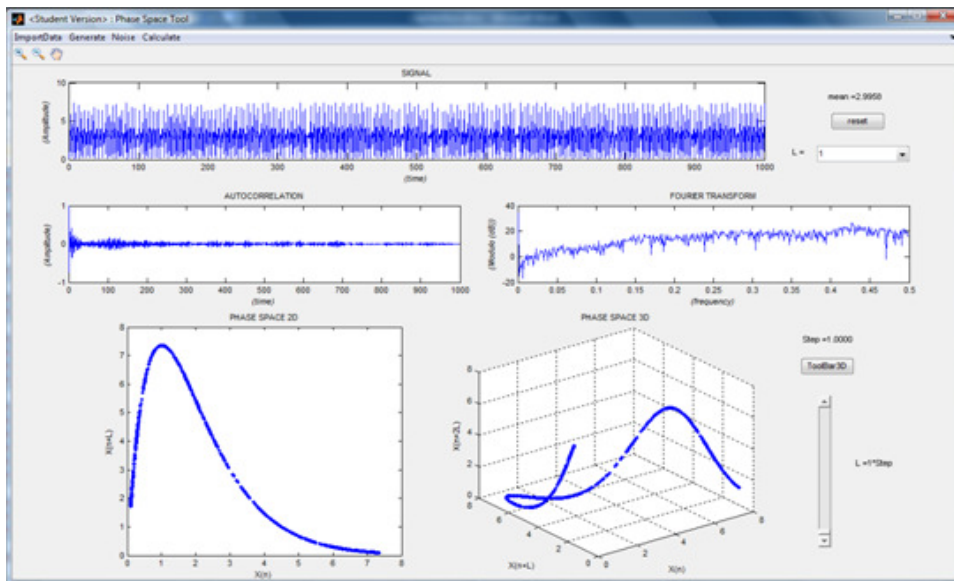


Figura 4.11 - Modelo populacional de Ricker's, com os parâmetros usuais, para 1000 amostras.

- Resultado para a dimensão de reconstrução óptima: 0% de falsos vizinhos próximos para dimensão de reconstrução 2.
- Resultado para o maior expoente de Lyapunov: 0.377, para $E_{\min}=0.176$ e $T_{\min}=34$.

3. Mapa de Hénon, com os parâmetros usuais, para 1000 amostras.

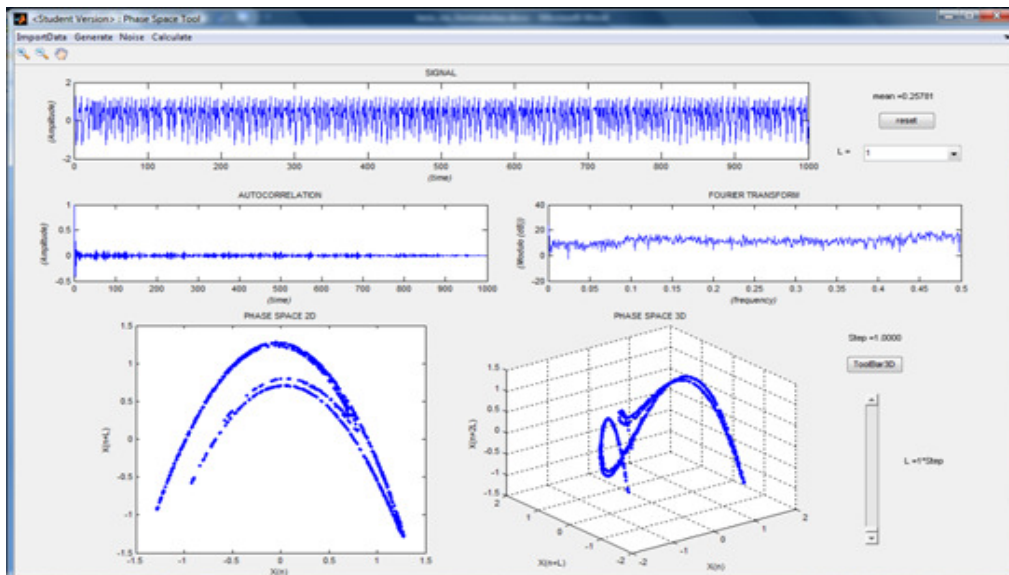


Figura 4.12 - Mapa de Hénon, com os parâmetros usuais, para 1000 amostras.

- Resultado para a dimensão de reconstrução óptima: 0% de falsos vizinhos próximos para dimensão de reconstrução 2.

- Resultado para o maior expoente de Lyapunov: 0.417, para $E_{\min}= 0.043$ e $T_{\min}= 33$.

4. Mapa de Lozi, com os parâmetros usuais, para 1000 amostras.

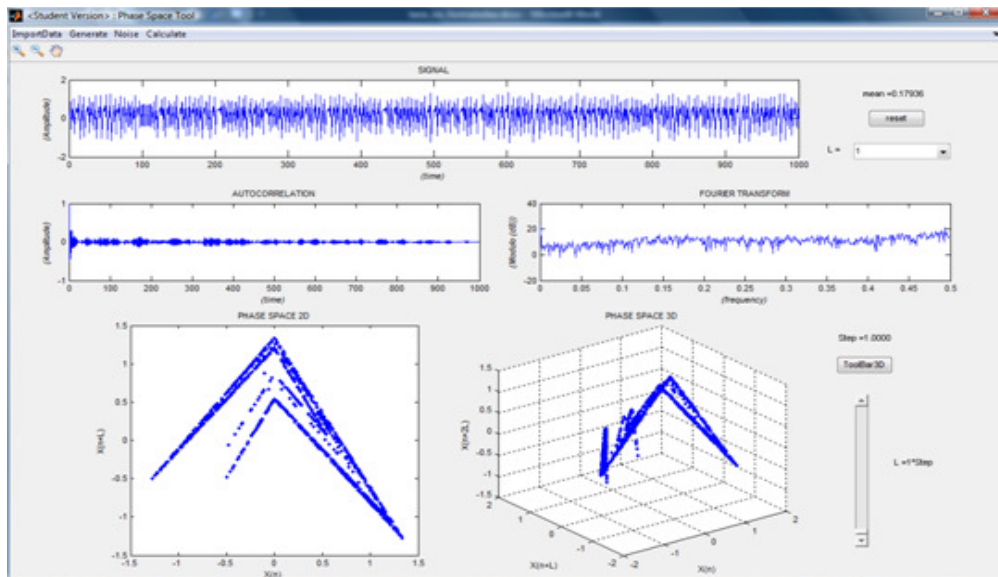


Figura 4.13 - Mapa de Lozi, com os parâmetros usuais, para 1000 amostras.

- Resultado para a dimensão de reconstrução óptima: 0% de falsos vizinhos próximos para dimensão de reconstrução 2.
- Resultado para o maior expoente de Lyapunov: 0.426, para $E_{\min}= 0.044$ e $T_{\min}= 33$.

5. Mapa logístico atrasado, com os parâmetros usuais, para 1000 amostras.

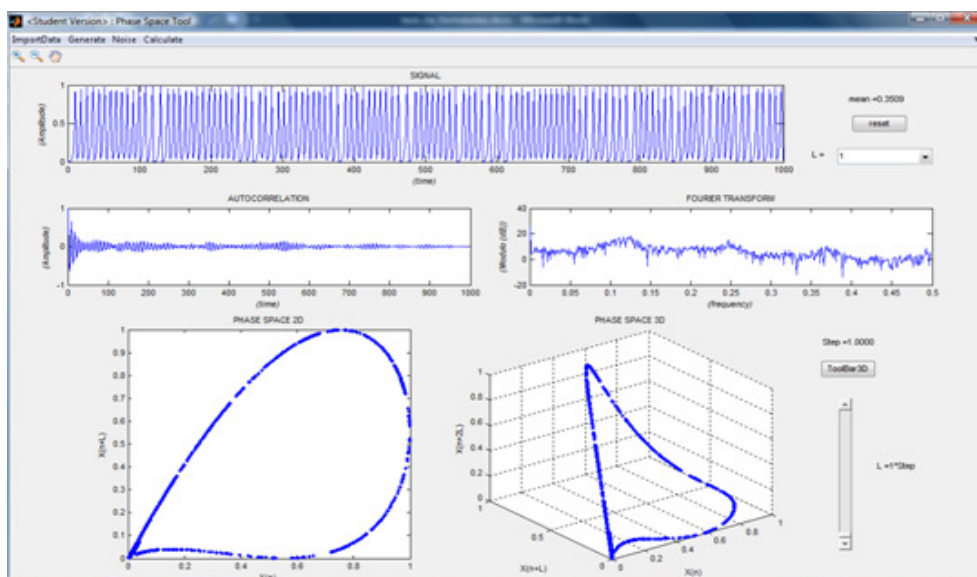


Figura 4.14 - Mapa logístico atrasado, com os parâmetros usuais, para 1000 amostras.

- Resultado para a dimensão de reconstrução óptima: 0% de falsos vizinhos próximos para dimensão de reconstrução 2.
- Resultado para o maior expoente de Lyapunov: 0.187, para $E_{\min}= 0.029$ e $T_{\min}= 33$.

Os resultados conseguidos para os sistemas testados, estão de acordo com os valores de referência obtidos.

No caso específico da informação mútua média, não foi possível obter valores de referência para os sistemas caóticos comuns testados. Com o objectivo de poder validar também este algoritmo, vamos a avaliar o resultado obtido na nossa aplicação para a informação mútua média no caso de uma série de valores correspondentes a um segmento de sinal sinusoidal.

