



UNIVERSIDADE DE LISBOA  
FACULDADE DE CIÊNCIAS  
DEPARTAMENTO DE ENGENHARIA GEOGRÁFICA, GEOFÍSICA E ENERGIA

## Previsão Sazonal de Índices de seca com modelos de Markov

João Miguel Barbosa Botelho de Sousa

Mestrado em Ciências Geofísicas  
Especialização em Meteorologia

2009



UNIVERSIDADE DE LISBOA  
FACULDADE DE CIÊNCIAS  
DEPARTAMENTO DE ENGENHARIA GEOGRÁFICA, GEOFÍSICA E ENERGIA

## Previsão Sazonal de Índices de seca com modelos de Markov

João Miguel Barbosa Botelho de Sousa

Mestrado em Ciências Geofísicas  
Especialização em Meteorologia

Trabalho de Projecto orientado pelo Prof. Doutor Carlos Alberto Leitão Pires

2009

Dissertação inserida no âmbito do projecto PTDC/AGR-AAM/71649/2006 – “Gestão do risco em secas: Identificação, monitorização, caracterização, predição e mitigação”, financiado pela Fundação para a Ciência e Tecnologia – Portugal.

# Agradecimentos

Gostaria em primeiro lugar de expressar um agradecimento especial ao meu Professor e Orientador Carlos Alberto Leitão Pires pela oportunidade que me deu de trabalhar neste projecto, pela sua orientação e pela sua disponibilidade na resolução de problemas e dúvidas que surgiram ao longo do mesmo.

Não poderia deixar de agradecer ao coordenador deste projecto denominado "*Gestão do risco em secas*" – PTDC/AGR-AAM/71649/2006 – CC4881, o Professor Luís Santos Pereira do Instituto Superior de Agronomia da Universidade Técnica de Lisboa, pela oportunidade de trabalhar neste projecto e ainda, em poder participar no 6º Simpósio de Meteorologia e Geofísica da APMG, Encontro Luso-Espanhol de Meteorologia, em Março de 2009 na Caparica e ainda na Assembleia Geral da EGU – *European Geosciences Union*, em Abril de 2009 na cidade de Viena, Áustria.

Queria fazer também um agradecimento à Professora Ana Paulo do Instituto Superior de Agronomia da Universidade Técnica de Lisboa por toda a ajuda que me disponibilizou e pelas séries de dados que me facultou sempre que precisei, nomeadamente as séries de dados do SPI para a região do Alentejo, obtidas através dos polígonos de Thiessen.

Agradeço também de uma forma geral a todos os meus colegas de laboratório pelo ambiente de trabalho tranquilo que é fundamental ao bom desempenho de trabalho.

Agradeço por fim à minha família que sempre me apoiou, aconselhou e motivou na execução deste trabalho e que sem eles teria sido muito difícil a sua execução.

## Resumo

A seca é um fenómeno recorrente que tem afectado as civilizações ao longo da história, com ocorrência aleatória e severidade incerta, resultante de valores anómalos da precipitação. Afecta os habitats naturais, ecossistemas e muitos sectores económicos e sociais desde a agricultura, transportes, fornecimento urbano de água e complexos industriais modernos.

Devido às características aleatórias da seca é adoptada uma abordagem estocástica ao estudo e caracterização deste fenómeno. Neste trabalho é utilizado um modelo estocástico de tipo cadeias de Markov ao índice SPI (Standard Precipitation Index), para identificar, caracterizar e monitorizar o fenómeno em estudo. Este modelo é igualmente útil como ferramenta de previsão e predição de secas. Dada a conhecida influência dos padrões de circulação atmosférica de larga escala na precipitação sazonal, são então, acrescentadas algumas alterações no modelo no sentido de aumentar a capacidade preditiva deste, incluindo para isso informação proveniente destes padrões através de índices de circulação global atmosférica (CGA) a eles associados.

A calibração destes dois modelos de Markov, que diferem entre si pela inclusão ou não dos índices de CGA (ou preditores meteorológicos de larga escala), é feita através de séries temporais de precipitação e de séries temporais de índices de CGA, para o local em estudo, o Alentejo. Estes dois modelos são então validados para instantes de tempo futuros e a sua qualidade preditiva aferida.

## **Abstract**

Drought is a recurrent phenomenon that has affected civilizations throughout history, with random occurrence and uncertain severity, due to outliers in rainfall. It affects the natural habitats, ecosystems and many social and economic sectors from agriculture, transport, urban water supply and modern industrial plants.

Due to the random characteristics of the drought, one has considered a stochastic approach to study and characterize this phenomenon. This paper used a stochastic model of type Markov chains to the SPI index (Standard Precipitation Index) in order to identify, characterize and monitor the phenomenon under study. This model is also useful as a tool for forecasting and prediction of the drought. Given the known influence of atmospheric circulation patterns in large-scale seasonal rainfall, we have added some changes to the Markov chain model in order to increase its predictive ability, including information from these patterns by using indexes of global atmospheric circulation (GAC) associated with them.

The calibration of these two Markov models, which differ by the inclusion or not of the GAC indexes (or meteorological predictors of large-scale), is done through time series of rainfall and time series of GAC indexes, to the site of study, the Alentejo region. These two models are then validated for time instants in the future and its predictive quality is measured.

## Lista de Figuras e Tabelas

FIGURA 2.1 SITUAÇÃO SINÓPTICA (PRESSÃO AO NÍVEL MÉDIO DO MAR) .....	3
FIGURA 2.2 SECAS EM PORTUGAL CONTINENTAL DE MAIOR EXTENSÃO DESDE 1940 (FONTE: IM). .....	4
FIGURA 2.3 ESQUEMA DA SEQUÊNCIA TEMPORAL DOS DIVERSOS TIPOS DE SECA (ADAPTADO DE WWW.DROUGHT.UNL.EDU, 2002). .....	6
FIGURA 2.4 TRANSFORMAÇÃO DOS VALORES DA PRECIPITAÇÃO, AJUSTADOS À FUNÇÃO GAMA, NUMA DISTRIBUIÇÃO NORMAL DE VALOR MÉDIO ZERO E DESVIO PADRÃO UNITÁRIO (DUTRA, 2006). .....	10
FIGURA 2.5 FASE POSITIVA DA NAO .....	12
FIGURA 2.6 FASE NEGATIVA DA NAO .....	13
FIGURA 2.7 MAPA DA CORRELAÇÃO ENTRE OS ÍNDICES NAO E SPI NO INVERNO (DJF) NA REGIÃO EURO-ATLÂNTICA (PIRES E PERDIGÃO, 2007).....	14
FIGURA 2.8 (A) E (B) FASES POSITIVA E NEGATIVA DA OSCILAÇÃO DO ÁRTICO, RESPECTIVAMENTE. (WALLACE J., UNIVERSIDADE DE WASHINGTON).....	15
FIGURA 2.9 ANOMALIAS DA PRESSÃO (MB) NOS CASOS DE EL NIÑO (À ESQUERDA) E LA NIÑA (À DIREITA). (CPC, 2005) .....	16
FIGURA 3.1 CAMINHOS POSSÍVEIS DE UMA REDE DE MARKOV COM 4 ESTADOS DE UM ESTADO $I=2$ , NUM INSTANTE $T$ , PARA UM ESTADO $J=1$ , NUM INSTANTE $T+2$ (PAULO <i>ET AL.</i> , 2005).....	25
FIGURA 4.1 ESTAÇÕES UDOMÉTRICAS DO ALENTEJO, ESCOLHIDAS PARA MODELAÇÃO DE MARKOV ATRAVÉS DE SÉRIES TEMPORAIS DE SPI (PAULO E PEREIRA, 2005).....	33
FIGURA 4.2 MALHA DE DADOS DE PRECIPITAÇÃO ACUMULADA MENSAL DAS REANÁLISES NCEP (2.5° LAT X 2.5° LONG). O PONTO A VERMELHO É O PONTO ESCOLHIDO.....	34
TABELA 2.1 REGIMES DE ESCASSEZ DE ÁGUA (PEREIRA <i>ET AL.</i> , 2002).....	5
TABELA 2.2 CLASSIFICAÇÃO DA INTENSIDADE DA SECA DE ACORDO COM O SPI E FREQUÊNCIA DE OCORRÊNCIA DE CADA CLASSE.....	9
TABELA 3.1 TABELA DE CONTINGÊNCIA PARA K CLASSES DE SEVERIDADE DE SECA (PREVISÃO/OBSERVAÇÃO) .....	31
TABELA 4.1 LOCALIZAÇÃO DAS 7 ESTAÇÕES UDOMÉTRICAS NO ALENTEJO.....	33
TABELA 4.2 CLASSIFICAÇÃO UTILIZADA NA DISCRETIZAÇÃO DAS SÉRIES TEMPORAIS DE SPI. ....	35
TABELA 5.1 CORRELAÇÃO ENTRE O ÍNDICE NAO E O ÍNDICE SPI-1 A VÁRIOS MESES DE ANTECEDÊNCIA PARA A REGIÃO DO ALENTEJO. ....	38
TABELA 5.2 CORRELAÇÃO ENTRE O ÍNDICE AO E O ÍNDICE SPI-1 A VÁRIOS MESES DE ANTECEDÊNCIA PARA A REGIÃO DO ALENTEJO. ....	38
TABELA 5.3 CORRELAÇÃO ENTRE O ÍNDICE SOI E O ÍNDICE SPI-1 A VÁRIOS MESES DE ANTECEDÊNCIA PARA A REGIÃO DO ALENTEJO. ....	38
TABELA 5.4 CORRELAÇÃO ENTRE O ÍNDICE NAO E O ÍNDICE SPI-1 (LAG=0), PARA TODOS OS MESES E PARA OS MESES DE OUTONO-INVERNO (OUT-MAR) E PRIMAVERA-VERÃO (ABR- SET). ....	39
TABELA 5.5 CORRELAÇÃO ENTRE O ÍNDICE NAO E O ÍNDICE SPI-12 A VÁRIOS MESES DE ANTECEDÊNCIA PARA A REGIÃO DO ALENTEJO. ....	39
TABELA 5.6 CORRELAÇÃO ENTRE O ÍNDICE AO E O ÍNDICE SPI-12 A VÁRIOS MESES DE ANTECEDÊNCIA PARA A REGIÃO DO ALENTEJO. ....	39
TABELA 5.7 CORRELAÇÃO ENTRE O ÍNDICE SOI E O ÍNDICE SPI-12 A VÁRIOS MESES DE ANTECEDÊNCIA PARA A REGIÃO DO ALENTEJO. ....	39
TABELA 5.8 INDICADORES HSS E PSS PARA AS PREVISÕES DE 1 A 3 MESES PARA OS MODELOS COM SPI-1 (MENSAL), COM O ÍNDICE DA NAO, AO E SOI E SEM ÍNDICE DE CGA NAS VÁRIAS OPÇÕES DE HOMOGENEIDADE E NÃO HOMOGENEIDADE, NA REGIÃO DO ALENTEJO.....	40
TABELA 5.9 INDICADOR PC PARA AS PREVISÕES DE 1 A 3 MESES PARA OS MODELOS COM SPI-1 (MENSAL), COM O ÍNDICE DA NAO, AO E SOI E SEM ÍNDICE DE CGA NAS VÁRIAS OPÇÕES DE HOMOGENEIDADE E NÃO HOMOGENEIDADE, NA REGIÃO DO ALENTEJO. ....	40

TABELA 5.10 INDICADORES HSS E PSS PARA AS PREVISÕES DE 1 A 3 MESES PARA OS MODELOS COM SPI-12 (ANUAL), COM O ÍNDICE DA NAO, AO E SOI E SEM ÍNDICE DE CGA NAS VÁRIAS OPÇÕES DE HOMOGENEIDADE E NÃO HOMOGENEIDADE, NA REGIÃO DO ALENTEJO.....	41
TABELA 5.11 INDICADOR PC PARA AS PREVISÕES DE 1 A 3 MESES PARA OS MODELOS COM SPI-12 (ANUAL), COM O ÍNDICE DA NAO, AO E SOI E SEM ÍNDICE DE CGA NAS VÁRIAS OPÇÕES DE HOMOGENEIDADE E NÃO HOMOGENEIDADE, NA REGIÃO DO ALENTEJO.....	41
TABELA 5.12 VECTOR DE EQUILÍBRIO DO ÍNDICE DE CGA PARA A REGIÃO DO ALENTEJO.....	42
TABELA 5.13 VECTOR DE EQUILÍBRIO DO SPI-1 E SPI-12 PARA A REGIÃO DO ALENTEJO.....	42
TABELA 5.14 TEMPO DE RESIDÊNCIA ESPERADO NA CLASSE I (MESES) PARA A REGIÃO DO ALENTEJO.....	43
TABELA 5.15 TEMPO DE RECORRÊNCIA (MESES) PARA AS VÁRIAS CLASSES DE SECA NOS CASOS DE SPI-1 E SPI-12, PARA A REGIÃO DO ALENTEJO.....	44
TABELA 5.16 TEMPO MÉDIO DE CHEGADA À CLASSE DE "NÃO-SECA" (N), PARTINDO NUM DADO MÊS DE UMA CLASSE DE SECA I (1, 2 OU 3), NOS CASOS DE SPI-1 E SPI-12, PARA A REGIÃO DO ALENTEJO.....	44
TABELA 5.17 PREVISÕES A CURTO PRAZO NOS CASOS DE SPI-1, NOS VÁRIOS MODELOS, PARA A REGIÃO DO ALENTEJO.....	45
TABELA 5.18 PREVISÕES A CURTO PRAZO NOS CASOS DE SPI-12, NOS VÁRIOS MODELOS, PARA A REGIÃO DO ALENTEJO.....	46

## Lista de Acrónimos

AO	<i>Arctic Oscillation</i>
CPC	<i>Climate Prediction Center</i>
CGA	Circulação Geral Atmosférica
EA	<i>Eastern Atlantic</i>
EOF	<i>Empirical Orthogonal Function</i>
EUA	Estados Unidos da América
ENSO	<i>El-Niño Southern Oscillation</i>
HSS	<i>Heidke Skill Score</i>
IM	Instituto de Meteorologia
INAG	Instituto da Água I. P.
NAO	<i>North Atlantic Oscillation</i>
NASA	<i>National Aeronautics and Space Administration</i>
NCAR	<i>National Center for Atmospheric Research</i>
NCEP	<i>National Center of Environmental Prediction</i>
NOAA	<i>National Oceanic and Atmospheric Administration</i>
PDSI	<i>Palmer Drought Severity Index</i>
PI	Península Ibérica
PSS	<i>Peirce Skill Score</i>
SLP	<i>Sea Level Pressure</i>
SOI	<i>Southern Oscillation Index</i>
SPI	<i>Standard Precipitation Index</i>
SST	<i>Sea Surface Temperature</i>

## Lista de Símbolos

$AO$	<i>Arctic Oscillation</i>
$CGA$	Circulação Geral Atmosférica
$DL$	Duração da seca
$DM$	Magnitude da seca
$DMM$	Magnitude média da seca ou Intensidade da seca
$e$	Número de estados da cadeia de Markov
$ENSO$	El-Nino Southern Oscillation
$g(x)$	Distribuição gama
$HSS$	<i>Heidke Skill Score</i>
$l_f$	Mês final do período de seca
$M$	Matriz estocástica de transição do índice $SPI$ num instante $t$ para um instante $t+1$ , no caso homogéneo
$M_{ij}$	Entrada da matriz $M$ correspondente à probabilidade de transição de $i$ para $j$ num passo de tempo
Markov-SPI	Modelo de Markov constrangido apenas pelo índice SPI
Markov-SPI-NAO	Modelo de Markov constrangido pelos índices SPI e NAO
$n_{ij}$	Número de transições entre o estado $i$ e o estado $j$
$NAO$	<i>North Atlantic Oscillation</i>
$NAO_t$	Vector da Oscilação do Atlântico Norte ( $2 \times 2$ )
$NAO_+^{(t)}$	Vector da fase positiva da Oscilação do Atlântico Norte ( $2 \times 2$ )
$NAO_-^{(t)}$	Vector da fase negativa da Oscilação do Atlântico Norte ( $2 \times 2$ )
$p_i$	Percentagem total de observações da classe $i$ da amostra numa tabela de contingência
$\hat{p}_i$	Percentagem total de previsões da classe $i$ da amostra numa tabela de contingência
$p_{ii}$	Percentagem de previsões concordantes com as observações para uma determinada classe de seca $i$ numa tabela de contingência
$P$	Matriz estocástica de transição de um estado de um determinado índice de $CGA$ num instante $t$ para um instante $t+1$ , no caso homogéneo
$p_{ij}$	Probabilidade da cadeia de Markov estar num próximo instante no estado $j$ , dado que no instante presente está no estado $i$ , no caso homogéneo
$p_{ij}^{t,t+1}$	Probabilidade da cadeia de Markov estar num próximo instante $t+1$ (mês) no estado $j$ , dado que no instante presente $t$ (mês) está no estado $i$ , no caso não homogéneo
$P_{i,j}$	Entrada da matriz $P$ correspondente à probabilidade de transição de $i$ para $j$ num passo de tempo