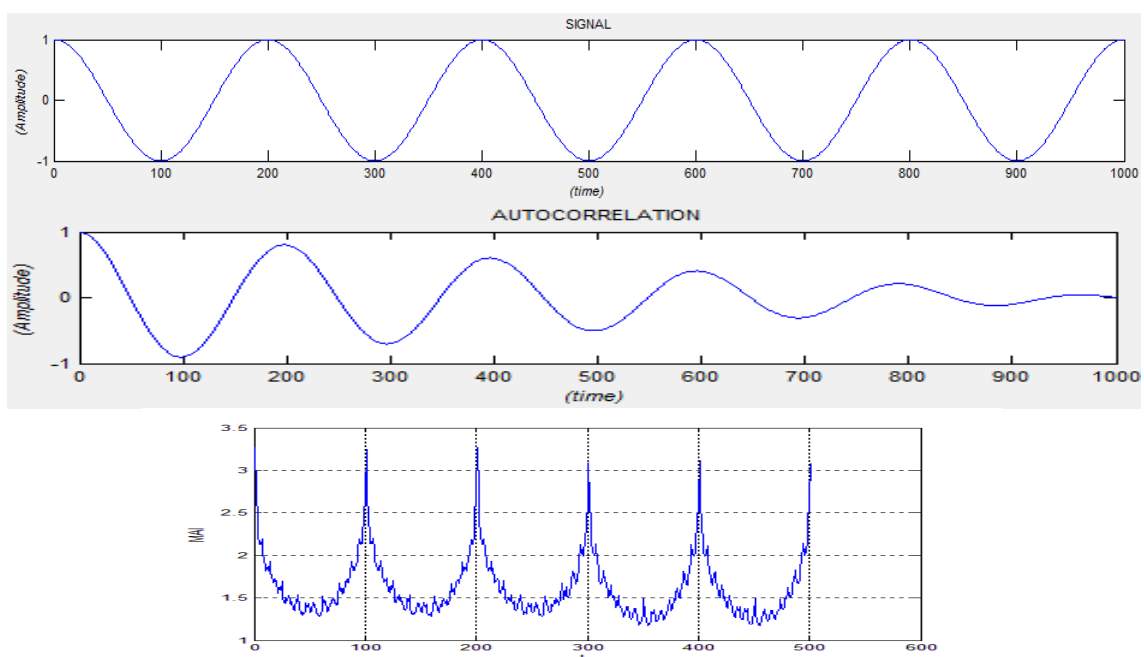


Figura 4.15 - Segmento de sinal sinusoidal com a sua correspondente autocorrelação e informação mútua média em função do desfasamento temporal L .

Tal e como foi descrito no capítulo três, a escolha de um desfasamento temporal igual ao primeiro mínimo da informação mútua média visa obter a máxima independência possível entre os valores da série temporal e um outro conjunto de valores formados pela mesma série temporal desfasada. Com base no anterior pode-se dizer que o resultado obtido está de acordo com o esperado, uma vez que os mínimos da função informação mútua média correspondem com os desfasamentos para os quais o sinal original e o sinal desfasado estão totalmente decorrelacionados, ou seja, estes mínimos correspondem com os zeros da função autocorrelação, garantindo-se assim a máxima independência sinal original e o sinal desfasado.

Uma vez concluídos satisfatoriamente os testes, estamos finalmente em condições de utilizar o interface gráfico como ferramenta de apoio no estudo e análise de séries temporais a traves de representações no espaço de fases.

Capítulo 5

Casos de estudo

No seguinte capítulo é feita a análise e caracterização de séries temporais no domínio das ciências naturais e sociais, as primeiras representativas de fenómenos pouco sujeitos à intervenção humana e as segundas de fenómenos em que a intervenção humana é determinante. Para este efeito faz-se, primeiramente, um estudo sobre a evolução temporal de caudais de rios e seguidamente o objectivo de estudo passa a ser a evolução temporal de algumas cotações de títulos na bolsa de valores nacional.

5.1 - Análise de caudais de rios

Em [20] são referenciadas algumas características na evolução temporal de caudais de rios que fazem deste tipo de dados interessantes do ponto de vista da análise não linear. O caudal de um rio é o resultado de uma medição de um sistema complexo em que intervém um grande número de variáveis tais como a precipitação, temperatura, humidade relativa do ar, existência de barragens, etc.

Foram analisados vários rios portugueses em diversos pontos de medição. De seguida são ilustradas, a partir de três casos representativos, as características gerais observadas para o caudal médio diário dos rios analisados. Os dados foram obtidos a partir do Sistema Nacional de Informação de Recursos Hídricos [23].

1. Caudal médio diário (m^3/seg) do rio Guadiana no posto de medição de Albernoa. Intervalo de observação: desde 31/10/1982 até 31/10/1997.

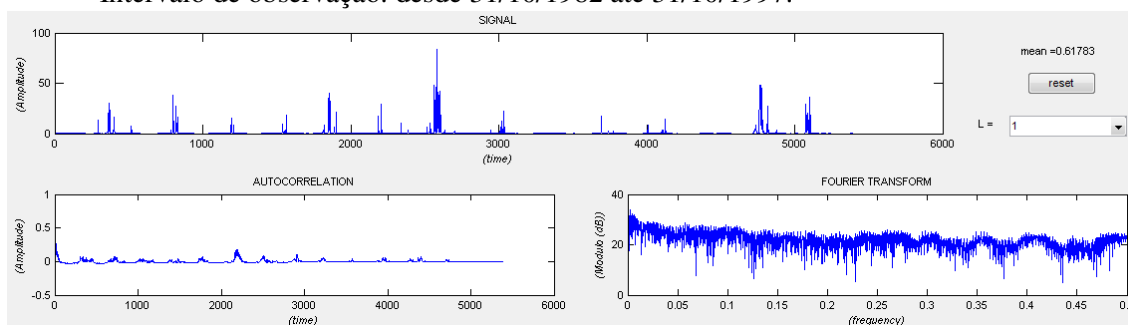


Figura 5.1 - Cronograma, autocorrelação e espectro do caudal médio diário do rio Guadiana no posto de medição de Albernoa, desde 31/10/1982 até 31/10/1997.

- Cronograma: Pode-se distinguir uma componente de variação sazonal, contudo há diferenças significativas de ano para ano. Não parece haver tendência. É de ressaltar ocorrência de subidas rápidas e valores muito afastados da média o qual é um indício de falta de linearidade.
- Autocorrelação: Coeficientes de autocorrelação em geral muito baixos indicam uma baixa correlação entre amostras, o qual permite concluir que estamos perante um processo dominado por alguma aleatoriedade.
- Espectro: Componentes de maior magnitude correspondem a baixas frequências. O facto anterior permite concluir que estamos perante um processo de variações muito lentas. Isto por sua vez sugere que tal vez seja necessário obter mais amostras ou diminuir a frequência de amostragem para contemplar um maior intervalo de tempo com o mesmo número de amostras. Por último, com auxílio da ferramenta zoom, pode-se verificar uma componente em frequência de grande amplitude para $f=1/365$ Hz, o que vem confirmar a presença duma componente sazonal.

Espaço de fases para o desfasamento temporal: 1 dia;

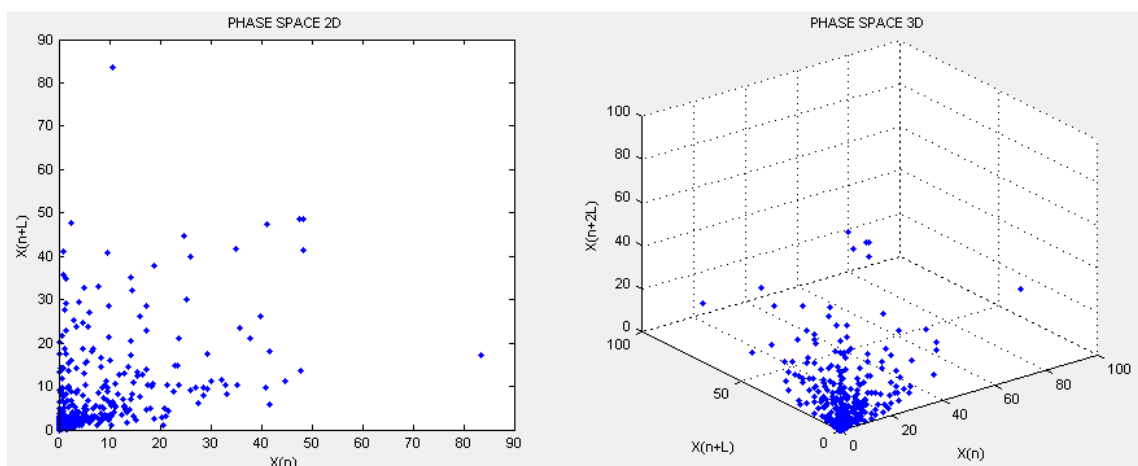


Figura 5.2 - Espaço de fases para o desfasamento temporal: 1 dia.

- Espaço de fases: Em geral bastante disperso, com o aparecimento de alguns valores discrepantes que correspondem a os valores de pico súbitos do cronograma. O resultado obtido para o cálculo da dimensão de reconstrução óptima é cinco ou superior, pelo que o espaço de fases em duas e três dimensões corresponde apenas a uma projecção do espaço de reconstrução ideal. Não foi possível efectuar o cálculo do maior expoente de Lyapunov para valores aceitáveis de E_{\min} e T_{\min} .

Espaço de fases para o desfasamento temporal: 1º zero da FAC= 58 dias;

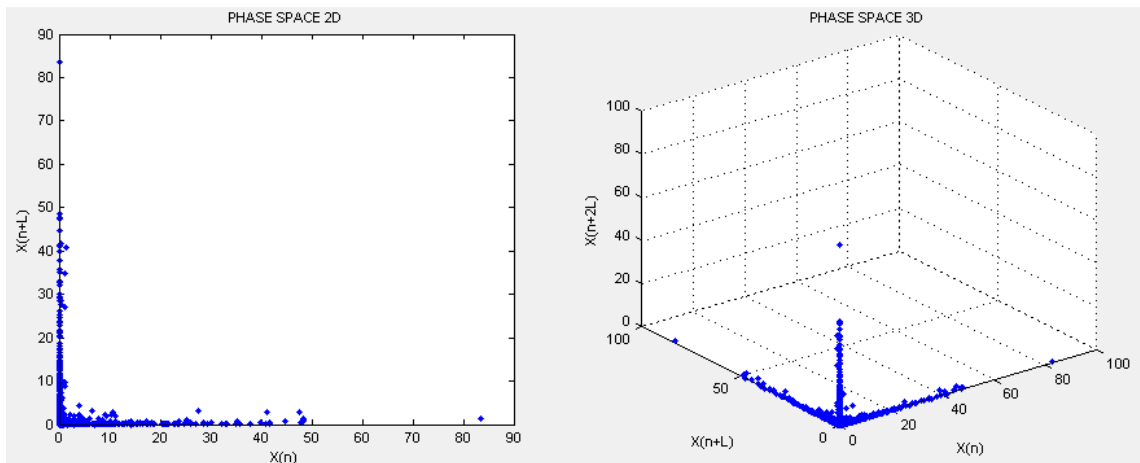


Figura 5.3 - Espaço de fases para o desfasamento temporal: 1º zero da FAC= 58 dias.

- Espaço de fases: Em geral concentrado numa região próxima a os eixos o qual indica variações elevadas entre valores da série separados por este desfasamento temporal. O resultado obtido para o cálculo da dimensão de reconstrução óptima é cinco ou superior, pelo que o espaço de fases em duas e três dimensões corresponde apenas a uma projecção do espaço de reconstrução ideal. O resultado do cálculo do maior expoente de Lyapunov foi de 0.396, para $E_{\min}=2$ e $T_{\min}= 171$. Um valor positivo para o maior expoente de Lyapunov confirma a presença de algum caos no sistema, mas sendo este um valor muito baixo, não há uma garantia sólida de caos no sistema.

Espaço de fases para o desfasamento temporal: 1º mínimo da IMM= 204 dias;

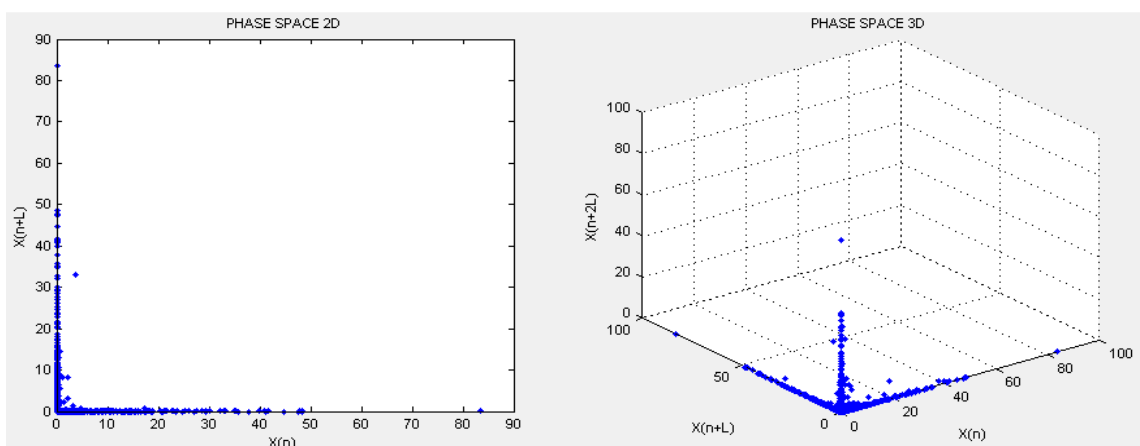


Figura 5.4 - Espaço de fases para o desfasamento temporal: 1º mínimo da IMM= 204 dias.

- Espaço de fases: Resultados semelhantes a os obtidos anteriormente para o caso do desfasamento temporal correspondente ao 1º zero da FAC.

2. Caudal médio diário (m^3/seg) do rio Lima no posto de medição de Forjões.
Intervalo de observação: desde 01/10/1982 até 30/09/1989.

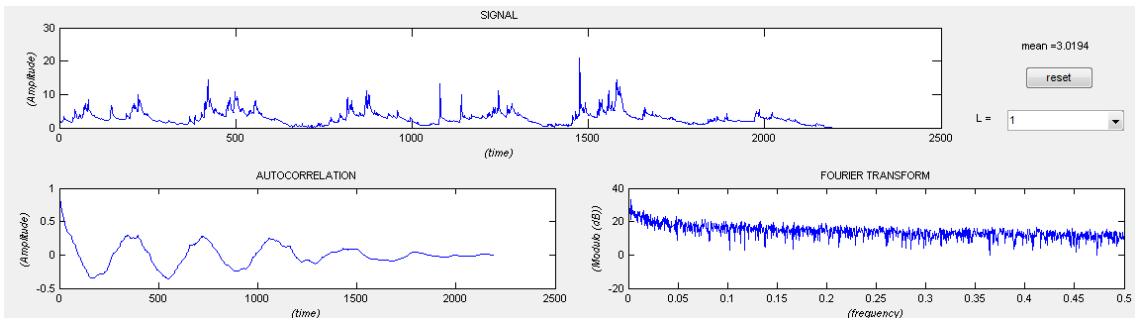


Figura 5.5 – Cronograma, autocorrelação e espectro do caudal médio diário (m^3/seg) do rio Lima no posto de medição de Forjões, desde 01/10/1982 até 30/09/1989.

- Cronograma: Pode-se distinguir uma componente de variação sazonal. Não parece haver tendência. É de ressaltar ocorrência de subidas rápidas e valores muito afastados da média o qual é um indício de falta de linearidade.
- Autocorrelação: É periódica, o que confirma a existência de periodicidade no cronograma. A existência de coeficientes de correlação elevados indica uma alta correlação entre os valores das amostras para certos desfasamentos.
- Espectro: Tal e como no caso anterior, as componentes de maior magnitude correspondem a baixas frequências o qual permite concluir que estamos perante um processo de variações muito lentas. Isto por sua vez sugere que tal vez seja necessário obter mais amostras ou diminuir a frequência de amostragem para contemplar um maior intervalo de tempo com o mesmo número de amostras. Por último, com auxílio da ferramenta zoom, pode-se verificar uma componente em frequência de grande amplitude para $f=1/365 \text{ Hz}$, o que vem confirmar a presença duma componente sazonal.

Espaço de fases para o desfasamento temporal: 1 dia;

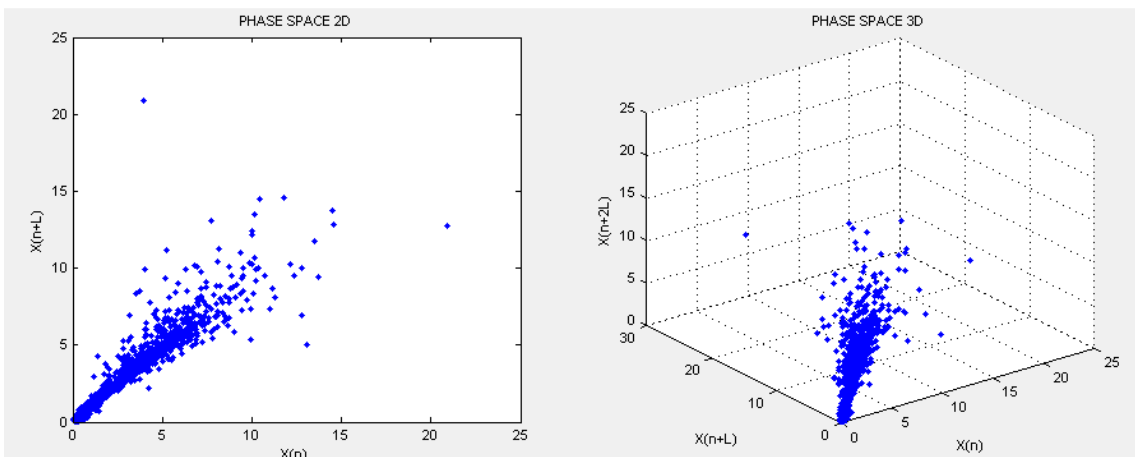


Figura 5.6 - Espaço de fases para o desfasamento temporal: 1 dia.

- Espaço de fases: Em geral muito concentrado na direcção da bissectriz, confirmando uma forte dependência entre estados consecutivos. O resultado obtido para o cálculo da dimensão de reconstrução óptima é quatro, visto que a percentagem de falsos vizinhos próximos para este valor é suficientemente baixa. O facto anterior implica que o espaço de fases em duas e três dimensões corresponde apenas a uma projecção do espaço de reconstrução ideal. Não foi possível efectuar o cálculo do maior expoente de Lyapunov para valores aceitáveis de E_{\min} e T_{\min} .

Espaço de fases para o desfasamento temporal: 1º zero da FAC= 96 dias;

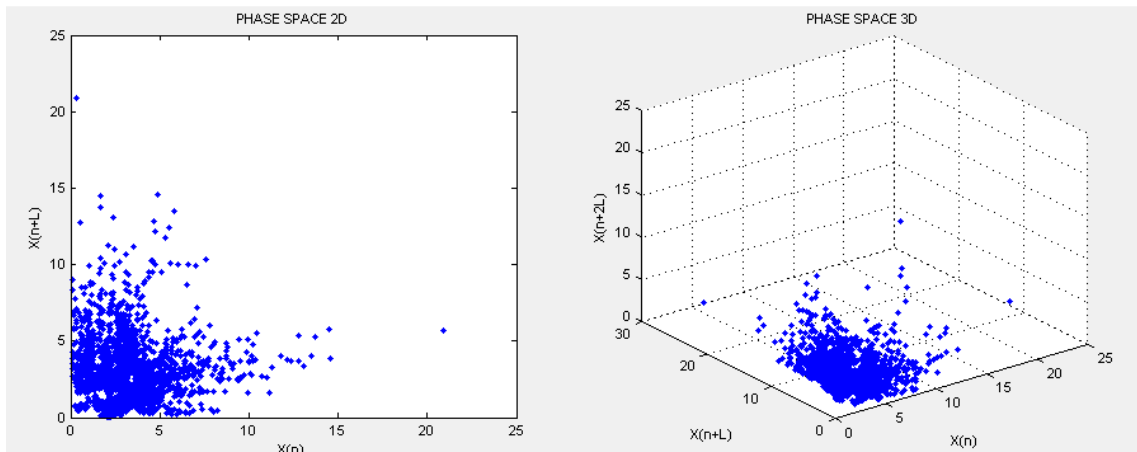


Figura 5.7 - Espaço de fases para o desfasamento temporal: 1º zero da FAC= 96 dias.

- Espaço de fases: Em geral muito disperso. O resultado obtido para o cálculo da dimensão de reconstrução óptima é quatro o qual implica que o espaço de fases em duas e três dimensões corresponde apenas a uma projecção do espaço de reconstrução ideal. Não foi possível efectuar o cálculo do maior expoente de Lyapunov para valores aceitáveis de E_{\min} e T_{\min} .