$P_{i,l}$	Precipitação do mês $j$ e do ano $i$
$PC$	<i>Proportion Correct</i>
$PSS$	<i>Peirce's Skill Score</i>
$r$	Coefficiente de correlação de <i>Pearson</i>
$S$	Conjunto de estado do índice <i>SPI</i> discretizado numa rede de Markov
$SPI$	<i>Standard Precipitation Index</i>
$SPI_t$	Variável Aleatória <i>SPI</i> num instante de tempo $t$
$t$	Tempo
$X_{i,j}^k$	Precipitação acumulada para um mês $j$ particular do ano $i$ numa escala de tempo de $k$ meses
$X_t$	Variável Aleatória $X$ num instante de tempo $t$
$\Gamma$	Função gama
$\alpha$	Parâmetro de forma da distribuição gama
$\beta$	Parâmetro de escala da distribuição gama

# Índice

<b>AGRADECIMENTOS</b> .....	<b>II</b>
<b>RESUMO</b> .....	<b>III</b>
<b>ABSTRACT</b> .....	<b>IV</b>
<b>LISTA DE FIGURAS E TABELAS</b> .....	<b>V</b>
<b>LISTA DE ACRÓNIMOS</b> .....	<b>VII</b>
<b>LISTA DE SÍMBOLOS</b> .....	<b>VIII</b>
<b>ÍNDICE</b> .....	<b>X</b>
<b>1. INTRODUÇÃO</b> .....	<b>1</b>
1.1    MOTIVAÇÃO.....	1
1.2    OBJECTIVOS .....	1
1.3    ESTRUTURA DA DISSERTAÇÃO .....	2
<b>2. O FENÓMENO DA SECA</b> .....	<b>3</b>
2.1    INTRODUÇÃO .....	3
2.2    ESCASSEZ DE ÁGUA.....	4
2.3    DEFINIÇÕES DE SECA .....	5
2.4    O ÍNDICE SPI.....	7
2.4.1    Processo de cálculo do SPI.....	7
2.4.2    Estatísticas associadas ao SPI.....	10
2.5    TELECONEXÕES E CIRCULAÇÃO ATMOSFÉRICA .....	10
2.5.1    O índice NAO ( <i>North Atlantic Oscillation</i> ) .....	11
2.5.2    O índice AO.....	14
2.5.3    O índice SOI ( <i>Southern Oscillation Index</i> ).....	15
<b>3. MODELAÇÃO DA SECA</b> .....	<b>17</b>
3.1    CADEIAS DE MARKOV .....	17
3.1.1    Algumas aplicações.....	17
3.1.2    Processo estocástico Markoviano .....	17
3.1.3    Matrizes e vectores estocásticos.....	18
3.1.4    Propriedades da Cadeia de Markov.....	19
3.1.4.1    Vector de equilíbrio – “ <i>Steady State Vector</i> ” .....	21
3.1.4.2    Tempo médio de permanência ininterrupta no estado <i>i</i> .....	23
3.1.4.3    Tempo médio para a primeira visita ao estado <i>j</i> .....	24
3.1.4.4    Previsões a curto prazo.....	25
3.2    APLICAÇÃO DOS ÍNDICES DE CGA ÀS CADEIAS DE MARKOV .....	26
3.2.1    Vector de equilíbrio – “ <i>Steady State Vector</i> ”.....	28
3.2.2    Tempo médio de permanência ininterrupta no estado <i>i</i> .....	29
3.2.3    Tempo médio para a primeira visita ao estado <i>j</i> .....	30
3.2.4    Previsões a curto prazo .....	30
3.3    INDICADORES DE PRECISÃO – “ <i>SKILL SCORES</i> ” .....	31
3.3.1    HSS ( <i>Heidke Skill Score</i> ) e PSS ( <i>Peirce’s Skill Score</i> ).....	31
3.3.2    Proporção Correcta (PC).....	32
3.3.3    Correlação.....	32
<b>4. DADOS</b> .....	<b>33</b>
<b>5. RESULTADOS E DISCUSSÃO</b> .....	<b>36</b>
5.1    MATRIZES DE TRANSIÇÃO .....	36
5.2    CORRELAÇÃO ENTRE OS ÍNDICES DE SECA E DE CGA.....	37
5.3    PERFORMANCE DOS MODELOS MARKOV-SPI E MARKOV-SPI-ÍNDICECGA.....	39
5.4    VECTOR DE EQUILÍBRIO .....	41
5.5    TEMPO MÉDIO DE PERMANÊNCIA ININTERRUPTA NO ESTADO I .....	42

5.6	TEMPO MÉDIO PARA A PRIMEIRA VISITA AO ESTADO J .....	43
5.7	PREVISÕES A CURTO PRAZO .....	45
<b>6.</b>	<b>CONCLUSÕES.....</b>	<b>48</b>
<b>7.</b>	<b>REFERÊNCIAS.....</b>	<b>50</b>

# 1. Introdução

## 1.1 Motivação

A seca é um fenómeno natural que resulta da persistência de valores de precipitação abaixo da média. A sua frequência, duração e severidade são incertas, o que a tornam num fenómeno pouco predictivo, e por isso também, perigoso e desastroso. Os impactos negativos da sua ocorrência, nos diversos ramos de actividade, agrícola, industrial e urbano e nos ecossistemas poderão ser minimizados, se se tiver um melhor conhecimento do fenómeno, das suas causas e da sua possível evolução.

Para tal, a abordagem estocástica contribuirá para uma melhor caracterização deste fenómeno aleatório e da sua evolução no tempo, constituindo também uma ferramenta auxiliar de predição.

Neste sentido, várias metodologias têm sido adoptadas na identificação, monitorização e caracterização das secas, dentro as quais, os modelos de cadeias de Markov e os modelos Log-Lineares são bons exemplos das mesmas.

Devido à forte influência dos padrões da Circulação Geral Atmosférica na precipitação sazonal existe a possibilidade de serem usados índices de Circulação Geral Atmosférica como preditores, com o intuito de melhorar a performance dos modelos de previsão/predição de seca. Estes índices são integrados nos modelos estocásticos.

## 1.2 Objectivos

Pretende-se dar continuidade a trabalhos anteriores Paulo e Pereira (2005, 2006 e 2007) e Paulo, (2004) no domínio das secas no sul de Portugal, mais concretamente em diversas zonas do Alentejo. Nestes trabalhos foram utilizados modelos Log-Lineares (Fonseca e Coelho, 1998) e Ferreira (2004) e modelos de cadeias de Markov simples (Doob, 1953), ou seja, definidos apenas por classes de severidade de seca a partir dos índices de seca, SPI (Standard Precipitation Index) e/ou PDSI (Palmer Drought Severity Index).

Pretende-se agora, neste trabalho desenvolver dois modelos de cadeias de Markov: o primeiro modelo, à semelhança dos trabalhos acima citados, apenas definido a partir de classes de severidade de seca pelo índice SPI, designado de **Markov-SPI**; e outro modelo designado **Markov-SPI-preditor**<sup>1</sup>, em que as classes de severidade de seca são definidas como no modelo anterior e ainda com o acoplamento de informação proveniente de índices de CGA (ex: *NAO, AO, SOI, etc...*).

Posteriormente são validados estes dois modelos em modo *Hindcast*, ou seja, em que as previsões correspondem a instantes já ocorridos no passado e são comparadas com as correspondentes observações. É avaliada e comparada a *performance* de cada um destes dois modelos através de indicadores de precisão

---

<sup>1</sup> *preditor* corresponde a qualquer um dos índices CGA. No caso de ser o índice NAO o modelo será **Markov-SPI-NAO**.

adequados a variáveis discretas, tais como *HSS* (*Heidke Skill Score*) e *PSS* (*Peirce Skill Score*).

São ainda, testadas as versões homogéneas e não-homogéneas destes dois modelos. Na versão homogénea, o modelo de Markov tem uma matriz de transição entre classes de seca (SPI) que é constante ao longo do ano. Na versão não homogénea a matriz de transição tem um ciclo anual dependendo do mês correspondente do vector de classes de seca e que evolui ciclicamente à medida que é integrada a cadeia de Markov. Em ambas as versões são determinados vários itens importantes para caracterizar secas tais como:

- ✓ *probabilidade climatológica da classe de seca,*
- ✓ *tempo médio de residência ininterrupta no estado ou classe de seca  $i$ ,*
- ✓ *tempo médio para a primeira visita ao estado ou classe de seca  $j$ ,*

A integração temporal dos modelos de Markov permite obter previsões probabilistas do vector de probabilidades das classes de seca. Medidas de qualidade dessas previsões são obtidas no modo Hindcast).

Todos os itens referidos são obtidos algebricamente de modo simples a partir da matriz de transição do modelo de cadeia de Markov. A matriz de transição de classes é suficiente para a formulação de um modelo probabilista. Previsões probabilistas sazonais de precipitação e índices de seca podem igualmente ser obtidas por previsões de ensemble multi-modelos, tal como, foi executado nos projectos europeus DEMETER (<http://www.ecmwf.int/research/demeter/>) e ENSEMBLES (<http://ensemble.jrc.ec.europa.eu/>), que tem obviamente em conta uma muito maior variedade de graus de liberdade físicos do sistema climático.

### **1.3 Estrutura da dissertação**

Esta dissertação está estruturada em 7 capítulos, que incluem um enquadramento teórico aos assuntos abordados, a apresentação de resultados com sua discussão e conclusões.

No capítulo 2 é abordado em linhas gerais o tema da seca, as suas características e o seu impacto socioeconómico. É descrito, como quantificar a seca através do índice SPI e como calculá-lo. São apresentados alguns índices de circulação geral atmosférica (CGA) associados a padrões atmosféricos de teleconexão que influenciam a precipitação no Península Ibérica.

No capítulo 3 é apresentado o modelo estocástico do tipo cadeias de Markov, bem como a sua adaptação ao fenómeno da seca em estudo, com e sem acoplamento de índices de CGA.

No capítulo 4 é mostrada a localização dos dados utilizados neste trabalho bem como as fontes destes.

No capítulo 5 segue-se a apresentação dos resultados obtidos e sua discussão. E nos capítulos 6 e 7 contêm as habituais conclusões e referencias essenciais ao caso de estudo.

## 2. O fenómeno da seca

### 2.1 Introdução

As situações de seca constituem uma ocorrência natural associada essencialmente à falta de precipitação, que se verifica todos os anos em diversas regiões do mundo. Contrariamente aos outros desastres naturais, que geralmente actuam de forma rápida e com impactos imediatos, a seca é o desastre natural de origem meteorológica e climatológica mais complexo e que afecta mais pessoas e durante mais tempo que qualquer outro.

Os impactos resultantes deste fenómeno variam conforme a escala espacial e temporal. Longos períodos de seca provocam graves prejuízos económicos, nomeadamente ao nível dos sectores agrícola, pecuário e recursos hídricos, originando muitas vezes o desenvolvimento e propagação de pragas e pestes, o que designadamente em países com economias débeis leva à escassez de alimentos e consequentemente à perda de um número muito significativo de vidas humanas (estatísticas apontam, em média, para um número superior a 30 000 por ano).

Sendo embora um desastre natural que não poderá ser evitado, os seus impactos poderão ser minimizados através da deslocação de grandes quantidades de água ou da promoção do estabelecimento de mecanismos para o seu armazenamento; por seu lado, a má gestão do uso do solo e práticas agrícolas inadequadas, contribuem para a degradação dos solos e dos recursos hídricos, aumentando a vulnerabilidade dos e populações a eventos de seca.

O problema das secas deve enquadrar-se em anomalias da circulação geral da atmosfera, a que correspondem flutuações do clima numa escala local ou regional. A situação geográfica do território de Portugal Continental é favorável à ocorrência de episódios de seca, quase sempre associados a situações de bloqueio em que o anticiclone subtropical do Atlântico Norte se mantém numa posição que impede que as perturbações da frente polar atinjam a Península Ibérica. Na Figura 2.1 apresenta-se uma situação de bloqueio no Inverno com o anticiclone dos Açores anormalmente desenvolvido.

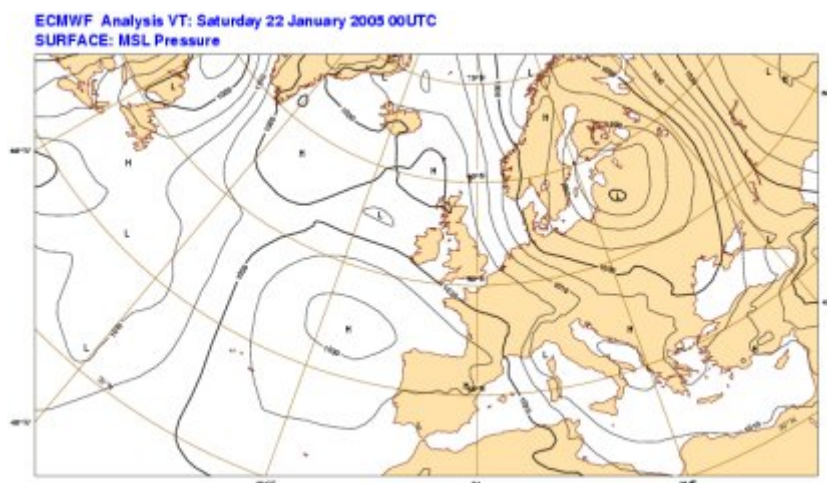


Figura 2.1 Situação sinóptica (pressão ao nível médio do mar) de 22 de Janeiro de 2005.

As situações de seca são frequentes em Portugal Continental, com consequências desastrosas na agricultura e na pecuária, nos recursos hídricos e no bem-estar das populações, sendo de destacar, nos últimos 65 anos, os anos de seca com maior severidade: 1944/46, 1965, 1976, 1980/81, 1991/92, 1994/95 e 1998/99, 2004/05. Nomeadamente as regiões a Sul do Tejo são as mais vulneráveis, e as que têm sido mais afectadas.

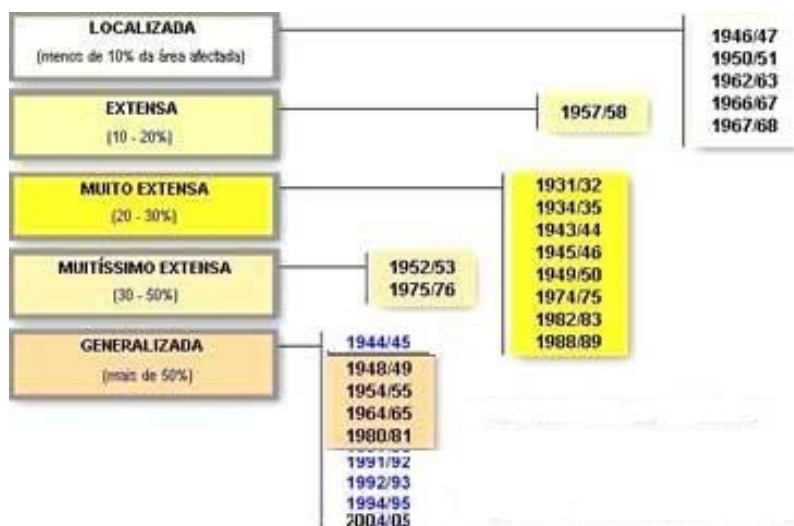


Figura 2.2 Secas em Portugal Continental de maior extensão desde 1940 (Fonte: IM).

## 2.2 Escassez de água

Há décadas atrás, a água era vista como um recurso natural ilimitado uma vez que era renovável todos os anos no decorrer das estações. O Homem progressivamente apropriou-se deste recurso e passou a usá-lo com poucas restrições. Evolução no controlo e desvio de águas superficiais, exploração de águas subterrâneas, e uso de recursos para uma variedade de finalidades têm sido realizados sem a necessária precaução à conservação deste recurso natural, evitando desperdícios e mau uso, e preservando a qualidade deste recurso. Assim, hoje em dia, a água tem-se tornado escassa não apenas em regiões áridas e propensas a secas, mas também em áreas onde a precipitação é relativamente abundante. A escassez é agora vista sobre a perspectiva das quantidades disponíveis para os usos económicos e sociais, bem como em relação aos requisitos de água para os ecossistemas naturais e artificiais. O conceito de escassez também abraça a qualidade de água uma vez que muitas vezes os recursos de água disponíveis estão degradados ou no melhor dos casos apenas marginalmente disponíveis para uso em sistemas humanos e naturais (Pereira *et al.*, 2002).

Mundialmente, a agricultura é o sector que tem a maior exigência de água. Como resultado da sua vasta utilização na agricultura, a irrigação é frequentemente considerada a principal causa de escassez de água. A irrigação é considerada, como sendo, uma má utilização de água, de produzir excessivos desperdícios e de degradar a qualidade da mesma. Contudo, a irrigação na agricultura fornece os meios de subsistência de uma enorme parte da população do mundo rural e ainda uma grande parte dos alimentos de todo o mundo. Na actualidade, a irrigação da agricultura é largamente afectada pela escassez de recursos hídricos. Existem esforços por parte de agências de financiamento e gestores no sentido de criar incentivos para inovar e melhorar as práticas de gestão da água, para controlar os impactos negativos da irrigação, para diversificar a utilização de água nos projectos

de irrigação, e para aumentar a produtividade e os rendimentos dos agricultores. Paralelamente, grandes progressos na engenharia e na gestão económica estão a produzir novas considerações para a utilização da água e controlo de qualidade da água para fins não agrícolas, particularmente para consumo doméstico e saneamento (Pereira *et al.*, 2002).