Espaço de fases para o desfasamento temporal: 1º mínimo da IMM= 83 dias;

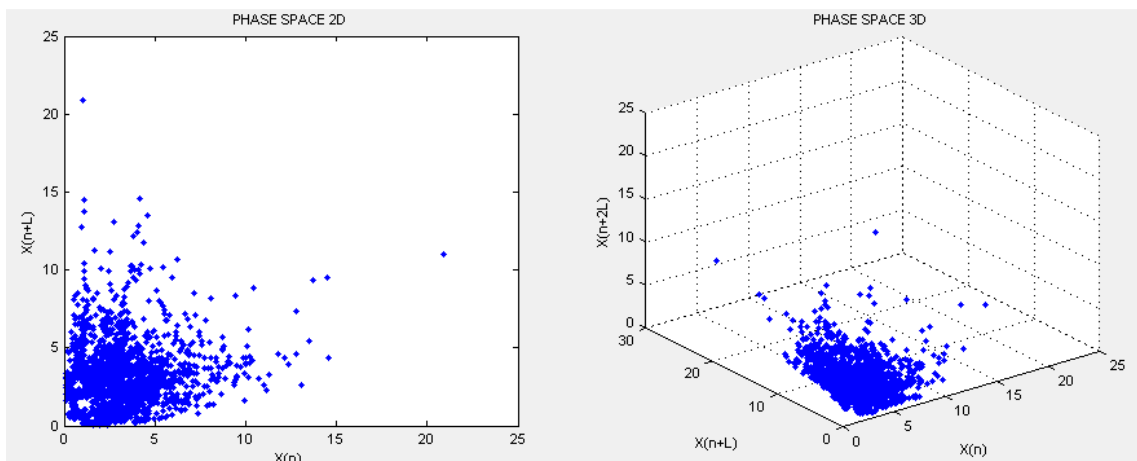


Figura 5.8 - Espaço de fases para o desfasamento temporal: 1º mínimo da IMM= 83 dias.

- Espaço de fases: Resultados semelhantes a os obtidos anteriormente para o caso do desfasamento temporal correspondente ao 1º zero da FAC.
3. Caudal médio diário (m^3/seg) do rio Mondego no posto de medição de Ponte Casével. Intervalo de observação: desde 01/10/1975 até 30/09/1990.

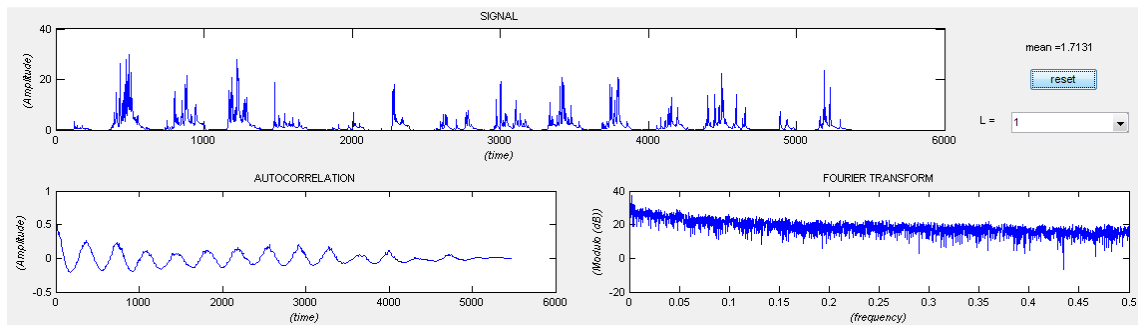


Figura 5.9 - Cronograma, autocorrelação e espectro do caudal médio diário (m^3/seg) do rio Mondego no posto de medição de Ponte Casével, desde 01/10/1975 até 30/09/1990.

- Cronograma: Mais uma vez pode-se distinguir uma componente de variação sazonal. Não parece haver tendência. É de ressaltar ocorrência de subidas rápidas e valores muito afastados da média o qual é um indício de falta de linearidade.
- Autocorrelação: Tal e como no caso anterior, é periódica, o que confirma a existência de periodicidade no cronograma. A existência de coeficientes de correlação elevados indica uma alta correlação entre os valores das amostras para certos desfasamentos.
- Espectro: Segue de encontro com os resultados anteriores onde as componentes de maior magnitude correspondem a baixas frequências permitido concluir que estamos perante um processo de variações muito lentas. Isto por sua vez sugere que tal vez seja necessário obter mais amostras ou diminuir a frequência de amostragem para contemplar um maior intervalo de tempo com o mesmo número de amostras. Por último, com auxílio da ferramenta zoom, pode-se verificar uma componente em frequência de grande amplitude para $f=1/365 \text{ Hz}$, o que vem confirmar a presença duma componente sazonal.

Espaço de fases para o desfasamento temporal: 1 dia;

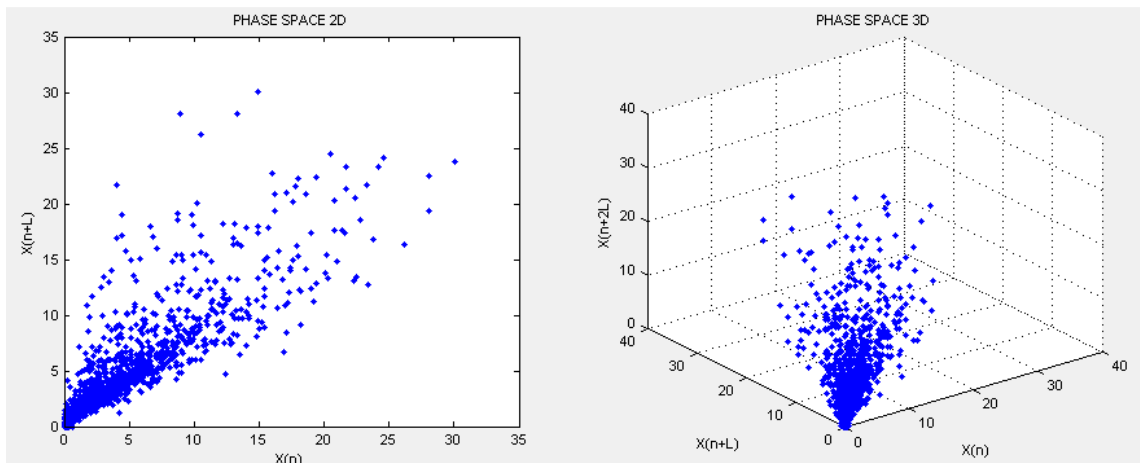


Figura 5.10 - Espaço de fases para o desfasamento temporal: 1 dia.

- Espaço de fases: Em geral disperso mas alinhado na direcção da bissectriz, confirmando uma dependência entre estados consecutivos. O resultado obtido para o cálculo da dimensão de reconstrução óptima é cinco ou superior, pelo que o espaço de fases em duas e três dimensões corresponde apenas a uma projecção do espaço de reconstrução ideal. Não foi possível efectuar o cálculo do maior expoente de Lyapunov para valores aceitáveis de E_{\min} e T_{\min} .

Espaço de fases para o desfasamento temporal: 1º zero da FAC= 106 dias;

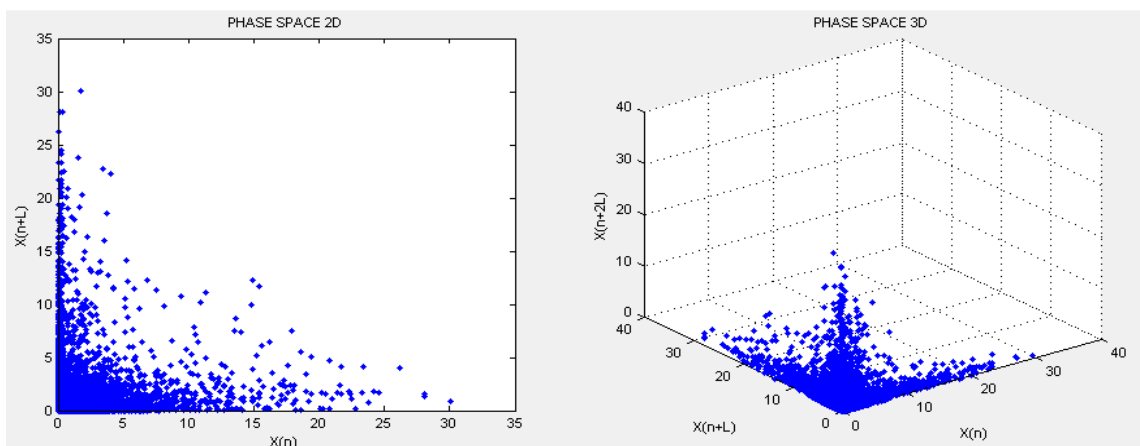


Figura 5.11 - Espaço de fases para o desfasamento temporal: 1º zero da FAC= 106 dias.

- Espaço de fases: Em geral muito disperso. O resultado obtido para o cálculo da dimensão de reconstrução óptima é cinco ou superior, pelo que o espaço de fases em duas e três dimensões corresponde apenas a uma projecção do espaço de reconstrução

ideal. Não foi possível efectuar o cálculo do maior expoente de Lyapunov para valores aceitáveis de E_{\min} e T_{\min} .

Espaço de fases para o desfasamento temporal: 1º mínimo da IMM= 130 dias;

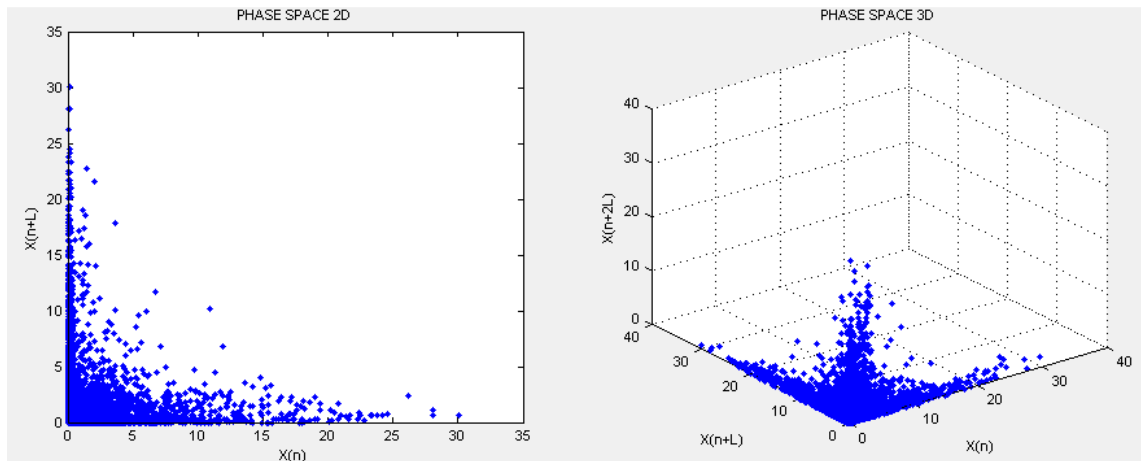


Figura 5.12 - Espaço de fases para o desfasamento temporal: 1º mínimo da IMM= 130 dias.

- Espaço de fases: Resultados semelhantes a os obtidos anteriormente para o caso do desfasamento temporal correspondente ao 1º zero da FAC.

O resultado obtido para as séries temporais correspondentes a caudais de rios mostra, em todos os casos, a presença de uma componente sazonal. Na análise espectral, componentes de maior amplitude a baixas frequências sugerem que estamos perante processos de variações lentas. Em geral espaços de fase correspondentes ao 1º zero da FAC e ao 1º mínimo da IMM mostraram características muito similares. Não foi possível encontrar padrões de caos determinístico, devido a representações no espaço de fases em geral com pontos muito dispersos e a impossibilidade, na maior parte dos casos, de efectuar o cálculo do maior expoente de Lyapunov.

5.2 - Análise de cotações de títulos na bolsa de valores nacional

Em [24] John Sterman provou, a partir da simulação de um sistema macroeconómico, que o processo de toma de decisões por parte de agentes económicos, pode ter comportamentos caóticos determinísticos. O mesmo autor explica que a descoberta de comportamentos caóticos determinísticos em fenómenos naturais há motivado a investigação de padrões similares em comportamentos humanos.

A cotação dum título na bolsa de valores [26] varia em função da procura e da oferta de cada acção, quanto maior for a procura, mais a cotação tenderá a subir. Por sua vez, a procura e a oferta são influenciadas pela opinião dos investidores sobre a situação da empresa (factores internos) e por diversos elementos de natureza social, política e económica que a empresa não

domina, mas que desempenham um papel preponderante (factores externos). Assim as cotações dos títulos são o resultado de um sistema complexo onde intervém um grande número de variáveis.

Foram analisados vários títulos da bolsa de valores portuguesa. De seguida ilustramos, a partir de três casos representativos, as características gerais observadas para os valores diários finais de acções em títulos do PSI20 (*Portuguese Stock Index*). Os dados analisados foram obtidos a partir do banco de dados do PSI20 [25].

1. Valor final diário (Euros) das acções do BCP no PSI20.
Intervalo de observação: desde 01/01/2003 até 11/05/2009.

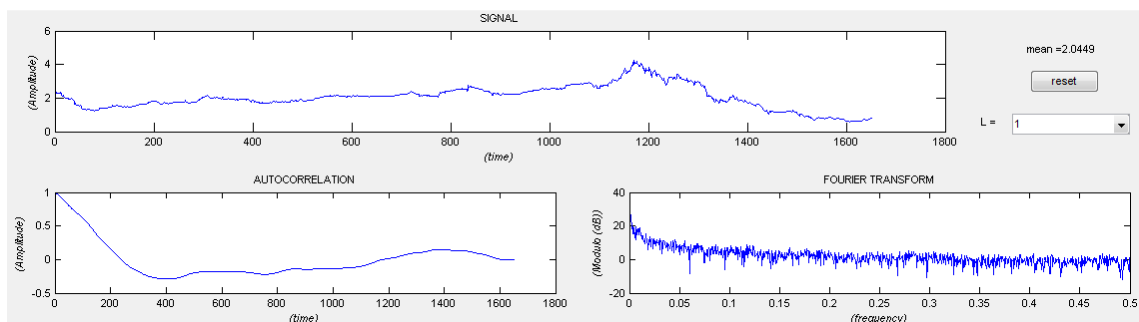


Figura 5.13 - Cronograma, autocorrelação e espectro para o valor final diário (Euros) das acções do BCP no PSI20, desde 01/01/2003 até 11/05/2009.

- Cronograma: Mostra dois segmentos de características diferentes divididos no dia 1170 onde ocorre o valor de máximo. O primeiro segmento correspondente a os 1170 primeiros dias é dominado por uma componente de ligeira componente de tendência crescente, sem apresentar indícios de oscilações sazonais ou cíclicas. O segundo segmento corresponde a os últimos 501 dias onde há uma forte componente de tendência decrescente e não há indícios de variações sazonais ou cíclicas.
- Autocorrelação: Não fornece grande informação, uma vez que o factor dominante é a tendência.
- Espectro: Componentes de maior magnitude correspondem a baixas frequências. O facto anterior permite concluir que estamos perante um processo de variações muito lentas. Isto por sua vez sugere que tal vez seja necessário obter mais amostras ou diminuir a frequência de amostragem para contemplar um maior intervalo de tempo com o mesmo número de amostras.

Espaço de fases para o desfasamento temporal: 1 dia;

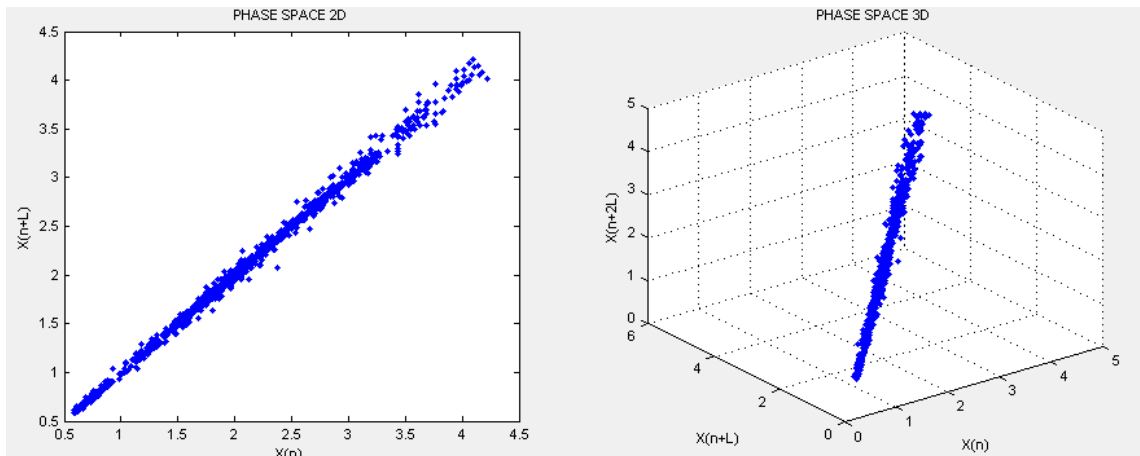


Figura 5.14 - Espaço de fases para o desfasamento temporal: 1 dia.

- Espaço de fases: Em geral fortemente alinhado na direcção da bissetriz, confirmando uma elevada dependência entre estados consecutivos. O resultado obtido para o cálculo da dimensão de reconstrução óptima é quatro, pelo que o espaço de fases em dois e três dimensões corresponde apenas a uma projecção do espaço de reconstrução ideal. O resultado do cálculo do maior expoente de Lyapunov foi de 0.0657, para $E_{\min}=0.248$ e $T_{\min}=18$. Um valor positivo para o maior expoente de Lyapunov confirma a presença de algum caos no sistema, mas sendo este um valor muito baixo, não há uma garantia sólida de caos no sistema.

Espaço de fases para o desfasamento temporal: 1º zero da FAC= 241 dias;

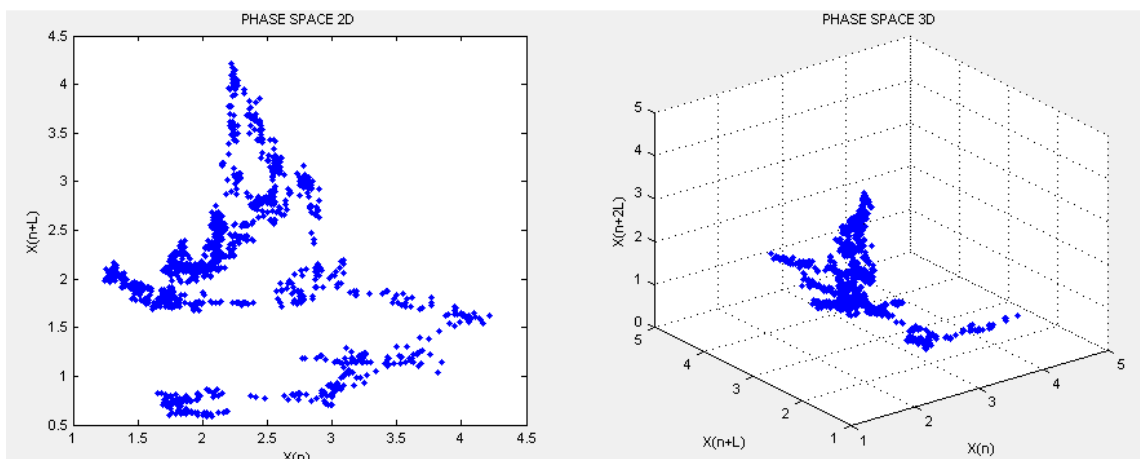


Figura 5.15 - Espaço de fases para o desfasamento temporal: 1º zero da FAC= 241 dias.

- Espaço de fases: Apesar de que os pontos no espaço de fases parecem ter alguma ordem geométrica, ainda não é possível reconhecer órbitas quase-periódicas, o qual representa

a principal característica dos atractores caóticos. O resultado obtido para o cálculo da dimensão de reconstrução óptima é quatro, pelo que o espaço de fases em dois e três dimensões corresponde apenas a uma projecção do espaço de reconstrução ideal. Não foi possível efectuar o cálculo do maior expoente de Lyapunov para valores aceitáveis de E_{\min} e T_{\min} .