A escassez de água pode resultar de um vasto leque de fenómenos que pode ser devida a causas naturais, induzida por actividades humanas, ou podem resultar da interacção de ambos, como indicado na Tabela 2.1 (Vlachos e James, 1983).

**Tabela 2.1 Regimes de escassez de água (Pereira *et al.*, 2002).**

Escassez de água	Natural	Origem humana
Permanente	<i>Aridez</i>	<i>Desertificação</i>
Temporária	<i>Seca</i>	<i>Penúria de água</i>

A **aridez** é um desequilíbrio natural permanente de água disponível consistindo numa baixa precipitação média anual, com grande variabilidade espacial e temporal, resultando em geral numa baixa humidade e reduzida capacidade de sustento dos ecossistemas (Pereira *et al.*, 2002).

A **seca** é um desequilíbrio natural embora temporário da água disponível, consistindo numa precipitação persistentemente inferior à média, com frequência incerta, duração e severidade de ocorrência imprevisível ou difícil de prever, resultando numa diminuição dos recursos de água disponíveis, e reduzida capacidade de suporte dos ecossistemas (Pereira *et al.*, 2002).

A **Desertificação** é um desequilíbrio da disponibilidade de água em regiões áridas e sub-húmidas devido à degradação da terra, nomeadamente à sobreexploração das águas subterrâneas e/ou superficiais, à degradação do solo, erosão, uso inadequado da terra, redução da infiltração, cheias rápidas mais frequentes, e perda de ecossistemas ripícolas<sup>2</sup> e outros (Paulo, 2004).

A **Penúria de água** é um desequilíbrio na disponibilidade de água incluindo sobreexploração de aquíferos, redução de capacidade dos reservatórios, uso inadequado da terra, degradação da qualidade da água e redução da capacidade de suporte dos ecossistemas (Pereira *et al.*, 2002, Paulo, 2004).

## 2.3 Definições de Seca

A seca tem sido definida pela comunidade meteorológica internacional em termos de "ausência prolongada ou deficiência marcada na precipitação", uma "deficiência da precipitação que resulta numa redução da água disponível para algumas actividades ou grupos", ou um "período de tempo anormalmente seco e suficientemente prolongado para que a ausência de precipitação cause um sério desequilíbrio hidrológico" (WMO, 1975; AMS, 1997).

<sup>2</sup> **Floresta ripícola** é aquela cujas plantas crescem principalmente ao longo de um curso de água e/ou cujas raízes alcançam a franja de movimento capilar aquoso.

Em termos muito gerais define-se seca como um período de persistência anómala de tempo seco de modo a causar problemas na agricultura, na pecuária e/ou no fornecimento de água.

A definição de seca depende do ponto de vista do utilizador. Em geral distingue-se entre seca meteorológica, seca agrícola, seca hidrológica e seca sócio-económica, (ver Figura 2.3), (Whilite e Glantz, 1987).

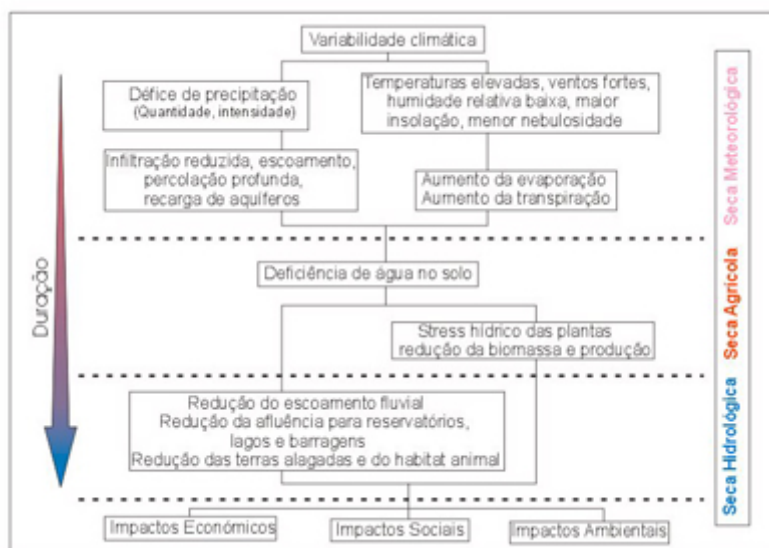


Figura 2.3 Esquema da sequência temporal dos diversos tipos de seca (adaptado de [www.drought.unl.edu](http://www.drought.unl.edu), 2002).

A **seca meteorológica** é uma medida do desvio da precipitação em relação ao valor normal; caracteriza-se pela falta de água induzida pelo desequilíbrio entre a precipitação e a evaporação, a qual depende de outros elementos como a velocidade do vento, temperatura e humidade do ar, insolação. A definição de seca meteorológica deve ser considerada como dependente da região, uma vez que, as condições atmosféricas que resultam em deficiências de precipitação podem ser muito diferentes de região para região.

A **seca agrícola** está associada à falta de água causada pelo desequilíbrio entre a água disponível no solo, a necessidade das culturas e a transpiração das plantas. Este tipo de seca está relacionado com as características das culturas, da vegetação natural, ou seja, dos sistemas agrícolas em geral.

A **seca hidrológica** está relacionada com a redução dos níveis médios de água nos reservatórios e com a depleção de água no solo. Este tipo de seca ocorre posteriormente à seca meteorológica e agrícola, dado que é necessário um período maior para que as deficiências na precipitação se manifestem nos diversos componentes do sistema hidrológico.

A **seca sócio-económica** está associada ao efeito conjunto dos impactos naturais e sociais que resultam da falta de água, devido ao desequilíbrio entre o fornecimento e a procura dos recursos de água e que vai afectar directamente as populações (Pires, 2003).

A quantidade e distribuição anual da precipitação, assim como valores elevados da temperatura do ar, são condições que determinam a intensidade e consequências de uma seca. De forma a estimar a possibilidade de ocorrência de uma seca, ou o seu grau de severidade, devem conhecer-se as condições climáticas. Neste

contexto para uma certa área, os dados meteorológicos mais importantes que se devem conhecer são a precipitação, a temperatura do ar, a humidade do ar e o conteúdo de água no solo.

Quando se dá início de uma seca, os efeitos desta fazem-se sentir de imediato na agricultura, uma vez que esta está directamente dependente do armazenamento de água no solo, podendo outras actividades que dependem das reservas superficiais vir a ser afectadas posteriormente. Geralmente, as actividades que dependem das reservas subterrâneas são as últimas a ser afectadas. Quando as condições normais de precipitação são restabelecidas, a reposição de água faz-se no sentido inverso: primeiramente as reservas de água no solo, seguidas dos caudais dos cursos de água e os reservatórios e lagos e por último as águas subterrâneas. O tempo de recuperação está dependente da duração e severidade da seca, e da precipitação verificada depois do seu término (Paulo, 2004).

## 2.4 O índice SPI

O SPI (Standard Precipitation Index) é um índice desenvolvido por McKee *et al.* (1993) com o propósito de definir e monitorizar secas locais. Foi concebido também, para identificar períodos de seca e severidade de secas em variadas escalas temporais. Embora tenha sido concebido para a identificação de períodos secos, também pode ser usado na identificação de períodos húmidos. As escalas temporais mais analisadas são geralmente a mensal (SPI-1 mês) e a sazonal ou trimensal (SPI-3 meses), embora se calculem também SPI anual (SPI-12 meses). Escalas temporais diferentes reflectem o atraso na resposta dos diferentes recursos hídricos às anomalias da precipitação. À medida que a escala temporal aumenta, a resposta do SPI tornam-se mais lenta a mudanças na precipitação (McKee *et al.*, 1993). Se, por exemplo, o SPI calculado numa escala de tempo de 3 meses indicar, para um determinado mês, uma situação de deficit, isto significa que a precipitação acumulada dos últimos 3 meses é inferior à precipitação histórica "normal" desses mesmos meses. Assim, diferentes tipos de seca (ex: seca meteorológica, seca agrícola, seca hidrológica, seca sócio-económica) requerem períodos de SPI diferentes em cada um dos casos. Assim uma escala de 3 meses poderá ser mais adequada para avaliar uma seca agrícola e uma escala de 12 ou 24 meses para uma seca hidrológica (Paulo, 2004).

Condições de humidade do solo respondem às anomalias da precipitação num espaço de tempo relativamente curto, enquanto que a quantidade de água no solo, a corrente dos rios e os reservatórios de armazenamento de água reflectem anomalias da precipitação de longo prazo (Paulo *et al.*, 2005).

O SPI é calculado com base na distribuição de probabilidades da precipitação do período escolhido. Deste modo, os valores do índice SPI obtidos dependem da função de distribuição escolhida, dos valores da amostra a partir dos quais são determinados os parâmetros da distribuição e, também, do método de estimação. O SPI permite fazer comparações entre diferentes locais e diferentes períodos de tempo, devido ao facto, de este índice indicar o afastamento entre a precipitação observada e a média para um mês considerado, e ser também normalizado pelo desvio padrão local da precipitação desse mês, (Paulo, 2004).

### 2.4.1 Processo de cálculo do SPI

A precipitação total  $X_{i,j}^k$  para um dado ano  $i$  e mês  $j$  depende da escala de tempo  $k$  escolhida ( $k$  em meses). Esta quantidade é simplesmente a precipitação acumulada nos  $k$  meses anteriores a partir da data (ano= $i$ , mês= $j$ ), inclusive o primeiro mês. Por exemplo  $X_{i=2008,j=1}^3$  é a soma das 3 precipitações mensais de (Nov. 2008), (Out. 2008) e (Set. 2008). Em termos de uma expressão matemática, têm-se por exemplo a precipitação acumulada  $X_{i,j}^k$  para um mês  $j$  particular do ano  $i$  com uma escala de tempo  $k=12$  meses é dada por (Paulo *et al.*, 2003):

$$X_{i,j}^k = \sum_{l=13-k+j}^{12} P_{i-1,l} + \sum_{l=1}^j P_{i,l}, \text{ se } j < k, \text{ e} \quad (2.1)$$

$$X_{i,j}^k = \sum_{l=j-k+1}^j P_{i,l}, \text{ se } j \geq k.$$

onde  $P_{i,l}$  é a precipitação do mês  $j$  e do ano  $i$  [mm].

Cada conjunto de dados é então ajustado à distribuição gama para definir a relação entre a probabilidade e a precipitação. A distribuição gama é definida para uma variável aleatória contínua positiva  $x$  pela função de densidade de probabilidade:

$$g(x) = \frac{1}{\beta} \left( \frac{x}{\beta} \right)^{\alpha-1} \frac{\exp\left(-\frac{x}{\beta}\right)}{\Gamma(\alpha)}, \quad x > 0 \quad (2.2)$$

onde  $\alpha > 0$  e  $\beta > 0$  são respectivamente, os parâmetros de forma e de escala. A função gama  $\Gamma(\alpha)$  na fórmula (2.2) é definida por:

$$\Gamma(\alpha) = \int_0^{\infty} y^{\alpha-1} e^{-y} dy \quad (2.3)$$

A distribuição da frequência de precipitação total é ajustada à função densidade de probabilidade gama, sendo os parâmetros  $\alpha$  e  $\beta$  estimados para cada escala de tempo (número de meses de acumulação da precipitação) e para cada mês do ano (mês final do período de acumulação) Esse ajuste é facilmente obtido através das expressões da média  $E(x) = \alpha\beta$  e variância de  $x$ :  $\text{var}(x) = \alpha\beta^2$ . Os estimadores de  $\alpha$  e  $\beta$  podem ser obtidos através da média e variância empíricas de  $x$ . Para valores suficientemente elevados de  $\alpha$ , a densidade de probabilidade  $g(x)$  aproxima-se de uma densidade Gaussiana. Para um dado valor da precipitação acumulada  $X_{i,j}^k$ , calcula-se a probabilidade  $p$ , integrada a partir da função densidade  $g(x)$ , de essa precipitação ser inferior a  $X_{i,j}^k$  ou seja:  $p = \int_0^{X_{i,j}^k} g(x) dx$ . O valor do SPI correspondente ao ano  $i$ , mês  $j$  e escala de tempo  $k$  (em meses):  $SPI_{i,j}^k$  é o quantil da função densidade normalizada (média nula e desvio padrão unitário) correspondente a  $p$  ou seja:

$$p = \int_0^{X_{i,j}^k} g(x) dx = \int_{-\infty}^{SPI_{i,j}^k} \varphi(x) dx \text{ ou } SPI_{i,j}^k = \Phi^{-1} \left( \int_0^{X_{i,j}^k} g(x) dx \right)$$

onde  $\varphi(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right)$  é a densidade Gaussiana standard e  $\Phi(\ )$  é a função de probabilidade acumulada.

Assim o SPI é o quantil Gaussiano correspondente à probabilidade  $p$  associada à função de densidade Gama. Há assim uma correspondência biunívoca entre a precipitação acumulada e o SPI (Dutra, 2006). Em termos pictóricos, tal é exemplificado na Figura 2.4.

O SPI calculado desta forma possui as seguintes características (McKee *et al.*, 1993):

- O SPI está unicamente relacionado com a probabilidade;
- A precipitação utilizada no SPI pode ser utilizada para calcular o déficit de precipitação para o período utilizado;
- A precipitação utilizada no SPI pode ser utilizada para calcular a percentagem corrente da média da precipitação para o período de tempo  $k$  meses;
- O SPI está normalmente distribuído, pelo que pode ser utilizado para monitorizar períodos secos como períodos húmidos;
- O SPI está normalizado, pelo que períodos secos e húmidos são representados de formas similares.

McKee *et al.* (1993) definiu arbitrariamente a severidade de uma seca segundo a Tabela 2.2. A frequência esperada em cada classe de severidade foi baseada na análise de um grande número de séries de precipitação no Colorado (EUA). As percentagens de frequência nas várias classes correspondem aos valores esperados de uma distribuição normal estandardizada do SPI.

**Tabela 2.2 Classificação da intensidade da seca de acordo com o SPI e frequência de ocorrência de cada classe.**

<b>Valores de SPI</b>	<b>Categoria de seca</b>	<b>Frequência na categoria</b>
0.00 a -0.99	Seca ligeira	34.0%
-1.00 a -1.49	Seca moderada	9.2%
-1.50 a -1.99	Seca severa	4.4%
≤ -2.00	Seca extrema	2.3%
<i>Total</i>		50.0%

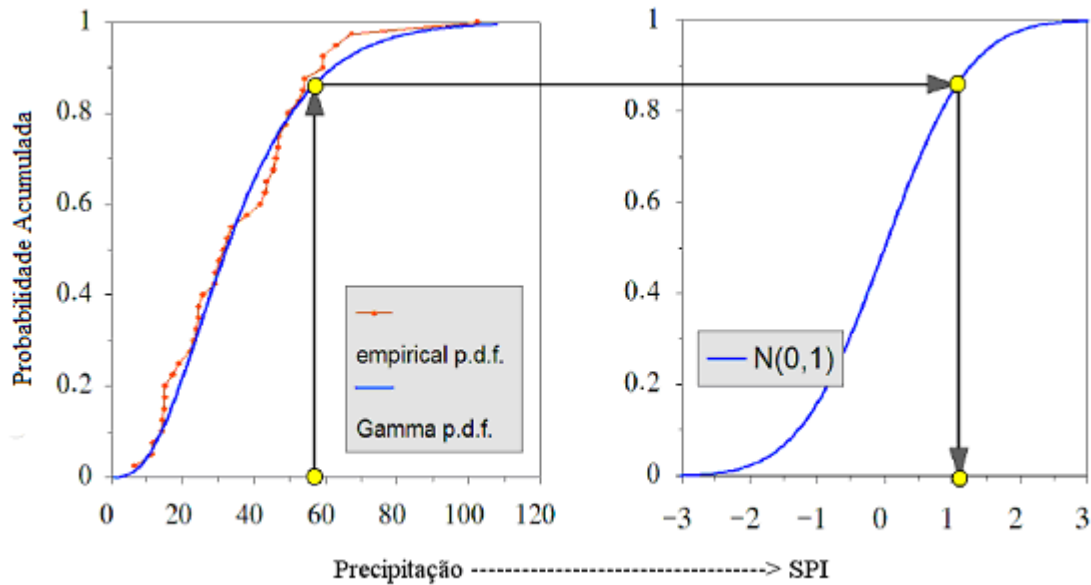


Figura 2.4 Transformação dos valores da precipitação, ajustados à função gama, numa distribuição normal de valor médio zero e desvio padrão unitário (Dutra, 2006).

## 2.4.2 Estatísticas associadas ao SPI

Os valores mensais do SPI variam num intervalo contínuo, isto é tem-se uma variável dependente contínua (SPI) de uma variável independente descontínua (mês). Uma vez que os valores de SPI são obtidos tendo em conta os ciclos anuais da média e do desvio padrão das precipitações mensais, poderemos admitir que a série temporal do SPI é uma realização de um processo estocástico estacionário, na hipótese de não haver tendência climática. Ao longo da série existem períodos de meses consecutivos mais ou menos extensos em que  $SPI \leq -1$  ou seja em que existiu seca, pelo menos moderada. O fim da seca surge no mês em que  $SPI > -1$ . Assim, do ponto de vista do SPI, seca corresponde a um período sequencial de um ou mais meses em que  $SPI \leq -1$ . Cada seca tem assim uma certa duração DL em meses, uma magnitude média DMM, que é a média dos valores de  $-SPI$  ao longo dos DL meses. Define-se igualmente magnitude da seca DM, em unidades de SPIxmeses como o produto:  $DM = DMM \times DL$ . Têm-se assim:

$$DMM = -\frac{1}{DL} \sum_{l=1}^{DL} SPI(l),$$

$$DM = DL \times DMM$$

## 2.5 Teleconexões e Circulação Atmosférica

A identificação de teleconexões e a análise de suas influências na estrutura horizontal da circulação atmosférica pode ser útil para a compreensão da ocorrência de eventos anómalos em várias regiões do globo. É bem sabido que a circulação atmosférica exhibe bastante variabilidade. Esta variabilidade reflecte padrões de tempo e sistemas de circulação que ocorrem com várias escalas de tempo distintas, perdurando desde alguns dias (característico de um sistema de tempestade normal e passagem de um sistema frontal), a várias semanas (próprio de aquecimentos a meio do Inverno ou de períodos húmidos a meio do Verão), a vários meses

(particularmente típico de invernos frios e verões quentes), a vários anos (característico de invernos anormais durante vários anos seguidos), a vários séculos (específico de mudanças climáticas de longo prazo).

O termo “padrão de teleconexão” refere-se a um recorrente e persistente padrão de larga escala das anomalias da pressão e circulação que alcança vastas áreas geográficas. Essas anomalias, mesmo que tomadas em locais muito distantes entre si, exibem assim correlações temporais significativas. Muitos padrões de teleconexão estão associados a variabilidade de baixa frequência (ou longas escalas temporais) da atmosfera. Apesar destes padrões perdurarem geralmente desde várias semanas a vários meses, podem por vezes estar activos durante vários anos consecutivos, reflectindo então uma parte importante da variabilidade interanual e interdecadal da circulação atmosférica. Muitos dos padrões de teleconexão são também de escala planetária, e atravessam bacias oceânicas inteiras e continentes. Por exemplo, certos padrões atravessam a bacia do Pacífico Norte inteira, enquanto que outros se estendem desde o leste da América do Norte até à Europa central. Outros ainda cobrem praticamente toda a Eurásia.

Os padrões das teleconexões na atmosfera podem ocorrer devido essencialmente a duas causas: a primeira é a ocorrência persistente durante meses e em larga escala espacial de anomalias dos forçamentos de superfície, nomeadamente através de anomalias da temperatura da superfície do mar, da humidade no solo, da quantidade de neve etc. São exemplos alguns padrões sobre o Pacífico Norte, forçados por alterações na temperatura superficial do mar (SST) e convecção tropical associadas aos ciclos ENSO (Mo e Livezey 1986, Barnston e Livezey 1991) e à Oscilação Madden-Julian (MJO). A segunda causa pode ser unicamente a variabilidade interna de baixa frequência da atmosfera, sem estar associada a variabilidade das condições fronteira de superfície.

Os padrões da teleconexões reflectem alterações a larga escala nas ondas atmosféricas e no padrão da corrente de jacto, precipitação, trajectórias de tempestades, e localização da corrente de jacto e sua intensidade sobre vastas áreas. Deste modo, as teleconexões são frequentemente responsáveis pelos padrões de tempo anómalos ocorrendo simultaneamente em vastas distâncias. A título de exemplo, o inverno de 1995 foi muito frio e nevado sobre grande parte da área do leste da América do Norte, enquanto que a Europa e a Escandinávia foram frios e o sul da Europa e Norte de África experimentaram condições muito húmidas e tempestuosas. Estas condições estiveram parcialmente relacionadas com o mesmo padrão de teleconexão: uma forte fase negativa da **Oscilação do Atlântico Norte (NAO)**.

### **2.5.1 O índice NAO (North Atlantic Oscillation)**

A Oscilação do Atlântico Norte é um fenómeno de variabilidade climática de larga escala que tem impactos importantes no tempo e no clima na região do Atlântico Norte, bem como nos continentes vizinhos, especialmente na Europa. É um dos padrões de teleconexões mais importante em todas as estações do ano (Barnston e Livezey, 1987), embora seja durante o inverno que é particularmente dominante. A NAO combina parte dos padrões do Atlântico Este e Atlântico Oeste originalmente identificados por Wallace e Gutzler (1981) para a estação do Inverno. Consiste num dipolo Norte-Sul de anomalias de pressão, com um dos centros localizado sobre a Gronelândia e o outro centro de sinal oposto “varrendo” as latitudes centrais do Atlântico Norte entre os 35°N e os 40°N. A NAO está também, relacionada com as alterações dos ventos de oeste sobre o Atlântico e representa uma oscilação meridional de massas atmosféricas. A intensidade  $I(\text{NAO})$  da NAO, é descrita pelo seu índice, sendo este definido pela diferença normalizada das pressões

normalizadas ao nível médio do mar entre duas estações situadas perto dos centros da depressão da Islândia e do Anticiclone dos Açores. Stykkisholmur na Islândia é usado como a estação Norte ao passo que Ponta Delgada, Lisboa ou Gibraltar são usadas para a estação a Sul. Assim, recorrendo a uma formula:

$$I' = \left( \frac{p - \bar{p}}{\sigma(p)} \right)_{Lisboa} - \left( \frac{p - \bar{p}}{\sigma(p)} \right)_{Islândia} ; I(NAO) = \frac{I'}{\sigma(I')} \quad (2.4)$$

onde  $\sigma$  significa desvio padrão. Existem índices alternativos, com elevada correlação temporal com o índice apresentado. Um deles é baseado na projecção sobre a primeira EOF (função empírica ortogonal) do campo da pressão nos meses de inverno na região Atlântica. A primeira EOF, ou seja, a que explica maior fracção da variância, exhibe uma estrutura dipolar da pressão no Atlântico que tem aproximadamente a mesma estrutura do campo do composto das anomalias quando  $I(NAO) > 0$  (fase positiva da NAO) e portanto podemos associar essa EOF à NAO.

A **fase positiva da NAO**, (ver Figura 2.5) mostra um centro de alta pressão mais forte que o normal na região dos Açores e uma depressão mais profunda que o normal na região da Gronelândia. O aumento do gradiente meridional de pressão em relação à climatologia ( $I(NAO)=0$ ), resulta na maior intensidade da corrente de jacto e no desvio para Norte da rota das tempestades (storm-track) de Inverno que atravessam o Oceano Atlântico. Consequentemente, os Invernos na Europa do Sul e na parte sudeste dos Estados Unidos (EU) são mais quentes e húmidos, sendo mais frios e secos no norte do Canadá, Gronelândia e Europa do Norte. Neste caso, a região Sul dos EU experimenta condições inverniais mais suaves e húmidas. A precipitação é mais abundante durante o Inverno, na Europa do Norte e Escandinávia, e abaixo do normal na Europa central e do sul (Lamb e Pepler, 1991, Hurrell, 1995). A título de exemplo o elevado índice da NAO no período inverno/primavera de 1989, 1990, e 1995, foi causado pelo deslocamento de uma massa de ar vinda do Ártico e das regiões da Islândia em direcção à faixa subtropical perto dos Açores e península Ibérica, tendo ainda os ventos de oeste aumentado de intensidade sobre a região do Atlântico Norte. Ventos de oeste mais fortes trazem mais ar quente e húmido para o continente Europeu no Sul e torna os invernos mais suaves nesta zona.

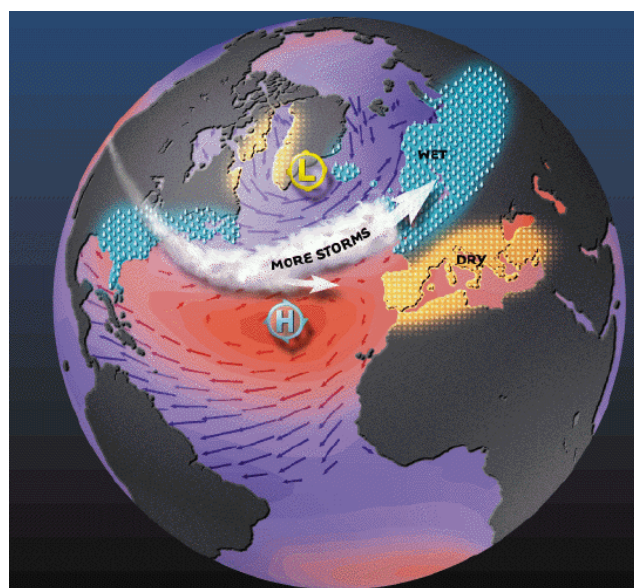
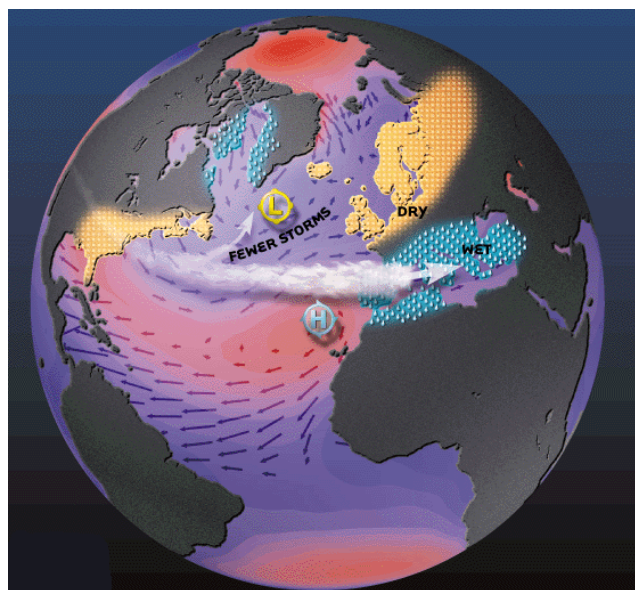


Figura 2.5 Fase positiva da NAO  
(Fonte: <http://www.atmosphere.mpg.de/enid/2.html>)

A **fase negativa da NAO**, (ver Figura 2.6), mostra um anticiclone subtropical fraco e uma fraca depressão Islandesa. O reduzido gradiente meridional de pressão resulta numa travessia de massas de ar menor e mais fraca e num sentido mais zonal, oeste-leste. Nesta fase da NAO há transporte de ar húmido para o Mediterrâneo e intrusão de ar frio, e por isso, condições de neve e anomalias positivas de precipitação na Europa do Sul devido ao desvio para sul da rota das depressões frontais ao longo do Atlântico. Como exemplos de baixos índices de NAO têm-se os períodos de Inverno/Primavera de 1917, 1936, 1963, e 1969 onde os ventos de oeste foram em média mais fracos sobre o Atlântico Norte com um respectivo Inverno Europeu mais frio que o normal.



**Figura 2.6 Fase negativa da NAO**  
(Fonte: <http://www.atmosphere.mpg.de/enid/2.html>)

Ambas as fases da NAO causam alterações na intensidade e localização da corrente de jacto e no percurso das tempestades do Atlântico Norte, bem como nos padrões normais de transporte zonal e meridional de calor e de humidade (Hurrell, 1995), que resultam em variações dos padrões de temperatura e de precipitação, estendendo-se geralmente desde o leste da América do Norte até à Europa central e ocidental (Walker e Bliss 1932, van Loon e Rogers 1978, Rogers e van Loon 1979).

A NAO tem uma variabilidade inter-sazonal (ou intranual), e inter-anual considerável, e é comum existirem períodos prolongados (vários meses) de ambas as fases, positiva e negativa. A NAO de Inverno também apresenta variabilidade multidecadal significativa (Hurrell 1995, Chelliah and Bell 2005).

### **Correlação entre os índices NAO e SPI**

A informação dada por índices de circulação geral atmosférica é uma boa ferramenta para o aumento da qualidade da previsão de secas, devido à sua forte correlação com a precipitação (índice *SPI*).

Testes de correlação feitos por Pires e Perdigão (2007), na península Ibérica, mostraram que a correlação entre o índice NAO e o índice *SPI* mensal atinge os -60% no inverno (DJF), o que tornam a NAO numa ferramenta fundamental para o melhoramento de previsões de seca (ver Figura 2.7).

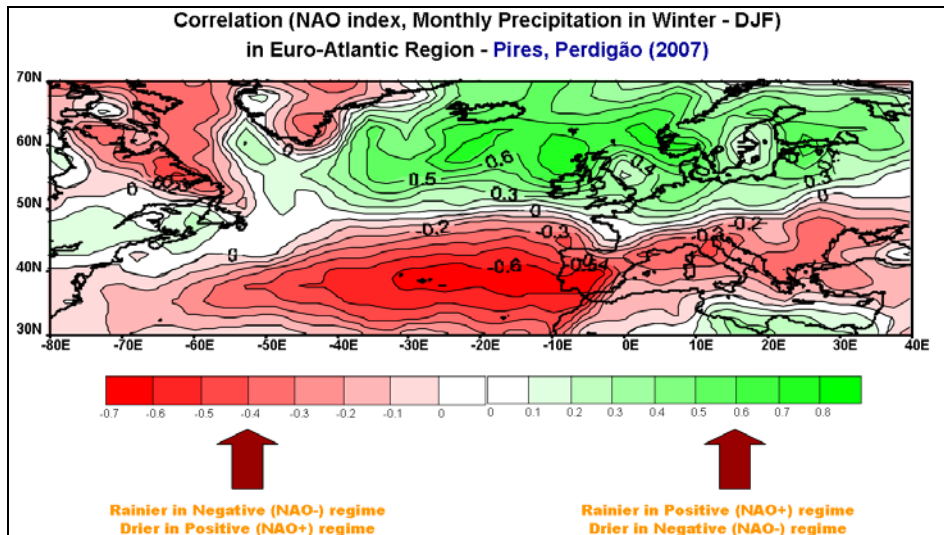


Figura 2.7 Mapa da correlação entre os índices *NAO* e *SPI* no inverno (DJF) na região Euro-Atlântica (Pires e Perdigão, 2007).

## 2.5.2 O índice AO

A Oscilação do Ártico é um modo de larga escala de variabilidade climática, também designada como modo anular do Hemisfério Norte. Este padrão de larga escala é caracterizado por uma oscilação da média zonal da componente zonal (oeste-leste) do vento, ou seja, uma oscilação da intensidade média da corrente geral de oeste à volta do Ártico à latitude aproximada de 55 graus Norte (ver Figura 2.8). Esta oscilação é um padrão no qual a pressão atmosférica polar e nas latitudes médias flutuam entre uma fase positiva e negativa. Na **fase positiva**, (ver Figura 2.8(a)), um anel de circulação de ventos fortes à volta do Pólo Norte confina ar mais frio ao longo das regiões polares. Nesta fase, as tempestades são mais para norte e trazem condições de meteorológicas mais húmidas para o Alasca, Escócia e Escandinávia, e condições mais secas para áreas como a Califórnia, Espanha e Médio Oriente. Esta fase está também, associada a uma intensificação do vórtice polar. Na **fase negativa**, (ver Figura 2.8(b)), as condições são opostas, o anel de ventos torna-se mais fraco e mais distorcido, permitindo mais facilmente, uma penetração no sentido sul de massas de ar árticas mais frias e um aumento de tempestades nas latitudes médias. A pressão sobre a região polar é maior que o normal e mais baixa que o normal à latitude de cerca de 45° Norte. Nesta fase, existe transporte de ar frio para centro-oeste dos Estados Unidos e Europa Ocidental, e tempestades e precipitação para o Mediterrâneo.