Espaço de fases para o desfasamento temporal: 1º mínimo da IMM= 214 dias;

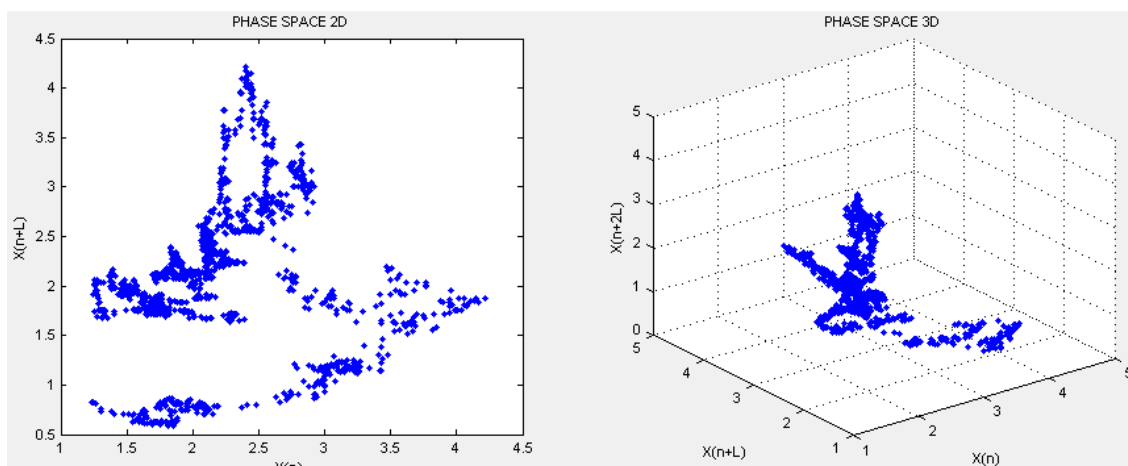


Figura 5.16 - Espaço de fases para o desfasamento temporal: 1º mínimo da IMM= 214 dias .

- Espaço de fases: Resultados semelhantes a os obtidos anteriormente para o caso do desfasamento temporal correspondente ao 1º zero da FAC.

2. Valor final diário (Euros) das acções da Brisa no PSI20.
Intervalo de observação: desde 01/01/2003 até 11/05/2009.

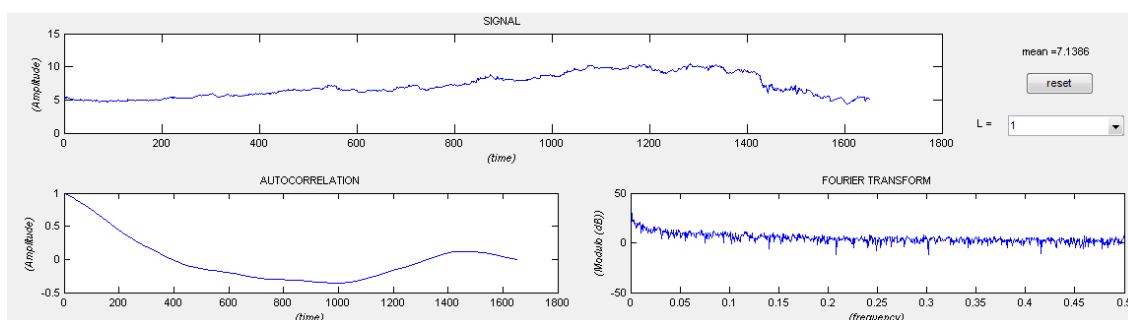


Figura 5.17 - Cronograma, autocorrelação e espectro do valor final diário (Euros) das acções da Brisa no PSI20, desde 01/01/2003 até 11/05/2009.

- Cronograma: Tal e como no caso anterior, mostra dois segmentos de características diferentes divididos desta vez no dia 1284 onde ocorre o valor de máximo. O primeiro segmento correspondente a os 1284 primeiros dias é dominado por uma componente de ligeira componente de tendência crescente, sem apresentar indícios de oscilações sazonais ou cíclicas. O segundo segmento corresponde a os últimos 387 dias onde há uma forte componente de tendência decrescente e não há indícios de variações sazonais ou cíclicas.
- Autocorrelação: Não fornece grande informação, uma vez que o factor dominante é a tendência.
- Espectro: Tal e como no caso anterior, componentes de maior magnitude correspondem a baixas frequências. Este facto permite concluir que estamos perante um processo de variações muito lentas. Isto por sua vez sugere que tal vez seja necessário obter mais amostras ou diminuir a frequência de amostragem para contemplar um maior intervalo de tempo com o mesmo número de amostras.

Espaço de fases para o desfasamento temporal: 1 dia;

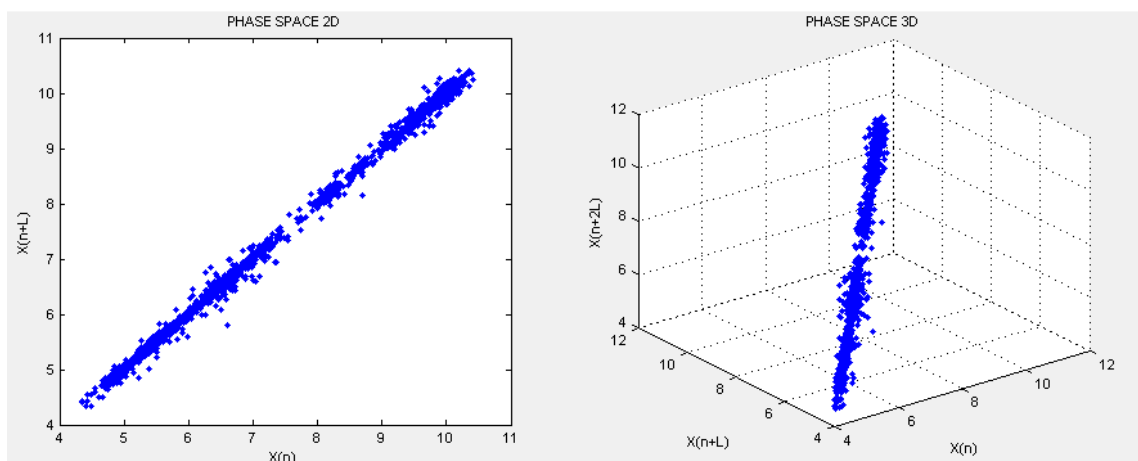


Figura 5.18 - Espaço de fases para o desfasamento temporal: 1 dia.

- Espaço de fases: Em geral fortemente alinhado na direcção da bissetriz, confirmando uma elevada dependência entre estados consecutivos. O resultado obtido para o cálculo da dimensão de reconstrução óptima é três, visto que a percentagem de falsos vizinhos próximos para este valor é suficientemente baixa. O resultado do cálculo do maior expoente de Lyapunov foi de 0.1354, para $E_{\min}=0.3035$ e $T_{\min}= 84$. Um valor positivo para o maior expoente de Lyapunov confirma a presença de algum caos no sistema, mas sendo este um valor muito baixo, não há uma garantia sólida de caos no sistema.

Espaço de fases para o desfasamento temporal: 1º zero da FAC= 396 dias;

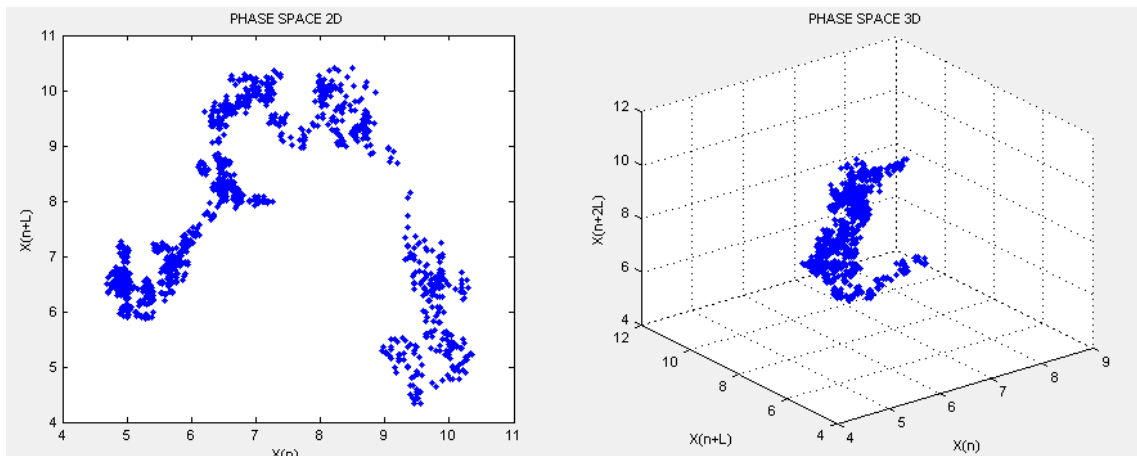


Figura 5.19 - Espaço de fases para o desfasamento temporal: 1º zero da FAC= 396 dias.

- Espaço de fases: Apesar de que os pontos no espaço de fases parecem ter alguma ordem geométrica, ainda não é possível reconhecer órbitas quase-periódicas, o qual representa a principal característica dos atratores caóticos. Não foi possível efectuar o cálculo da dimensão de imersão óptima e do maior expoente de Lyapunov devido a que o valor para o desfasamento temporal é demasiado elevado.

Espaço de fases para o desfasamento temporal: 1º mínimo da IMM= 104 dias;

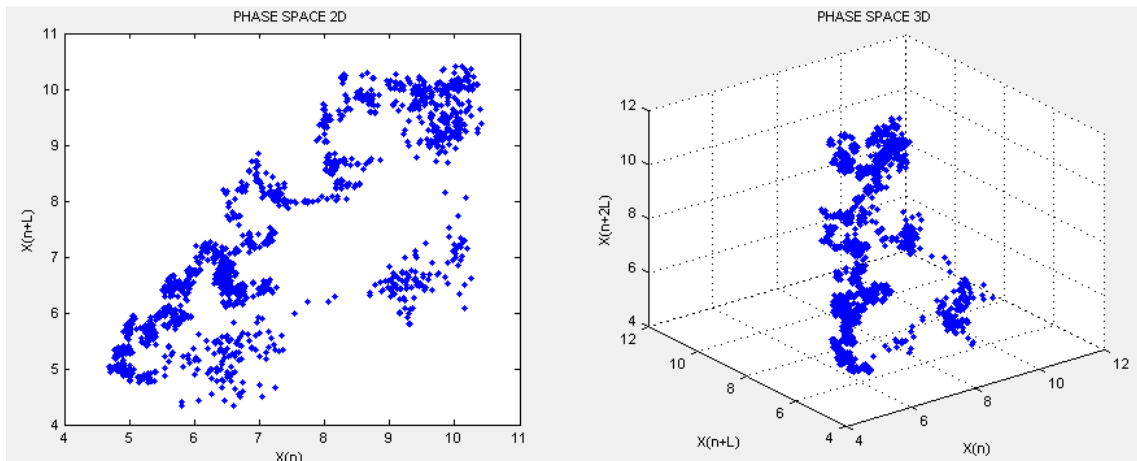


Figura 5.20 - Espaço de fases para o desfasamento temporal: 1º mínimo da IMM= 104 dias.