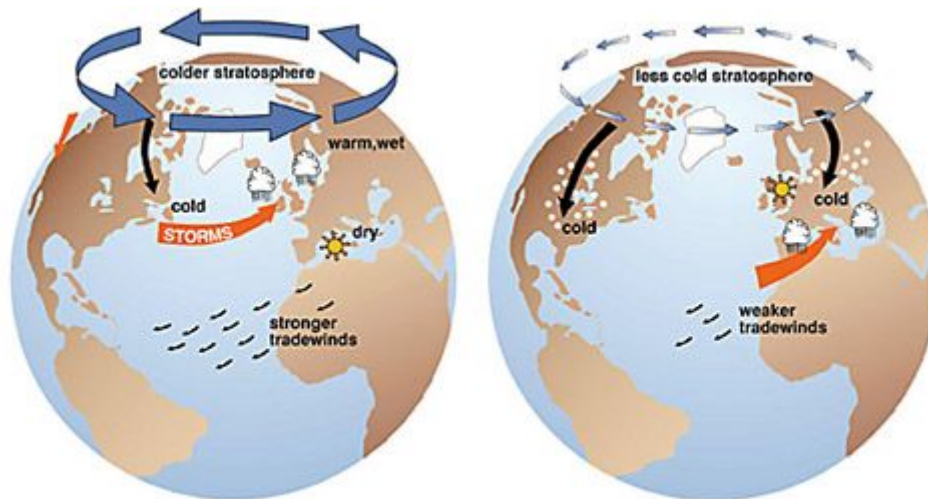


Figura 2.8 (a) e (b) Fases positiva e negativa da Oscilação do Ártico, respectivamente. (Wallace J., Universidade de Washington)

Recentemente, foi sugerido que a NAO poderia ser parte de um modo hemisférico de variabilidade. Thompson e Wallace (1998) e Thompson *et al.* (2000a,b) discutem o conceito de "Oscilação do Ártico", ou de "modo anular" do Hemisfério Norte, que a definiram como sendo a Função Empírica Ortogonal (EOF) principal da anomalia da pressão ao nível do mar (SLP) mensal média de inverno, do Hemisfério Norte. Wallace (2000) defende que a NAO e a AO representam um único fenómeno visto por dois paradigmas. A NAO pode ser vista como uma expressão regional da AO. A oscilação do Ártico pode ser interpretada como o principal modo de variabilidade da SLP de baixa frequência hemisférica de inverno, enquanto que a NAO é o modo principal na bacia Atlântica.

### 2.5.3 O índice SOI (Southern Oscillation Index)

O índice de Oscilação do Sul (SOI) é um índice estandardizado baseado nas diferenças de pressão observada ao nível médio do mar entre Tahiti no Pacífico e Darwin na Austrália. O SOI é uma medida das flutuações de larga escala da pressão atmosférica que ocorre entre o Pacífico Oeste e o Pacífico leste tropical, (isto é, o estado da Oscilação do Sul), durante os episódios de El Niño e La Niña, correspondentes respectivamente a anomalias negativas e positivas da temperatura de superfície do mar no Pacífico tropical a leste da América do Sul. Em geral, séries temporais do SOI, correspondem bastante bem a alterações na temperatura do oceano ao longo do Pacífico tropical. A **fase negativa** do SOI representa uma pressão atmosférica mais baixa que o normal em Tahiti e acima do normal em Darwin. Períodos prolongados de **valores negativos** de SOI coincidem com **águas anormalmente quentes** ao longo do Pacífico leste tropical típico de episódios de **El Niño**. Ao invés, períodos prolongados de **valores positivos** de SOI coincidem com **águas anormalmente frias** ao longo do Pacífico leste tropical típico de episódios de **La Niña**.

# PRESSURE DEPARTURES (mb)

## EL NIÑO

Jan-Mar 1998

## LA NIÑA

Jan-Mar 1989

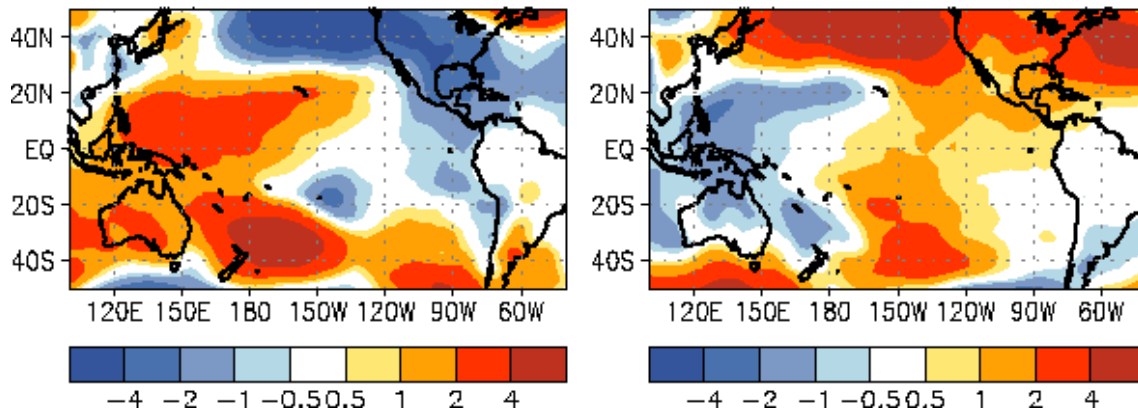


Figura 2.9 Anomalias da pressão (mb) nos casos de El Niño (à esquerda) e La Niña (à direita). (CPC, 2005)

Séries temporais do índice SOI e da temperatura da superfície do mar no Pacífico equatorial indicam que o ciclo do ENSO tem um período médio de cerca de quatro anos, embora os registos históricos tenham variações entre dois a sete anos. Os anos 80 e 90 foram caracterizados por um ciclo ENSO bastante activo, com 5 episódios de El Niño (1982/83, 1986/87, 1991-1993, 1994/95, e 1997/98) e 3 episódios de La Niña (1984/85, 1988/89, 1995/96). Este período foi também caracterizado por dois dos episódios de El Niño mais fortes do século (1982/83 e 1997/98), bem como dois períodos consecutivos de condições de El Niño durante 1991-1995 sem a ocorrência de um episódio frio. Existe também variabilidade considerável de uma década para outra, por exemplo, décadas durante as quais o ciclo foi relativamente inactivo e décadas em que este foi muito pronunciado (CPC, 2005).

Vicente-Serrano (2005) determinou o impacto de fases extremas do SOI nas secas da Península Ibérica. Usou 51 séries de precipitação (1910-2000) e classificou espacialmente esta área com base nos registos da precipitação mensal para identificar regiões homogéneas e analisar diferenças espaciais na influência destas fases extremas. Para cada região calculou um índice de seca SPI em escalas de tempo de 1, 3, 6 e 12 meses. Os anos de El Niño e La Niña foram identificados usando o índice SOI. Os seus resultados mostraram que as fases extremas do SOI afectam significativamente a ocorrência de secas na Península Ibérica. Foram também identificadas diferenças espaciais e temporais na escala de tempo usada. Vastas áreas na Península Ibérica são afectadas por valores significativamente negativos de SPI durante os meses finais de um ano de La Niña e nos meses iniciais do ano seguinte. Em contraste, outras áreas são afectadas por condições de seca durante os primeiros meses de um ano de El Niño, assim como, durante os Verões e Outonos do ano seguinte. As diferenças espaciais nas condições de seca durante as fases extremas do SOI são notórias, e o sinal da seca é mais consistente para o caso da La Niña que no caso dos anos de El Niño.

### 3. Modelação da seca

#### 3.1 Cadeias de Markov

##### 3.1.1 Algumas aplicações

No domínio da hidrologia, as **cadeias de Markov**<sup>3</sup> têm sido utilizadas em diversos estudos:

- ❖ Na Índia, em regiões afectadas pelas monções, é modelada a precipitação diária, para a qual é usada uma classificação binária, com dois estados, seco *versus* húmido, e na modelação de reservatórios (Banik *et al.*, 2002).
- ❖ Modelação de séries de precipitação (Banik *e tal.*, 2002), séries temporais de índices de seca (Lohani *et al.*, 1998; Steinemann, 2003) e séries temporais de precipitação, de escoamento e de índices de seca (Cancelliere e Salas, 2004).
- ❖ Ochola e Kerkides (2003) usaram-nas na predição de períodos secos (*dry spells*<sup>4</sup>) no Quênia.
- ❖ Aplicação a séries temporais do índice de Palmer, no estado de Virgínia, EUA, para a caracterização da possível evolução da severidade da seca e como ferramenta planeamento e gestão dos recursos hídricos (Lohani e Loganathan, 1997; Lohani *et al.*, 1998).

##### 3.1.2 Processo estocástico Markoviano

Uma cadeia de Markov é um caso particular de um processo estocástico. Um processo estocástico é uma colecção indexada de variáveis aleatórias,  $\{X_t : t \in T\}$ , as quais, quando o índice  $t$  percorre um dado conjunto  $T$ , tomam valores num dado conjunto  $E$ . Este conjunto  $E$  designa-se por espaço de estados e o conjunto  $T$  por conjunto de parâmetros. Se  $T$  é numerável o processo estocástico diz-se com parâmetros discretos, caso contrário com parâmetros contínuos. É usual considerar que  $t$  representa o tempo e considerar  $X_t$  como o estado ou posição do processo no tempo  $t$  (Çinlar, 1975). Por exemplo  $X_t$  pode representar o estado correspondente ao valor de um dado índice de seca referente ao mês  $t$  (e.g. classe SPI).

Os processos estocásticos têm interesse para descrever o comportamento de um sistema operando ao longo de um dado período de tempo. Um processo estocástico multivariado discreto geralmente obedece à seguinte estrutura:

- A situação actual do sistema pode ser classificada em qualquer uma das  $e$  categorias mutuamente exclusivas,  $1, \dots, e$ , designadas estados.
- A variável aleatória  $X_t$  representa o estado do sistema no momento  $t$ , e toma um dos valores possíveis  $1, \dots, e$ .

---

<sup>3</sup> As **cadeias de Markov** são assim chamadas em homenagem ao matemático *Andrei Andreyevich Markov*.

<sup>4</sup> Situações temporárias e pouco prolongadas de défice que não são consideradas como secas.

- O sistema é observado em determinados pontos no tempo,  $t = 0, 1, 2, \dots$ . Assim o processo estocástico  $\{X_t\} = \{X_0, X_1, X_2, \dots\}$  fornece uma representação do modo como a situação do sistema evolui no tempo.

Um processo estocástico desta natureza é um processo estocástico discreto no tempo, com um espaço de estados finito (Paulo, 2004).

Um processo estocástico Markoviano satisfaz a uma propriedade expressa em termos de probabilidades condicionais:

$$P(X_k | X_{k-1}, X_{k-2}, \dots, X_0) = P(X_k | X_{k-1}) \quad (k \geq 0)$$

Isto é a probabilidade do estado no instante actual  $k > 0$ , condicionada a toda a história passada do sistema:  $X_{k-1}, \dots, X_0$ , depende unicamente do estado no instante precedente  $k-1$ , ou por outras palavras, não mais que o estado no instante precedente  $k-1$  irá condicionar a probabilidade do estado no instante actual  $k$ . Os processos Markovianos podem modelar-se através de matrizes de transição que passaremos a explicar.

Um processo com as características apresentadas é uma cadeia de Markov simples ou de primeira ordem. Uma cadeia de Markov em que o valor do processo no momento  $k$  dependa dos valores em vários momentos imediatamente anteriores  $k-1, k-2, \dots$ , é uma cadeia de Markov de ordem superior (Paulo, 2004).

### 3.1.3 Matrizes e vectores estocásticos

**Matriz estocástica, matriz de probabilidade, ou matriz de transição** é usada para descrever a transições de uma rede de Markov. Encontra-se em teoria das probabilidades, estatística e álgebra linear, bem como nas ciências computacionais. Existem várias definições e tipos de matrizes estocásticas:

Uma **matriz estocástica direita** é uma matriz quadrada cujas linhas consistem em números reais não negativos, com soma 1.

Uma **matriz estocástica esquerda** é uma matriz quadrada cujas colunas consistem em números reais não negativos cuja soma é 1.

Uma **matriz duplamente estocástica** é uma matriz quadrada cujas entradas são números reais não negativos e todas as linhas e colunas somam 1.

Da mesma forma pode-se definir um **vector estocástico ou vector de probabilidades** como sendo, um vector cujos elementos consistem em números reais não negativos cuja soma é 1. Assim, cada linha (ou coluna, conforme o caso) de uma matriz estocástica é um vector estocástico.

#### Definições e propriedades

Uma matriz estocástica descreve uma cadeia de Markov  $X_t$  sobre um espaço de estados finito  $S$ .

Se a probabilidade de transição do estado  $i$  para o estado  $j$  no instante subsequente é a probabilidade condicional:  $\Pr(j|i) = P_{i,j}$ , (leia-se probabilidade de  $j$ , dado  $i$ ), a matriz estocástica  $P$  é dada  $P_{i,j}$  como o elemento da  $i$ -ésima linha e da  $j$ -ésima coluna:

$$P = \begin{pmatrix} p_{1,1} & p_{1,2} & \cdots & p_{1,j} & \cdots \\ p_{2,1} & p_{2,2} & \cdots & p_{2,j} & \cdots \\ \vdots & \vdots & \ddots & \vdots & \ddots \\ p_{i,1} & p_{i,2} & \cdots & p_{i,j} & \cdots \\ \vdots & \vdots & \ddots & \vdots & \ddots \end{pmatrix} \quad (3.1)$$

Como a probabilidade de transição do estado  $i$  para outro estado tem de ser 1, tem-se que esta matriz é uma matriz estocástica direita, e então:

$$\sum_j P_{i,j} = 1 \quad (3.2)$$

A probabilidade de transição de  $i$  para  $j$  em dois passos é dado pelo  $(i, j)$ -ésimo elemento do produto de matrizes  $PP = P^2$ :

$$(P^2)_{i,j} \quad (3.3)$$

Em geral, as probabilidades de transição de um qualquer estado para outro numa rede de Markov, após  $k$  instantes, é dada pelo produto de  $k$  matrizes:  $P^k$ .

### 3.1.4 Propriedades da Cadeia de Markov

Uma cadeia ou rede de Markov é um processo estocástico  $X$ , desde que a condição seguinte se verifique:

$$P\{X_{t+1} = j | X_0, X_1, \dots, X_t\} = P\{X_{t+1} = j | X_t = i\} \quad \forall i, j \in S, t \in T \quad (3.4)$$

Assim sendo, uma cadeia de Markov é uma sequencia temporal de variáveis aleatórias  $X_t$ , tal que, a cada qualquer instante de tempo  $t$ ,  $X_{t+1}$  é condicionalmente independente de  $X_0, X_1, X_2, \dots, X_{t-1}$ , dado  $X_t$ , ou seja, a probabilidade de  $X_{t+1}$  tomar um valor particular  $j$  depende apenas do passado apenas no seu valor mais recente  $X_t$ .

Um processo com as características apresentadas é uma cadeia de Markov simples ou de primeira ordem. Uma cadeia de Markov em que o valor do processo no momento  $t$  dependa dos valores em vários momentos imediatamente anteriores  $t-1, t-2, \dots$ , é uma cadeia de Markov de ordem superior (Paulo, 2004).

Uma rede de Markov é caracterizada por um conjunto de estados,  $S$ , e pela probabilidade de transição  $p_{ij}$ . A probabilidade de transição  $p_{ij}$  é a probabilidade da cadeia de Markov estar num próximo instante no estado  $j$ , dado que no instante presente está no estado  $i$ . Quando se verifica a probabilidade condicional de transição entre estados,  $P\{X_{t+1} = j | X_t = i\} = p_{ij}$ , a cadeia de Markov é independente de  $t$ , dependendo apenas dos estados  $i$  e  $j$ , e o mecanismo de transição designa-se **homogéneo** ou estacionário no tempo (Çınlar, 1975). No caso da cadeia de Markov ser **não homogénea** ou não estacionária, a probabilidade

condicional é descrita por  $P\{X_{t+1} = j | X_t = i\} = p_{ij}^{t,t+1}$  e além de depender dos estados  $i$  e  $j$ , depende também de  $t$ .

Uma rede de Markov **homogénea** tem, portanto, a seguinte propriedade:

$$P\{X_{t+1} = j | X_t = i\} = P\{X_t = j | X_{t-1} = i\} \quad \forall i, j \in S, t \in T \quad (3.5)$$

Na modelação de séries de índices de seca a formulação **não-homogénea**, é geralmente mais adequada que a **homogénea**, e tem-se em conta o mês em que a transição de um determinado estado  $i$  para um determinado estado  $j$  se dá. Nesse caso a rede de Markov é ciclo-estacionária ou seja, a matriz de transição admite uma periodicidade que pode ser o ciclo anual das estações ou dos meses.

Assim, no caso **não-homogéneo**, a probabilidade condicional  $P\{X_{t+1} = j | X_t = i\}$  depende não só dos estados  $i$  e  $j$  mas também de  $t$ .

$$p_{ij}^{(t,t+1)} = P\{X_{t+1} = j | X_t = i\}, \quad \forall i, j \in S \quad (3.6)$$

Consideram-se, então, doze matrizes mensais de probabilidades de transição,  $P^{(Mês)}$  que descrevem as probabilidades de transição entre estados dado o mês inicial. As matrizes  $P^{(Jan)}$ ,  $P^{(Fev)}$ , ...,  $P^{(Dez)}$  contêm as probabilidades de transição entre os meses de Janeiro e Fevereiro, Fevereiro e Março, ..., Dezembro e Janeiro, respectivamente. Considera-se ainda, que estas probabilidades só dependem do mês em questão e não do ano (Paulo, 2004).

Considerando uma série temporal mensal de valores de índices de seca, SPI, faz-se corresponder estes valores a classes de severidade de seca, por exemplo os da Tabela 2.2, obtendo-se assim uma série temporal discreta de valores de severidade de seca.