- Espaço de fases: Em geral disperso mas alinhado na direcção da bissectriz. O resultado obtido para o cálculo da dimensão de reconstrução óptima é três, visto que a percentagem de falsos vizinhos próximos para este valor é suficientemente baixa. Não

foi possível efectuar o cálculo do maior expoente de Lyapunov para valores aceitáveis de E_{\min} e T_{\min} .

3. Valor final diário (Euros) das acções da Sonae no PSI20.
Intervalo de observação: desde 30/06/2000 até 11/05/2009.

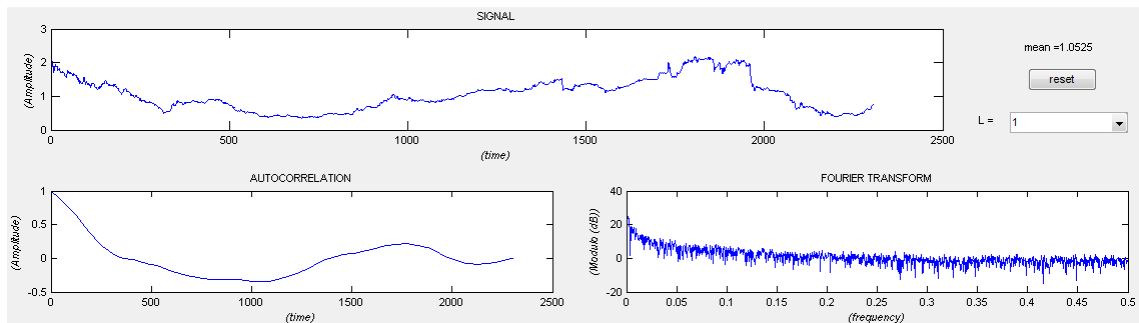


Figura 5.21 - Cronograma, autocorrelação e espectro do valor final diário (Euros) das acções da Sonae no PSI20, desde 30/06/2000 até 11/05/2009.

- Cronograma: O gráfico mostra um comportamento similar a os dois casos anteriores a partir do ano 2003, o que vem a confirmar que os valores das acções de diferentes títulos no PSI20 estão correlacionados entre si e seguem uma tendência global.
- Autocorrelação: Não fornece grande informação, uma vez que o factor dominante é a tendência.
- Espectro: Como era de esperar, componentes de maior magnitude correspondem a baixas frequências. Este facto permite concluir que estamos perante um processo de variações muito lentas. Isto por sua vez sugere que tal vez seja necessário obter mais amostras ou diminuir a frequência de amostragem para contemplar um maior intervalo de tempo com o mesmo número de amostras.

Espaço de fases para o desfasamento temporal: 1 dia;

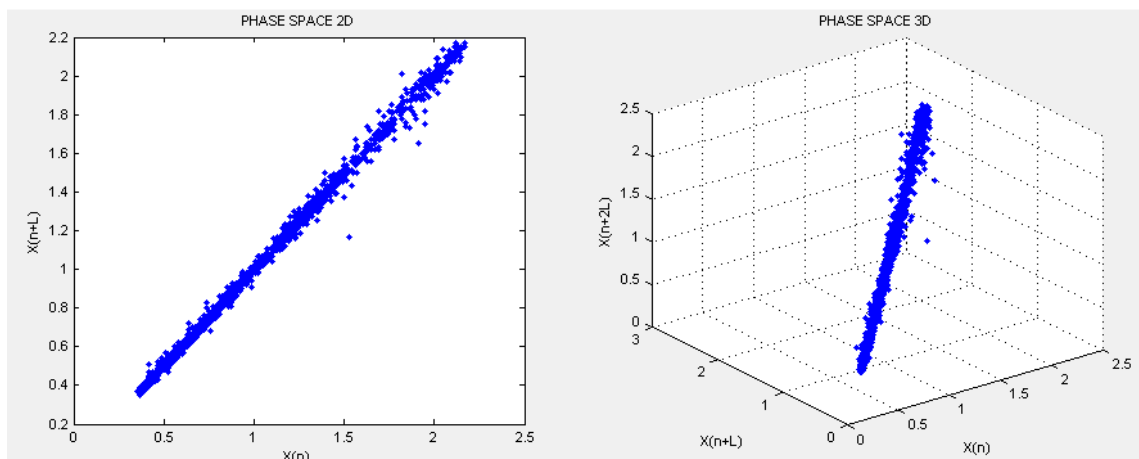


Figura 5.22 - Espaço de fases para o desfasamento temporal: 1 dia.

- Espaço de fases: Em geral fortemente alinhado na direcção da bissectriz, confirmando uma elevada dependência entre estados consecutivos. O resultado obtido para o cálculo da dimensão de reconstrução óptima é cinco ou superior, pelo que o espaço de fases em dois e três dimensões corresponde apenas a uma projecção do espaço de reconstrução ideal. O resultado do cálculo do maior expoente de Lyapunov foi de 0.075, para $E_{\min}=0.096$ e $T_{\min}= 26$. Um valor positivo para o maior expoente de Lyapunov confirma a presença de algum caos no sistema, mas sendo este um valor muito baixo, não há uma garantia sólida de caos no sistema.

Espaço de fases para o desfasamento temporal: 1º zero da FAC= 366 dias;

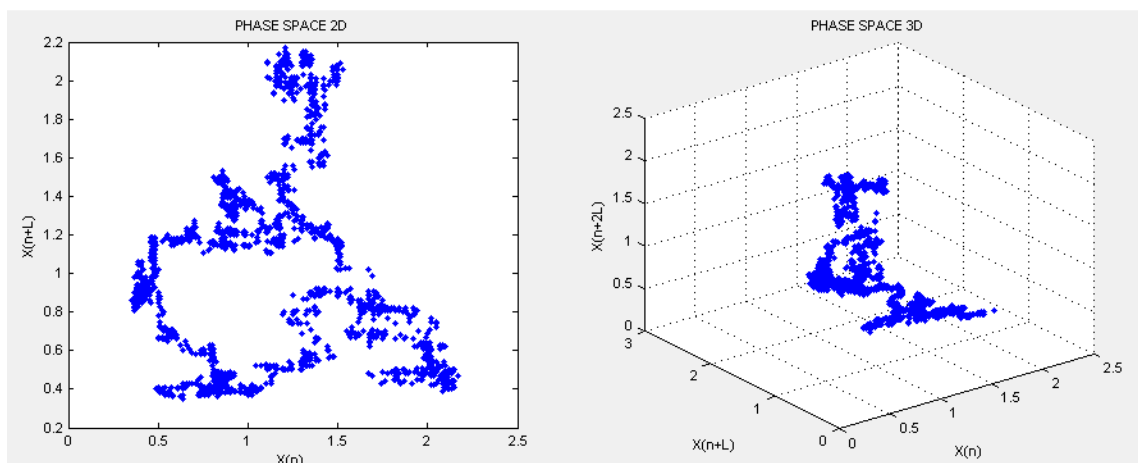


Figura 5.23 - Espaço de fases para o desfasamento temporal: 1º zero da FAC= 366 dias.

- Espaço de fases: Apesar de que os pontos no espaço de fases parecem ter alguma ordem geométrica, ainda não é possível reconhecer órbitas quase-periódicas, o qual representa a principal característica dos atractores caóticos. O resultado obtido para o cálculo da dimensão de reconstrução óptima é quatro, visto que a percentagem de falsos vizinhos próximos para este valor é suficientemente baixa. O facto anterior implica que o espaço de fases em dois e três dimensões corresponde apenas a uma projecção do espaço de reconstrução ideal. Não foi possível efectuar o cálculo do maior expoente de Lyapunov para valores aceitáveis de E_{\min} e T_{\min} .

Espaço de fases para o desfasamento temporal: 1º mínimo da IMM= 211 dias;

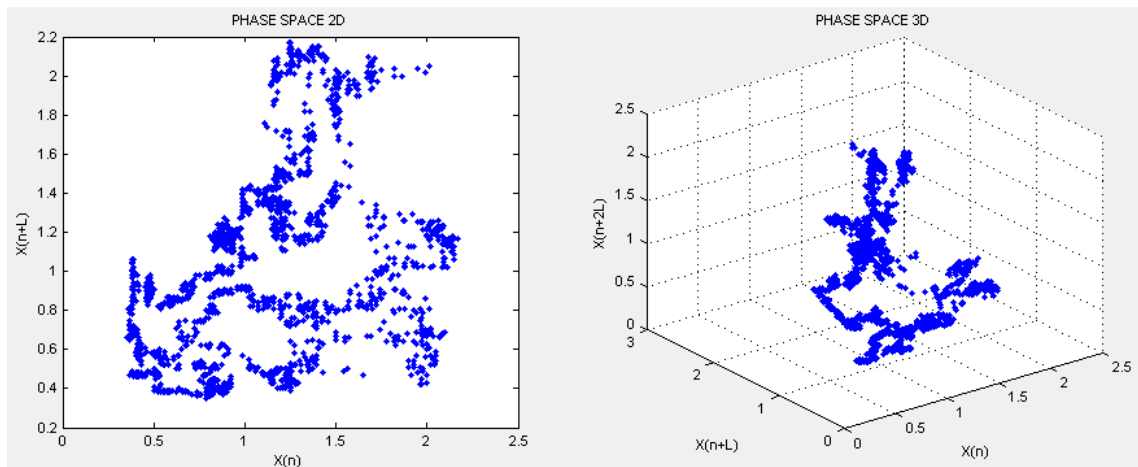


Figura 5.24 - Espaço de fases para o desfasamento temporal: 1º mínimo da IMM= 211 dias.