Numa cadeia de Markov de primeira ordem, homogénea no tempo, a matriz das probabilidades de transição  $P = [p_{ij}] = P\{X_{t+1} = j | X_t = i\}$  é estimada a partir dos dados da amostra, contando o número de vezes que o SPI passa do estado  $i$  para o estado  $j$ ,  $n_{ij}$ :

$$\Pr(j | i) = \hat{p}_{ij} = \frac{n_{ij}}{\sum_j n_{ij}} \quad (3.7)$$

A dimensão da amostra e o número de estados influencia a precisão das estimativas; o número de parâmetros do modelo depende do número de estados.

A matriz de transição de probabilidades  $P = [p_{ij}]$  verifica ainda, a propriedade de normalização das probabilidades (3.2).

As estimativas das probabilidades de transição da matriz  $P$  calculadas pela expressão (3.7) devem ser obtidas a partir de amostras grandes, para que estas estimativas sejam de confiança. Por outro lado, quanto maior é o número de estados em que o processo está dividido menos rigorosas são as estimativas (Haan, 1997).

A probabilidade da cadeia se encontrar num estado  $i_m$  num instante  $t+m$ , tendo passado sucessivamente pelos estados  $i_0, i_1, \dots, i_{m-1}$  e sabendo que no instante  $t$  se encontrava no estado  $i_0$  é obtida pela seguinte expressão resultante do produto das sucessivas probabilidades de transição entre estados:

$$P\{X_{t+1} = i_1, \dots, X_m = i_m | X_t = i_0\} = p_{i_0 i_1} p_{i_1 i_2} \dots p_{i_{m-1} i_m} \quad (3.8)$$

Se a probabilidade do estado inicial for  $P\{X_0 = i\} = \pi_0(i)$  então a distribuição anterior será definida por:

$$P\{X_0 = i_0, X_1 = i_1, \dots, X_m = i_m\} = \pi_0(i_0) p_{i_0 i_1} p_{i_1 i_2} \dots p_{i_{m-1} i_m} \quad (3.9)$$

A probabilidade de que a cadeia se mova de um estado  $i$  para um estado  $j$  em  $m$  passos, corresponde ao elemento  $(i, j)$  de  $P^m$ , é dada por:

$$P\{X_{t+m} = j | X_t = i\} = (P^m)_{i,j} \quad \forall i, j \in S, t \in T \quad (3.10)$$

Verifica-se ainda que:

$$P^{m+n} = P^m P^n \quad \forall m, n \in \mathbb{N} \quad (3.11)$$

Portanto, se um processo  $X$  que comece num estado  $i$ , para que chegue a um estado  $j$  depois de  $m+n$  transições, deverá estar num qualquer estado intermédio  $k$  a seguir ao estado  $m$  e transitar do estado  $k$  para o estado  $j$  nos  $n$  passos que restam, o que é traduzido pela equação de Chapman-Kolmogorov (Çinlar, 1975; Isaacson e Madsen, 1976):

$$p_{ij}^{m+n} = \sum_k p_{ik}^m p_{kj}^n \quad (3.12)$$

### 3.1.4.1 Vector de equilíbrio – “Steady State Vector”

A longo prazo, no caso homogéneo, a probabilidade da classe de seca é independente do estado inicial da rede de Markov, isto porque a probabilidade de transição de  $i$  para  $j$  em  $n$  passos,  $p_{ij}^n$ , sendo  $n$  suficientemente grande, pode convergir para valores de equilíbrio que não dependem do estado inicial.

A convergência é assegurada na modelação de séries temporais de estados de severidade de seca, uma vez que as cadeias de Markov são formadas por uma só classe recorrente, com todos os estados acessíveis entre si e podendo ser visitados um número infinito de vezes (Paulo, 2004).

Numa cadeia de Markov com uma única classe recorrente, não periódica, a probabilidade de equilíbrio do estado  $j$ ,  $\pi_j$ , é independente do estado inicial da cadeia (Isaacson e Madsen, 1976):

$$\pi_j = \lim_{t \rightarrow \infty} \Pr(X_t = j) = \lim_{t \rightarrow \infty, \forall i} \Pr(X_t = j | X_0 = i), \quad i \wedge j \in S \quad (3.13)$$

Os valores de  $\pi_j$ , constituem uma solução do sistema de equações:

$$\begin{cases} \pi_j = \sum_{k \in S} \pi_k p_{kj}, & j \in S \\ \sum_{j \in S} \pi_j = 1, \forall j \in S \end{cases} \quad (3.14)$$

Em que  $\pi_j$  é o vector de equilíbrio,  $p_{kj}$  é a probabilidade de transição de um qualquer estado  $k$  do conjunto  $S$ , e  $\pi_k$  é o vector que representa qualquer estado  $k$  do conjunto  $S$ . Esse vector  $\Pi$  está associado a valores próprios unitários de  $P$ . Prova-se que numa matriz de transição existe sempre um estado estacionário, no entanto a unicidade não é garantida. Bastará para tal considerar o caso trivial em que  $P$  é a matriz identidade.

O vector de equilíbrio,  $\Pi$ , também pode ser obtido pela potenciação sucessiva da matriz de probabilidades de transição. Para  $n$  suficientemente grande, as linhas da matriz  $P^n$  (matriz de equilíbrio) são todas iguais entre si e correspondem ao vector  $\Pi$  (Paulo, 2004).

No caso ***não-homogéneo***, tal como no caso ***homogéneo***, admite-se que a longo prazo a probabilidade de transição do estado  $i$  no mês  $t$  para o estado  $j$  no mês  $t+n$ , converge para valores de equilíbrio, não dependendo assim do estado inicial. Deste modo existem doze vectores de equilíbrio, referentes a cada mês do ano e transição é de natureza cíclica. Ex: as probabilidades de transição entre os meses 13 e 14 são iguais às dos meses 1 e 2.

Assim, adoptando a nomenclatura usada por Paulo (2004), no mês de Maio, a matriz de equilíbrio será:

$$\Theta^{(Mai, \infty)} = \left\{ P^{(Mai)} P^{(Jun)} \dots P^{(Mar)} P^{(Abr)} \right\} \cdot \left\{ P^{(Mai)} P^{(Jun)} \dots P^{(Mar)} P^{(Abr)} \right\} \dots \quad (3.15)$$

sendo o vector de probabilidades estacionárias  $\pi^{(Mai)}$  uma qualquer linha de  $\Theta^{(Mai, \infty)}$ .

Da mesma forma, para o mês de Junho será:

$$\Theta^{(Jun, \infty)} = \left\{ P^{(Jun)} P^{(Jul)} \dots P^{(Abr)} P^{(Mai)} \right\} \cdot \left\{ P^{(Jun)} P^{(Jul)} \dots P^{(Abr)} P^{(Mai)} \right\} \dots \quad (3.16)$$

É possível verificar as seguintes propriedades das matrizes (e vectores) de equilíbrio no caso ***não-homogéneo*** (Lohani *et al.*, 1998):

$$\Theta^{(Jun, \infty)} = \Theta^{(Mai, \infty)} P^{(Mai)} \quad (3.17)$$

Com,

$$\pi^{(Jun)} = \pi^{(Mai)} P^{(Mai)} \quad (3.18)$$

Para os restantes meses o processo é análogo:

$$\pi^{(Jul)} = \pi^{(Jun)} P^{(Jun)} \quad (3.19)$$

...

$$\pi^{(Mai)} = \pi^{(Abr)} P^{(Abr)} \quad (3.20)$$

### 3.1.4.2 Tempo médio de permanência ininterrupta no estado $i$

Um processo permanece num qualquer estado  $i$  da cadeia durante  $m$  meses, sem migrar para outro estado quando  $\{i = X_1 = X_2 = \dots = X_{m-1}; X_m \neq i\}$ .

Se  $m$  representar o número de meses (ou períodos de tempo) de permanência ininterrupta num determinado estado  $i$ , tem-se então para  $m = 1$ ,  $m = 2$ :

$$P(m = 1 | X_0 = i) = P(X_1 \neq i | X_0 = i) \quad (3.21)$$

$$P(m = 2 | X_0 = i) = P(X_2 \neq i | X_0 = i, X_1 = i) \dots \quad (3.22)$$

No caso **homogéneo**, a probabilidade de permanência ininterrupta, durante  $m$  meses, num determinado estado  $i$  da cadeia é calculada pelo produto das probabilidades das transições sucessivas entre o estado  $X_0$  e o estado  $X_m$ , e é dada por:

$$\begin{aligned} & P(X_1 = i | X_0 = i) P(X_2 = i | X_1 = i) \dots P(X_{m-1} = i | X_{m-2} = i) P(X_m \neq i | X_{m-1} = i) = \\ & = P_{ii}^{m-1} (1 - P_{ii}) = P(m | X_0 = i) \end{aligned} \quad (3.23)$$

Sabendo as probabilidades de permanência ininterrupta em cada classe de severidade  $i$  durante  $m$  meses ( $m = 1, 2, \dots, k \dots$ ), o tempo de residência esperado em cada classe  $i$ ,  $E(T_i | X_0)$ , é calculado por (Paulo, 2004):

$$E(T_i | X_0) = \sum_k k P(m = k | X_0 = i) \quad (3.24)$$

No caso **não-homogéneo** as probabilidades de transição são definidas para cada mês, e desta forma, a probabilidade de permanência ininterrupta da cadeia num estado  $i$  durante  $m$  meses depende do mês inicial  $t_0$  (com  $t_0 = 1, 2, \dots, 12$ ):

$$P(m = 1 | X_{t_0} = i) = P(X_{t_0+1} \neq i | X_{t_0} = i) \quad (3.25)$$

$$P(m = 2 | X_{t_0} = i) = P(X_{t_0+2} \neq i | X_{t_0} = i, X_{t_0+1} = i) \dots \quad (3.26)$$

Por exemplo, a probabilidade de permanência ininterrupta de 3 meses na classe  $i$ , sendo Maio o mês inicial, será:

$$\begin{aligned} P(m=3 | X_{Mai} = i) &= P(X_{Ago} \neq i | X_{Mai} = i, X_{Jun} = i, X_{Jul} = i) = \\ &= p_{ii}^{(Mai)} p_{ii}^{(Jun)} (1 - p_{ii}^{(Jul)}) \end{aligned} \quad (3.27)$$

O valor esperado para o tempo de residência ininterrupta na classe  $i$ , dado o mês inicial  $t_0$ ,  $E(T_i | X_{t_0})$ , é dado por:

$$E(T_i | X_{t_0}) = \sum_t t P(m=t | X_{t_0} = i) \quad (3.28)$$

### 3.1.4.3 Tempo médio para a primeira visita ao estado $j$

O tempo esperado,  $t_{ij}$ , para atingir pela primeira vez o estado  $j$  partindo do estado  $i$ , ou seja, o tempo esperado para a primeira visita ao estado  $j$  é a solução única do sistema de equações lineares (Paulo, 2004), no caso **homogêneo**:

$$t_{ij} = 1 + \sum_{k \neq j} p_{ik} t_{kj} \quad \forall i, j \in S \quad (3.29)$$

onde  $p_{ik}$  é a probabilidade de transição entre um estado  $i$  e o estado  $k$ .

Um caso particular do tempo esperado para a primeira passagem é o **tempo médio de recorrência** do estado  $i$ ,  $t_{ii}$ .

$$t_{ii} = \frac{1}{\pi_i} \quad (3.30)$$

onde  $\pi_i$  é a probabilidade de equilíbrio do estado  $i$ , Equação (3.13).

Para o caso **não-homogêneo**, o tempo esperado, para atingir pela primeira vez o estado  $j$  partindo do estado  $i$ , no mês inicial  $t_0$ ,  $t_{ij}^{(t_0)}$ , é solução única do sistema de equações lineares (Paulo, 2004):

$$t_{ij}^{(t_0)} = 1 + \sum_{k \neq j} p_{ik}^{(t_0, t_0+1)} t_{kj}^{(t_0+1)} \quad \forall i, j \in S \quad (3.31)$$

$p_{ik}^{(t_0, t_0+1)}$  é a probabilidade de transição entre o estado  $i$ , no mês  $t_0$  e o estado  $k$ , no mês seguinte,  $t_0 + 1$ .

O **tempo médio de recorrência**,  $t_{ii}^{(t_0)}$ , do estado  $i$ , no mês  $t_0$ , pode-se escrever como:

$$t_{ii}^{(t_0)} = \frac{1}{\pi_i^{(t_0)}} \quad (3.32)$$

onde  $\pi_i^{(t_0)}$  é a probabilidade de equilíbrio do estado  $i$ , no mês  $t_0$ .

### 3.1.4.4 Previsões a curto prazo

As cadeias de Markov não são modelos adequados para previsões a longo prazo, mas antes, para previsões de curto prazo, com um, dois e três meses de antecedência (Paulo, 2004).

Para determinar as previsões a curto prazo, no caso **homogéneo**, há que avaliar a probabilidade de todos os possíveis caminhos de um estado presente  $i$  num instante de tempo  $t$  ( $X_t = i$ ), para um estado futuro  $j$  a ocorrer  $m$  meses depois ( $X_{t+m} = j$ ). Assim, na formulação homogénea, para estimar a classe mais provável no espaço de tempo de 1 a 3 meses seguintes, as seguintes probabilidades tem de ser calculadas:

$$P(X_{t+1} = j | X_t = i) \quad (3.33)$$

$$P(X_{t+2} = j | X_t = i) = \sum_{\forall k \in S} P(X_{t+2} = j | X_{t+1} = k) \cdot P(X_{t+1} = k | X_t = i) \quad (3.34)$$

$$P(X_{t+3} = j | X_t = i) = \sum_{\forall l \in S} \sum_{\forall k \in S} P(X_{t+3} = j | X_{t+2} = l) \cdot P(X_{t+2} = l | X_{t+1} = k) \cdot P(X_{t+1} = k | X_t = i) \quad (3.35)$$

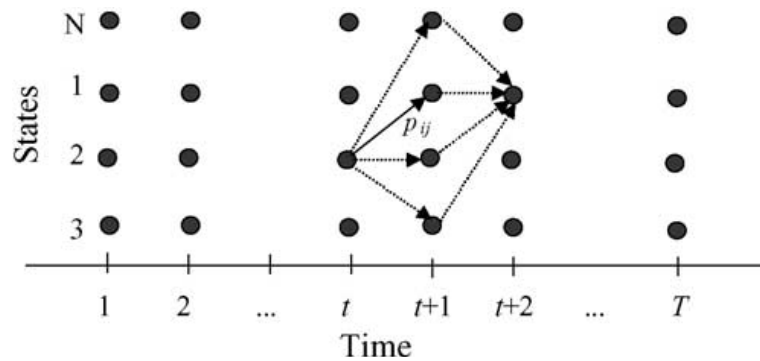


Figura 3.1 Caminhos possíveis de uma rede de Markov com 4 estados de um estado  $i=2$ , num instante  $t$ , para um estado  $j=1$ , num instante  $t+2$  (Paulo *et al.*, 2005).

A classe de seca mais provável num horizonte de previsão de  $m$  meses é aquela que apresentar o valor máximo de  $P(X_{t+m} = j | X_t = i)$ .

No caso **não-homogéneo**, as três expressões anteriores, (3.33), (3.34) e (3.35) podem-se escrever da seguinte maneira:

$$P(X_{t_0+1} = j | X_{t_0} = i) = p_{ij}^{(t_0)} \quad (3.36)$$

$$P(X_{t_0+2} = j | X_{t_0} = i) = \sum_k p_{kj}^{(t_0+1)} \cdot p_{ik}^{(t_0)} \quad (3.37)$$

$$P(X_{t_0+3} = j | X_{t_0} = i) = \sum_l \sum_k p_{lj}^{(t_0+2)} \cdot p_{kl}^{(t_0+1)} \cdot p_{ik}^{(t_0)} \quad (3.38)$$

para  $i \in S$ ,  $j, k, l = 1, \dots, n$ , onde  $n$  = número de classes de seca.

A classe de seca mais provável num horizonte de previsão de  $m$  meses no caso **não-homogéneo**, dado o mês inicial  $t_0$ , corresponde ao valor de  $j$  para o qual a probabilidade  $P(X_{t_0+m} = j | X_{t_0} = i)$  é máxima.

### 3.2 Aplicação dos índices de CGA às cadeias de Markov

A variável aleatória em estudo é o índice  $SPI$  discretizado em classes de categoria de seca. Para facilitar, a variável aleatória anteriormente referida por  $X$ , é agora, substituída pela variável aleatória  $SPI$ . Uma variável discreta é uma variável que assume apenas um valor num conjunto finito de hipóteses. São usadas as designações anteriormente introduzidas para os dois modelos de Markov, o modelo MARKOV-SPI e o modelo MARKOV-SPI-ÍNDICE.

#### MODELO MARKOV-SPI

Como foi já referido na secção 3.1.4, e adoptando a nova nomenclatura  $X \equiv SPI$ , a transição de um estado  $SPI_t$  para  $SPI_{t+1}$  faz-se pela multiplicação do vector estado  $SPI$  pela matriz  $M$  de transição (entre as várias classes de categoria de seca determinadas pelo valor do  $SPI$  discretizado), da seguinte forma, no **caso homogéneo**:

$$SPI_{t+1} = SPI_t M \quad (3.39)$$

Em que a matriz  $M$  é uma matriz  $(n \times n)$ ,  $P_{(n \times n)}$ , em que  $n$  é o número de classes de severidade de seca:

$$M = \begin{pmatrix} \Pr(SPI_{t+1} = 1 | SPI_t = 1) & \Pr(SPI_{t+1} = 2 | SPI_t = 1) & \cdots & \Pr(SPI_{t+1} = n | SPI_t = 1) \\ \Pr(SPI_{t+1} = 1 | SPI_t = 2) & \Pr(SPI_{t+1} = 2 | SPI_t = 2) & \cdots & \Pr(SPI_{t+1} = n | SPI_t = 2) \\ \vdots & \vdots & \ddots & \vdots \\ \Pr(SPI_{t+1} = 1 | SPI_t = n) & \Pr(SPI_{t+1} = 2 | SPI_t = n) & \cdots & \Pr(SPI_{t+1} = n | SPI_t = n) \end{pmatrix} \quad (3.40)$$

Ou seja, para cada entrada  $j$  do vector  $SPI$  no instante de tempo  $t+1$ :

$$SPI_j = \sum_{k \in S} SPI_k M_{kj}, \quad j \in S \quad (3.41)$$

O vector probabilístico  $SPI$  e a matriz de transição  $M$ , obedecem ainda às seguintes propriedades (Equação (3.2)):

$$\sum_{j=1}^N SPI_j = 1 \quad (3.42)$$

e

$$\sum_{j=1}^N M_{ij} = 1, \quad \forall i \in S \quad (3.43)$$

#### MODELO MARKOV-SPI-ÍNDICE

No caso da inclusão do índice de CGA no modelo de Markov, a previsão do índice *SPI* fica dependente do regime de CGA em questão, uma vez que se optou por considerar apenas dois estados, positivo e negativo, do índice de CGA. Considerando a cadeia de Markov num instante  $t$ , para um instante de tempo seguinte  $t+1$ , a transição de uma classe de *SPI* para outra, é função de duas matrizes de transição das classes do *SPI*, uma para todas as transições entre classes, quando o índice de CGA é positivo e outra para todas as transições entre classes quando o mesmo é negativo. Deste modo, é possível ter em conta o impacto que uma fase positiva ou negativa de um determinado regime de CGA tem na precipitação, (e conseqüentemente no índice de seca *SPI*). A título de exemplo, sabe-se por estudos anteriores, que a fase positiva da *NAO* favorece a escassez de precipitação em Portugal e portanto a seca, e, ao invés, a fase negativa da *NAO* favorece maior abundância de precipitação.

Para descrever este modelo que incorpora o índice de CGA e as suas transições, é mais simples e adequado particularizar as equações para um índice de CGA específico. Assim sendo, opta-se pelo índice da *NAO*, padrão de teleconexão com mais influência na precipitação em Portugal. Para este modelo há que ter em conta não uma, mas duas variáveis aleatórias referentes ao índice de seca e ao índice de CGA. Assim, estas variáveis aleatórias são *SPI* e *NAO*, respectivamente. A matriz de transição do índice de CGA será a matriz  $P$ , matriz esta que contém as probabilidades condicionais de transição entre todos os estados (positivo e negativo), do índice de CGA. A matriz  $P$  é assim, uma matriz  $(2 \times 2)$ . No caso da variável *SPI*, as matrizes  $M_+$  e/ou  $M_-$  representam as probabilidades condicionais das transições entre as classes de severidade de seca, entre  $(t, t+1)$  quando o regime em  $(t+1)$  de *NAO* é positivo e/ou negativo.

Seguem-se as expressões para o cálculo do índice *SPI* num instante seguinte,  $t+1$ :

O vector probabilidade *NAO* no instante seguinte,  $t+1$ , é calculado de forma análoga ao vector probabilidade *SPI* no modelo MARKOV-SPI:

$$NAO_{t+1} = NAO_t P \quad (3.44)$$

Como já foi referido, considerou-se apenas os estados positivo e negativo do índice *NAO*, e assim, a matriz  $P$  é então, uma matriz  $(2 \times 2)$ ,  $P_{(2 \times 2)}$ :

$$P = \begin{pmatrix} \Pr(NAO_+^{(t+1)} | NAO_+^{(t)}) & \Pr(NAO_-^{(t+1)} | NAO_+^{(t)}) \\ \Pr(NAO_+^{(t+1)} | NAO_-^{(t)}) & \Pr(NAO_-^{(t+1)} | NAO_-^{(t)}) \end{pmatrix} \quad (3.45)$$

Num determinado momento  $t+1$  da cadeia o vector *NAO* tem uma configuração deste tipo:

$$\Pr(NAO_{t+1}) = [a \quad b] \quad (3.46)$$

Com,

$$a = \Pr(NAO_+^{(t+1)}), \quad b = \Pr(NAO_-^{(t+1)}), \quad a, b \in \mathbb{R} \quad (3.47)$$

e em que,

$$a + b = \Pr(NAO_+^{(t+1)}) + \Pr(NAO_-^{(t+1)}) = 1 \quad (3.48)$$

Aplicando o teorema de Bayes e tomando a variável discreta X associada à classe SPI de seca escreve-se:

$$\begin{aligned} \Pr(X_{t+1} | X_t, NAO^t) &= \Pr(X_{t+1} | X_t, NAO_+^{t+1}, NAO^t) \Pr(NAO_+^{t+1} | X_t, NAO^t) + \\ &+ \Pr(X_{t+1} | X_t, NAO_-^{t+1}, NAO^t) \Pr(NAO_-^{t+1} | X_t, NAO^t) \end{aligned} \quad (3.49)$$

Admitindo que o índice da NAO satisfaz a um processo de Markov e que o índice NAO em t+1 depende unicamente do índice NAO no instante anterior t, tem-se:

$$\Pr(NAO^{t+1} | X_t, NAO^t) = \Pr(NAO^{t+1} | NAO^t) \quad (3.50)$$

o que permite escrever o estado SPI em t+1 na forma:

$$SPI_{t+1} = SPI_t [aM_+ + bM_-] \quad (3.51)$$

onde a, b evoluem de acordo com a cadeia de Markov que governa o índice da NAO e em que  $M_+$  e  $M_-$  são as matrizes de transição quando o regime de NAO é positivo ou negativo, (+/-), respectivamente, definidas analogamente à matriz da expressão (3.40):

$$M_{+/-} = \begin{pmatrix} \Pr(SPI_{t+1} = 1 | SPI_t = 1, NAO_{+/-}^{(t+1)}) & \cdots & \Pr(SPI_{t+1} = n | SPI_t = 1, NAO_{+/-}^{(t+1)}) \\ \vdots & \ddots & \vdots \\ \Pr(SPI_{t+1} = 1 | SPI_t = n, NAO_{+/-}^{(t+1)}) & \cdots & \Pr(SPI_{t+1} = n | SPI_t = n, NAO_{+/-}^{(t+1)}) \end{pmatrix} \quad (3.52)$$

Portanto, como expressão final fica,

$$\boxed{SPI_{t+1} = SPI_t [\Pr(NAO_+^{(t+1)})M_+ + \Pr(NAO_-^{(t+1)})M_-]} \quad (3.53)$$

Deste modo, a transição de um estado  $SPI_t$  para  $SPI_{t+1}$  faz-se pela multiplicação entre o vector de estado SPI no mês t, e uma matriz que é resultante da média pesada das duas matrizes de transição designadas por  $M_+$  e  $M_-$ , ou seja, a matriz de transição de um estado i de SPI para um estado j de SPI quando o índice de CGA presente é positivo ou negativo, respectivamente.

### 3.2.1 Vector de equilíbrio – “Steady State Vector”

Tal como foi referido na secção 3.1.4.1, para o caso do **modelo MARKOV-SPI e versão homogénea**, a longo prazo, a probabilidade da classe de seca é independente do estado inicial da rede de Markov, isto porque a probabilidade de transição de i para j em n passos,  $p_{ij}^n$ , sendo n suficientemente grande, pode convergir para valores de equilíbrio que não dependem do estado inicial.

Assim o vector  $\pi_j$  da expressão (3.13) é equivalente a qualquer uma das linhas da matriz  $M^n$  obtida pela potenciação sucessiva da matriz de probabilidades de transição  $M$  :

$$\pi_j = \lim_{n \rightarrow \infty} (M^n)_j \quad (3.54)$$

Com,

$$\lim_{n \rightarrow \infty} M^n = M.M.M.M...M \quad (3.55)$$

Para o caso do **modelo MARKOV-SPI-NAO**, há que se achar primeiro o vector de equilíbrio da *NAO* , através da potenciação sucessiva da matriz  $P$  :

$$\lim_{n \rightarrow \infty} P^{(n)} = P.P.P.P...P = \begin{bmatrix} c & d \\ c & d \end{bmatrix} \quad (3.56)$$

Em que,

$$c = \Pr\left(\lim_{n \rightarrow \infty} NAO_+^{(n)}\right), \quad d = \Pr\left(\lim_{n \rightarrow \infty} NAO_-^{(n)}\right), \quad c, d \in \mathbb{R} \quad (3.57)$$

Com  $c + d = 1$ .

Finalmente o vector de equilíbrio,  $\varphi_j$ , corresponde a cada linha da matriz resultante da seguinte expressão:

$$\varphi_j = \lim_{n \rightarrow \infty} [cM_+ + dM_-]^n = [cM_+ + dM_-] \cdot [cM_+ + dM_-] \dots [cM_+ + dM_-] \quad (3.58)$$

O vector de equilíbrio corresponde às linhas da matriz resultante da potenciação de  $cM_+ + dM_-$ . Esta matriz é resultante da média pesada das duas matrizes de transição  $M_+$  e  $M_-$ .

### 3.2.2 Tempo médio de permanência ininterrupta no estado $i$

Tal como descrito na secção 3.1.4.2, no caso do modelo MARKOV-SPI, a probabilidade de permanência ininterrupta durante  $m$  meses num determinado estado  $i$  da cadeia, e o tempo de residência esperado em cada classe  $i$ , nos **casos homogéneos**, são dados pelas expressões, (3.23) e (3.24), respectivamente.

No caso do modelo MARKOV-SPI-NAO, é necessário distinguir os casos nos dois regimes diferentes, isto é, *NAO* positiva e *NAO* negativa. Assim, determina-se a probabilidade de permanência ininterrupta durante  $m$  meses num determinado estado  $i$  da cadeia, e o tempo de residência esperado em cada classe  $i$ , em cada um dos casos supracitados.

A probabilidade de permanência ininterrupta durante  $m$  meses num determinado estado  $i$  da cadeia, aquando de uma situação de *NAO* positiva:

$$\begin{aligned}
& \Pr(SPI_1 = i | SPI_0 = i, NAO_+^{(0)}) \Pr(SPI_2 = i | SPI_1 = i, NAO_+^{(1)}) \dots \\
& \Pr(SPI_{m-1} = i | SPI_{m-2} = i, NAO_+^{(m-2)}) \Pr(SPI_m \neq i | SPI_{m-1} = i, NAO_+^{(m-1)}) = \quad (3.59) \\
& = M_{+ii}^{m-1} (1 - M_{+ii}) = \Pr_{NAO_+}(m | SPI_0 = i)
\end{aligned}$$

E analogamente, numa situação de *NAO* negativa:

$$\begin{aligned}
& \Pr(SPI_1 = i | SPI_0 = i, NAO_-^{(0)}) \Pr(SPI_2 = i | SPI_1 = i, NAO_-^{(1)}) \dots \\
& \Pr(SPI_{m-1} = i | SPI_{m-2} = i, NAO_-^{(m-2)}) \Pr(SPI_m \neq i | SPI_{m-1} = i, NAO_-^{(m-1)}) = \quad (3.60) \\
& = M_{-ii}^{m-1} (1 - M_{-ii}) = \Pr_{NAO_-}(m | SPI_0 = i)
\end{aligned}$$

E o tempo de residência esperado em cada classe *i*, será, nos casos de *NAO* positiva e negativa, respectivamente:

$$E_{NAO_+}(T_i | SPI_0) = \sum_k k P_{NAO_+}(m = k | SPI_0 = i) \quad (3.61)$$

$$E_{NAO_-}(T_i | SPI_0) = \sum_k k P_{NAO_-}(m = k | SPI_0 = i) \quad (3.62)$$

### 3.2.3 Tempo médio para a primeira visita ao estado *j*

No modelo MARKOV-SPI, o tempo esperado para a primeira visita ao estado *j*, partindo do estado *i*,  $t_{ij}$ , e o tempo médio de recorrência do estado *i*,  $t_{ii}$ , nas suas versões homogêneas, são dados pelas expressões (3.29) e (3.30) da secção 3.1.4.3, respectivamente.

Para o modelo MARKOV-SPI-NAO, estas duas expressões escrevem-se para os dois casos de *NAO* positiva e negativa, da seguinte forma, respectivamente:

$$t_{ij} = 1 + \sum_{k \neq j} p_{ik}^{NAO_{+/-}} t_{kj} \quad \forall i, j \in S \quad (3.63)$$

e,

$$t_{ii} = \frac{1}{\varphi_j} \quad (3.64)$$

Onde  $\varphi_j$  é o vector de equilíbrio e  $t_{ii}$  o tempo de recorrência na classe *i*.

### 3.2.4 Previsões a curto prazo

As previsões a curto prazo, no modelo MARKOV-SPI, nas suas versões homogêneas, são dadas pelas expressões (3.33), (3.34) e (3.35) da secção 3.1.4.4.

Para o modelo MARKOV-SPI-NAO, estas duas expressões escrevem-se para os dois casos de *NAO* positiva e negativa, da seguinte forma:

$$\Pr(SPI_{t+1} = j | SPI_t = i, NAO_{+/-}^{(t)}) \quad (3.65)$$

$$\Pr(SPI_{t+2} = j | SPI_t = i, NAO_{+/-}^{(t)}) = \sum_{\forall k \in S} \Pr(SPI_{t+2} = j | SPI_{t+1} = k, NAO_{+/-}^{(t)}) \Pr(SPI_{t+1} = k | SPI_t = i, NAO_{+/-}^{(t)}) \quad (3.66)$$

$$\Pr(SPI_{t+3} = j | SPI_t = i, NAO_{+/-}^{(t)}) = \sum_{\forall l \in S} \sum_{\forall k \in S} \Pr(SPI_{t+3} = j | SPI_{t+2} = l, NAO_{+/-}^{(t)}) \Pr(SPI_{t+2} = l | SPI_{t+1} = k, NAO_{+/-}^{(t)}) \Pr(SPI_{t+1} = k | SPI_t = i, NAO_{+/-}^{(t)}) \quad (3.67)$$

A classe de seca mais provável num horizonte de previsão de  $m$  meses é aquela que apresentar o valor máximo de  $P(X_{t+m} = j | X_t = i, NAO_{+/-}^{(t)})$ .

### 3.3 Indicadores de precisão – “Skill Scores”

Indicadores de precisão, são medidas estatísticas que avaliam a performance de uma previsão. Essa avaliação é feita tendo em conta a correspondência entre cada previsão e observação. No caso em estudo, as classes de categoria de seca são variáveis discretas, pelo que a avaliação das previsões é feita através do uso de tabelas de contingência entre as classes observadas e as classes previstas e o uso de indicadores apropriados a esta avaliação.

Tabela 3.1 Tabela de contingência para K classes de severidade de seca (previsão/observação)

<b>Observação Previsão</b>	Classe 1	...	Classe K	<b>% Total Previsão</b>
Classe 1	$p_{11}$	$p_{1j}$	$p_{1K}$	$\hat{p}_1$
$\vdots$	$p_{i1}$	$p_{ii}$	$p_{iK}$	$\hat{p}_i$
Classe K	$p_{K1}$	$p_{Kj}$	$p_{KK}$	$\hat{p}_K$
<b>% Total Observações</b>	$p_1$	$p_i$	$p_K$	100

#### 3.3.1 HSS (Heidke Skill Score) e PSS (Peirce’s Skill Score)

Neste trabalho os Indicadores usados foram o HSS – *Heidke Skill Score* e o PSS – *Peirce’s Skill Score*, que são definidos, respectivamente, por:

$$HSS = \left( \sum_{i=1}^K p_{ii} - \sum_{i=1}^K p_i \hat{p}_i \right) / \left( 1 - \sum_{i=1}^K p_i \hat{p}_i \right), \quad (3.68)$$

$$PSS = \left( \sum_{i=1}^K p_{ii} - \sum_{i=1}^K p_i \hat{p}_i \right) / \left( 1 - \sum_{i=1}^K p_i p_i \right), \quad (3.69)$$

Onde:

$p_{ii}$  corresponde à percentagem de previsões concordantes com as observações para uma determinada classe de seca  $i$  (entrada  $(i, i)$  da tabela de contingência);

$p_i$  é a percentagem total de observações da classe  $i$  da amostra (para qualquer classe  $j$  da previsão correspondente).

$\hat{p}_i$  é a percentagem total de previsões da classe  $i$  da amostra (para qualquer classe  $j$  da observação correspondente).