- Espaço de fases: Resultados semelhantes a os obtidos anteriormente para o caso do desfasamento temporal correspondente ao 1º zero da FAC.

No resultado obtido para séries temporais correspondentes a cotações de títulos na bolsa de valores foram evidenciadas características gerais comuns como uma forte dependência entre valores de dias consecutivos e o aparecimento de componentes de tendência. Na análise espectral, componentes de maior amplitude a baixas frequências sugerem que estamos perante processos de variações lentas. Nas representações no espaço de fases foram obtidas representações com alguma ordem geométrica e expoentes de Lyapunov positivos para desfasamentos temporais de um dia o qual fornece indícios de uma possível presença de caos determinístico em este tipo de fenómenos.

Capítulo 6

Conclusões

Neste trabalho foram abordados alguns dos principais modelos e métodos na análise de séries temporais, em especial aqueles que vão orientados a identificação de padrões de caos determinístico em séries temporais a traves de representações no espaço de fases.

Foi desenvolvido um interface gráfico capaz de servir como ferramenta de apoio na análise de séries temporais a partir de representações no espaço de fases. Para tal foram ultrapassadas algumas dificuldades de implementação como à obtenção de tempos de execução aceitáveis em algoritmos altamente iterativos e a implementação de funções como *zoom* e *pan* numa figura que integra sistemas de coordenadas de diferentes dimensões.

Na análise de séries temporais, algumas incertezas em torno da escolha do número de amostras ideal e do período de amostragem mais indicado, fazem difícil a validação dos resultados e das conclusões. Ao anterior soma-se o facto de nunca haver certeza absoluta de que o atractor reconstruído com base no teorema de Takens reproduza fielmente a dinâmica do sistema, visto que tal teorema não contempla factores que na prática devem ser considerados como erros de medição, ruído e amostras limitadas. Ainda cabe mencionar que esta reconstrução é baseada em métodos consensualmente aceites mais que não são infalíveis.

Por outro lado padrões de caos determinístico são difíceis de detectar devido ao grande número de evoluções diferentes que um sistema deste tipo pode tomar para condições iniciais muito próximas.

Em particular no caso de séries temporais de caudais de rios a característica mais significativa foi a presença de uma componente sazonal. Contudo não foi possível neste caso encontrar padrões de caos determinístico, devido a representações no espaço de fases em geral com pontos muito dispersos e a impossibilidade, na maior parte dos casos, de efectuar o cálculo do maior expoente de Lyapunov.

Já na análise de séries temporais de cotações de títulos na bolsa de valores foram evidenciadas características gerais como uma forte dependência entre valores consecutivos e o aparecimento de componentes de tendência. Nas representações no espaço de fases foram obtidas representações com alguma ordem geométrica e expoentes de Lyapunov positivos para desfaseamentos temporais de um dia o qual fornece indícios de uma possível presença de caos determinístico em este tipo de fenómenos.

Com base nestes resultados animadores mas pouco contundentes, pensamos que a análise de séries temporais no campo da economia a partir de representações no espaço de fases é um campo de investigação que poderá ser promissório. A ideia anterior vê-se reforçada pelo facto de ainda haver poucos estudos deste género na actualidade.

Anexo

Sistemas Caóticos Comuns

A seguir são descritos alguns valores de referência para os sistemas caóticos comuns referenciados ao longo do trabalho [12; p417].

1. Mapa logístico

$$x(n + 1) = A \cdot x(n) \cdot (1 - x(n)).$$

Parâmetros usuais: $A= 4$, $x(0)= 0.1$.

Dimensão de reconstrução: 2.

Maior Expoente de Lyapunov: 0.6931.

2. Modelo populacional de Ricker's

$$x(n + 1) = A \cdot x(n) \cdot e^{-x(n)}.$$

Parâmetros usuais: $A= 20$, $x(0)= 0.1$.

Dimensão de reconstrução: 2.

Maior Expoente de Lyapunov: 0.3848.

3. Mapa de Hénon

$$x(n + 1) = 1 - a \cdot x^2(n) + b \cdot y(n),$$

$$y(n + 1) = x(n).$$

Parâmetros usuais: $a= 1.4$, $b= 0.3$, $x(0)= 0$, $y(0)= 0.9$.

Dimensão de reconstrução: 2.

Maior Expoente de Lyapunov: 0.4192.

4. Mapa de Lozi

$$x(n+1) = 1 - a \cdot |x(n)| + b \cdot y(n),$$

$$y(n+1) = x(n).$$

Parâmetros usuais: $a=1.7$, $b=0.5$, $x(0)=-0.1$, $y(0)=0.1$.

Dimensão de reconstrução: 2.

Maior Expoente de Lyapunov: 0.47.

5. Mapa logístico atrasado

$$x(n+1) = A \cdot x(n) \cdot (1 - y(n)),$$

$$y(n+1) = x(n).$$

Parâmetros usuais: $a=2.27$, $x(0)=0.001$, $y(0)=0.001$.

Dimensão de reconstrução: 2.

Maior Expoente de Lyapunov: 0.1831.

Refêrencias

- [1] Ricardo S. Ehlers, "Análise de Séries Temporais", artigo escolar. Disponível em: <http://www.icmc.usp.br/~ehlers/notas/stemp.pdf>. Acesso 20/04/2009.
- [2] Bento J.F. Murteira, Daniel A. Müller, K. Feridun Turkman, *Análise de Sucessões Cronológicas*, McGraw-Hill de Portugal, 1993.
- [3] http://data.giss.nasa.gov/precip_cru/graphs/. Acesso 21/04/2009.
- [4] Christopher Chatfield, *Time-séries Forecasting*, Chapman & Hall/CRC, 2001. Via google books.
- [5] Sir Maurice Kendall, J. Keith Ord, *Time Séries*, Edward Arnold, 1990.
- [6] Christopher Chatfield, *The analysis of time séries: theory and practice*, Chapman & Hall, 1975.
- [7] Alan V Oppenheim, Alan S. Willsky, S Hamid Nawab, *Señales y Sistemas*, Pearson Educación, Mexico, 1998.
- [8] Henry D. I. Abarbanel, *Analysis of Observed Chaotic data*, Springer, 1996. Via google books.
- [9] Kyle Mandli, "Reconstruction of Phase Space: Embedding Dimensions", artigo escolar. Disponível em http://www.amath.washington.edu/~mandli/school/amath_575_talk.pdf. Acesso 07/05/2009.
- [10] Ricardo Vicente Solé, Susanna C. Manrubia, *Orden y Caos en sistemas complejos. Fundamentos*, Universitat Politècnica de Catalunya, 2001, Espanha. Via google books.
- [11] Adriana Campanharo, Elbert Macau, Fernando M Ramos, "Detectando a presença de caos em uma série temporal". Disponível em http://www.sbmac.org.br/eventos/cnmac/cd_xxviii_cnmac/resumos%20estendidos/andriana_campanharo_ST21.pdf. Acesso 08/05/2009. Acesso 08/05/2009.
- [12] Julien Clinton Sprott, *Chaos and Time-Séries Analysis*, Oxford University Press, 2003.

- [13] Edward Ott, *Chaos in Dynamical Systems*, Cambridge University Press, 1993.
- [14] Robert C. Hilborn, *Chaos and Nonlinear Dynamics. An Introduction For Scientists and engineer*, Oxford University Press, 1994.
- [15] Marcelo Amorim Savi, *Dinâmica Não-linear e Caos*, e-papers. Via google books.
- [16] Marcelo Catanoe Gandur, “Comportamento Dinâmico Complexo em Despelamento de Fitas Adesivas”, tese de doutoramento, Universidade Estadual de Campinas, 2001. Disponível em <http://www.biq.iqm.unicamp.br/arquivos/teses/ficha48910.htm>. Acesso 12/05/2009.
- [17] <http://docentes.if.ufg.br/~pferrari/projetocaos/minicurso/aula1.htm>. Acesso 13/05/2009.
- [18] Andrew M. Fraser, Harry L. Swinney, “Independent coordinates for strange attractors from mutual information”, *Physical Review A* 33-2, 1134-1140. Ano 1986. Disponível em <http://chaos.utexas.edu/manuscripts/1064949034.pdf>. Acesso 18/05/2009.
- [19] Andreia Raquel Simoni, “Análise de Séries Temporais Aeroelásticas Experimentais Não Lineares”, tese de doutoramento, Escola de Engenharia de São Carlos da Universidade de São Paulo, 2008. Disponível em <http://www.teses.usp.br/teses/disponiveis/18/18148/tde-12062008-113731/>. Acesso 19/05/2009.
- [20] Rui Jorge de Almeida Soares Gonçalves, “Sistemas naturais complexos: que tendências? O estocástico e o determinístico”, tese de doutoramento, Faculdade de Engenharia da Universidade do Porto, 2005.
- [21] Sílvio Abrantes Moreira, “Medidas de Informação”, artigo escolar. Disponível em: <http://paginas.fe.up.pt/~sam/TI/apontamentos/apontamentos.html>. Acesso 26/05/2009.
- [22] Howell Tong, *Non-linear Time Séries, A Dynamical System Approach*, Clarendon Press Oxford, 1990.
- [23] <http://snirh.pt>. Acesso 05/06/2009.
- [24] John Sterman, “Deterministic Chaos in an Experimental Economic System”, Massachusetts Institute of Technology. 1998. Disponível em <http://openlibrary.org/b/OL17942017M/Deterministic-chaos-in-an-experimental-economic-system>. Acesso 05/06/2009.
- [25] <http://www.psi20.net>. Acesso 05/06/2009.
- [26] <http://www.protestepoupanca.pt/map/show/122826/src/34970.htm>. Acesso 08/09/2009.
- [27] <http://www.dpi.physik.uni-goettingen.de/en/home.html>. Acesso 17/06/2009.

[28] <http://www.mpipks-dresden.mpg.de/~tisean/>. Acesso 17/06/2009.

[29] <http://www.ixellence.com/dataplore/>. Acesso 17/06/2009.

[30] Vagner Morais, *Cláudio Silva Vieira, Matlab 7 & 6 – Curso Completo*, FCA-Editora de Informática, 2006.

[31] <http://www.mathworks.com/Matlabcentral/fileexchange/16994>. Acesso 20/06/2009.

[32] http://www.educ.fc.ul.pt/icm/icm2000/icm24/teoria_do_caos.htm. Acesso 26/06/2009.