$K$  é o total de pares observação/previsão da amostra.

### 3.3.2 Proporção Correcta (PC)

O indicador PC, dá informação sobre a percentagem de casos em que a classe prevista é igual à classe observada:

$$PC = \sum_{i=1}^K p_{ii} \quad (3.70)$$

### 3.3.3 Correlação

O **coeficiente de correlação de Pearson**, também chamado de *coeficiente de correlação produto-momento* ou simplesmente de *r de Pearson* mede o grau da correlação (e a direcção dessa correlação - se positiva ou negativa) entre duas variáveis.

O coeficiente de correlação de *Pearson* calcula-se segundo a seguinte fórmula:

$$r = \frac{\sum_{i=1}^n x_i' y_i'}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{i=1}^n (y_i - \bar{y})^2}} \quad (3.71)$$

Onde  $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$  e  $\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i$ .

Este coeficiente, normalmente representado pela letra "r" assume apenas valores entre -1 e 1. Quando  $r = 1$ , significa uma correlação perfeita positiva entre as duas variáveis. Se  $r = -1$ , significa uma correlação negativa perfeita entre as duas variáveis - isto é, se uma aumenta, a outra sempre diminui. No caso de  $r = 0$ , significa que as duas variáveis não dependem linearmente uma da outra. Quanto maior o valor de  $|r|$ , mais forte é a relação. Portanto, para  $r > 0$  ambas as variáveis crescem juntas, e para  $r < 0$  quando uma das variáveis diminui, a outra aumenta.

## 4. Dados

A região do Alentejo localiza-se no sul de Portugal entre as latitudes 37°20' N – 39°40' N e entre as longitudes 6°55' – 8°50' W (ver Figura 4.1) e abrangendo uma área de 27000 km<sup>2</sup>. Esta área é caracterizada pela agricultura intensiva (de sequeiro<sup>5</sup>), maioritariamente cereais, oliveiras, vinhas, florestas de azinho e sobreiro. A precipitação é, de um modo geral, relativamente abundante durante o período de Outono-Inverno e extremamente escassa durante o verão. A exploração agrícola ocupa 69% da área total e a superfície irrigável representa apenas 6% da superfície agrícola total (GPPAA, 2000). A agricultura está assim, fortemente dependente da ocorrência e distribuição da precipitação durante o ciclo de desenvolvimento das culturas.



Figura 4.1 Estações udométricas do Alentejo, escolhidas para modelação de Markov através de séries temporais de SPI (Paulo e Pereira, 2005).

Na modelação de cadeias de Markov, foram utilizados séries temporais de dados provenientes de:

- ❖ Reanálises do NCEP-NCAR da precipitação acumulada mensal.
- ❖ Várias estações udométricas da precipitação acumulada mensal (ver Figura 4.1 e Tabela 2.1) do INAG, na região do Alentejo.

Tabela 4.1 Localização das 7 estações udométricas no Alentejo.

Código da estação	Local da estação	Latitude (N)	Longitude (W)	Elevação (m)
18M01	Portalegre	39°17'	7°25'	596
20O01	Elvas	38°53'	7°09'	208
22J01	Évora	38°34'	7°54'	309
25J02	Beja	38°01'	7°52'	246
25P01	Barrancos	38°08'	6°58'	380
26G01	Alvalade	37°57'	8°24'	61
28I01	Almodôvar	37°31'	8°04'	270

<sup>5</sup> Agricultura de sequeiro é uma técnica agrícola para cultivar terrenos onde a pluviosidade é diminuta.

As séries mensais das estações do INAG têm uma dimensão de  $n=805$  correspondente ao número de meses entre Setembro de 1932 e Setembro de 1999. No caso das séries mensais das reanálises NCEP ( $2.5^\circ$  lat x  $2.5^\circ$  long) a sua dimensão é  $n=672$  correspondente ao número de meses do período de tempo entre Janeiro de 1950 a Dezembro de 2005, sendo que nesta malhas apenas foi usado um ponto (a vermelho), o único mais perto da área do Alentejo (ver Figura 4.2).

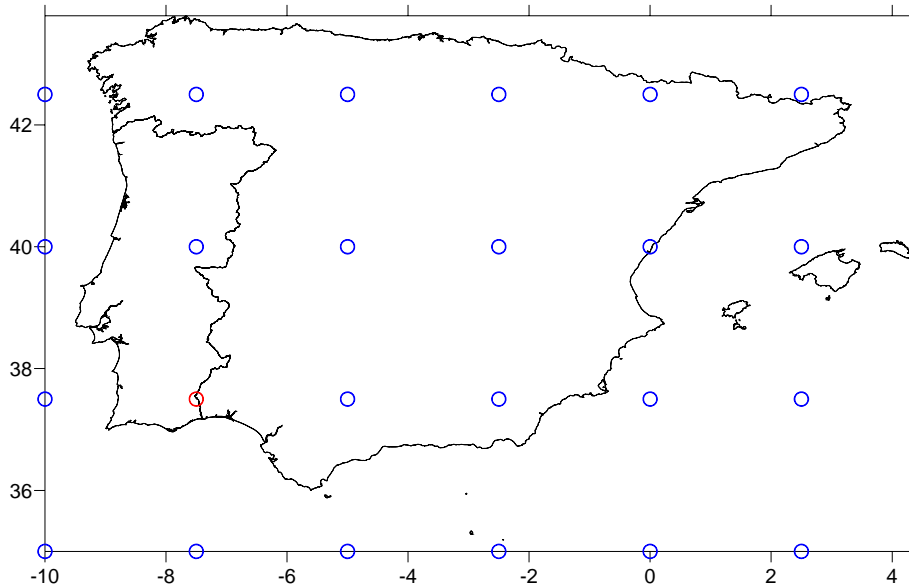


Figura 4.2 Malha de dados de precipitação acumulada mensal das reanálises NCEP ( $2.5^\circ$  lat x  $2.5^\circ$  long). O ponto a vermelho é o ponto escolhido

As séries temporais de SPI foram obtidas a partir das séries temporais da precipitação acumulada anteriores, pelo processo descrito na secção 2.4.1. Obtiveram-se séries temporais de SPI mensal (SPI-1), trimensal (SPI-3) e anual (SPI-12).

Foram usadas ainda, séries temporais de SPI determinadas a partir de séries da precipitação acumulada mensal de 49 estações udométricas do Alentejo, através da utilização da técnica dos polígonos de Thiessen. Deste modo, estas séries caracterizam o SPI regional representativo de toda a área do Alentejo.

Foram usadas também, séries temporais dos índices da NAO e da AO, obtidas a partir do *Climate Prediction Center* (CPC), em que a informação disponível correspondente ao período de tempo de Janeiro de 1950 até à actualidade (excepto o mês actual).

Para o caso da Oscilação do Sul, a série usada é designada como série temporal mensal do índice bivariado do ENSO (BEST), retirada do site do *Global Change Master Directory* da NASA. O índice BEST foi concebido para ser de cálculo simples e para permitir um índice ENSO de longo período para propósitos de pesquisa. O Nino 3.4 (média das temperaturas nas regiões 3 e 4 do El Niño) tem sido usado como medida da intensidade do ENSO no Pacífico tropical. Contudo, o seu uso isolado negligencia processos atmosféricos explícitos. Combinando-o com, o índice de Oscilação do Sul ou "SOI" (diferença de pressão entre Tahiti e Darwin), estes processos ficam incluídos. Além disso, valores mais antigos das SST são pelo menos parcialmente reconstruídos e não explicitamente medidos. Pela inclusão do SOI, que é a melhor medida historicamente, os efeitos do desvio médio (Viés) nos dados das SST introduzidos pela técnica de reconstrução são reduzidos. ([http://gcmd.nasa.gov/records/GCMD\\_NOAA\\_NWS\\_CPC\\_BEST.html](http://gcmd.nasa.gov/records/GCMD_NOAA_NWS_CPC_BEST.html)).

Esta série temporal é baseada na combinação da componente atmosférica do fenómeno do ENSO (o índice de Oscilação do Sul ou "SOI") e uma componente oceânica (Nino 3.4 SST que é definida como sendo a média da SST sobre a região 5N-5S e 170W-120W). O índice SOI é retirado de Phil Jones do *Climatic Research Unit* da Ânglia Oriental, Inglaterra, com um período de 1855 a 1997. Os valores seguintes do SOI são retirados do CPC. O índice da SST Nino 3.4 é retirado da base de dados GISST (até Abril de 1998). Os valores seguintes são da base de dados SST Reynolds.

(<http://www.esrl.noaa.gov/psd/people/cathy.smith/best/details.html>)

Para a discretização das séries temporais de valores contínuos de SPI foram usadas quatro classes de seca correspondentes às da Tabela 2.2, com as duas últimas classes a fundirem-se numa só devido ao reduzido número de transições de e para esta classe. Assim sendo, as quatro novas classes são:

**Tabela 4.2 Classificação utilizada na discretização das séries temporais de SPI.**

<b>Valores de SPI</b>	<b>Categoria de seca</b>
$SPI \geq 0$	Não-seca (N)
$-1 < SPI < 0$	Seca ligeira (1)
$-1.5 \leq SPI < -1$	Seca moderada (2)
$SPI \leq -1.5$	Seca severa ou extrema (3)

## 5. Resultados e Discussão

### 5.1 Matrizes de transição

Em seguida, são mostrados alguns exemplos de matrizes de transição associadas às cadeias de Markov usadas.

$$P = \begin{bmatrix} 0.5660 & 0.4340 \\ 0.4488 & 0.5512 \end{bmatrix} = \begin{bmatrix} \Pr(\text{NAO}_+^{(t+1)} | \text{NAO}_+^{(t)}) & \Pr(\text{NAO}_-^{(t+1)} | \text{NAO}_+^{(t)}) \\ \Pr(\text{NAO}_+^{(t+1)} | \text{NAO}_-^{(t)}) & \Pr(\text{NAO}_-^{(t+1)} | \text{NAO}_-^{(t)}) \end{bmatrix} \quad (5.1)$$

A matriz de transição do índice NAO (5.1) é a matriz das transições associada à série temporal usada a partir do CPC. Para o caso do modelo MARKOV-SPI-NAO, é possível verificar observando esta matriz que em qualquer dos casos de NAO+ e NAO-, a probabilidade de persistência de ambos os regimes é maior do que a transição de um regime para outro, bastando verificar que a diagonal da matriz tem os valores maiores. Como em qualquer matriz de transição, os elementos desta são positivos sendo a soma dos elementos das linhas igual a 1.

A matriz das transições da variável aleatória  $X$  condicionada a  $\text{NAO}^{(t+1)} > 0$  e para as classes de seca usadas  $\{X_t\} = \{N, 1, 2, 3\}$ , é definida da seguinte forma:

$$M_+ = \begin{bmatrix} \Pr(X_{t+1} = N | X_t = N, \text{NAO}_+^{(t+1)}) & \cdots & \Pr(X_{t+1} = 3 | X_t = N, \text{NAO}_+^{(t+1)}) \\ \vdots & \ddots & \vdots \\ \Pr(X_{t+1} = N | X_t = 3, \text{NAO}_+^{(t+1)}) & \cdots & \Pr(X_{t+1} = 3 | X_t = 3, \text{NAO}_+^{(t+1)}) \end{bmatrix} \quad (5.2)$$

No caso do SPI-1 (mensal) as matrizes  $M_+$  e  $M_-$  são as seguintes:

$$M_+ = \begin{bmatrix} 0.4430 & 0.3811 & 0.0912 & 0.0847 \\ 0.4811 & 0.3302 & 0.1226 & 0.0660 \\ 0.3818 & 0.4364 & 0.0364 & 0.1455 \\ 0.1250 & 0.5625 & 0.1250 & 0.1875 \end{bmatrix} \quad (5.3)$$

$$M_- = \begin{bmatrix} 0.6959 & 0.2351 & 0.0376 & 0.0313 \\ 0.5824 & 0.3077 & 0.0769 & 0.0330 \\ 0.4634 & 0.4390 & 0.0488 & 0.0488 \\ 0.4091 & 0.3182 & 0.1818 & 0.0909 \end{bmatrix} \quad (5.4)$$

É possível observar as diferenças dos valores entre as matrizes (5.3) e (5.4). A matriz  $M_+$ , que é condicionada a  $\text{NAO}^{(t+1)} > 0$  tem menor probabilidade de permanência na classe N ("Não-seca") em relação à matriz  $M_-$ , condicionada a  $\text{NAO}^{(t+1)} < 0$ , respectivamente, 0.4430 e 0.6959. Tal corrobora o facto de a "Não-seca" ou precipitação acima da mediana no Alentejo ser favorecido por condições de fase negativa da NAO. Mais ainda, é possível verificar as diferenças entre as

entradas da coluna 1 das matrizes  $M_-$  e  $M_+$ . Esta coluna tem a seguinte probabilidade:  $\Pr(X_{t+1} = N | X_t = x) \quad \forall x \in \{N, 1, 2, 3\}$ . Assim, verifica-se que as entradas desta coluna são sempre maiores no caso da matriz  $M_-$  face à sua homóloga  $M_+$ . Este resultado é esperado uma vez que a fase negativa da NAO é mais favorável à ocorrência de precipitação que a fase positiva, sendo esta mais favorável à ocorrência de secas. O contrário se observa na coluna 4 de ambas as matrizes cuja probabilidade é  $\Pr(X_{t+1} = 3 | X_t = x) \quad \forall x \in \{N, 1, 2, 3\}$ . Todas as entradas desta coluna têm valores maiores no caso da matriz  $M_+$  comparativamente à matriz  $M_-$  pelas razões já apresentadas.

No caso não-homogéneo, a cada mês está associada uma matriz de transição específica. Assim existem, para cada mês duas matrizes condicionadas a  $NAO^{(t+1)} > 0$  e  $NAO^{(t+1)} < 0$ . São apresentadas, um exemplo destas duas matrizes para o caso do mês de Dezembro:

$$M_+^{(Dez)} = \begin{bmatrix} 0.5882 & 0.4118 & 0.0000 & 0.0000 \\ 0.4400 & 0.4800 & 0.0000 & 0.0800 \\ 0.2500 & 0.7500 & 0.0000 & 0.0000 \\ 0.3333 & 0.3333 & 0.0000 & 0.3333 \end{bmatrix} \quad (5.5)$$

$$M_-^{(Dez)} = \begin{bmatrix} 0.6061 & 0.3939 & 0.0000 & 0.0000 \\ 1.0000 & 0.0000 & 0.0000 & 0.0000 \\ 0.0000 & 0.0000 & 0.0000 & 1.0000 \\ 0.0000 & 0.0000 & 0.0000 & 0.0000 \end{bmatrix} \quad (5.6)$$

À semelhança do que acontece com as matrizes do resto dos meses do ano, há uma grande quantidade de entradas nulas. De uma forma geral, a matriz  $M_-^{(Dez)}$  tem os valores não nulos confinados às classes associadas à "Não-seca" exceptuando o caso da transição entre classes 3→2 com a probabilidade de 1.0. A matriz  $M_+^{(Dez)}$  tem uma maior quantidade de entradas não nulas havendo assim maior quantidade de visitas a classes de severidade de seca maiores (fase positiva da NAO).

## 5.2 Correlação entre os índices de seca e de CGA

De forma a poder quantificar a relação entre a quantidade de precipitação numa determinada área e os índices de CGA, são apresentadas as correlações entre estes dois índices (séries mensais de SPI e dos índices de CGA já referidos anteriormente), para a área do Alentejo e com SPI-1 (mensal) e SPI-12 (anual). No âmbito deste trabalho, pretende-se estudar em particular a escassez de precipitação (que está correlacionada com as fases positivas da NAO e AO), pelo que são calculadas também as correlações assimétricas para as fases positivas e negativas dos regimes de CGA, nas tabelas seguintes. São ainda calculadas as correlações entre os índices  $NAO$  e  $SPI$  a vários meses de antecedência (Lag), ou seja,  $r_{(NAO_{t-lag}, SPI_t)}$ .

Tabela 5.1 Correlação entre o índice *NAO* e o índice *SPI-1* a vários meses de antecedência para a região do Alentejo.

Correlação <i>r</i>	Lag=0	Lag=1	Lag=2
<i>r</i>	-0.228	-0.040	-0.011
$r_{NAO>0}$	0.000	-0.065	0.007
$r_{NAO<0}$	-0.182	-0.036	0.030

Tabela 5.2 Correlação entre o índice *AO* e o índice *SPI-1* a vários meses de antecedência para a região do Alentejo.

Correlação <i>r</i>	Lag=0	Lag=1	Lag=2
<i>r</i>	-0.294	-0.107	0.001
$r_{AO>0}$	-0.128	-0.041	0.086
$r_{AO<0}$	-0.233	-0.064	-0.003

Tabela 5.3 Correlação entre o índice *SOI* e o índice *SPI-1* a vários meses de antecedência para a região do Alentejo.

Correlação <i>r</i>	Lag=0	Lag=1	Lag=2
<i>r</i>	0.036	0.066	0.074
$r_{SOI>0}$	0.023	0.070	0.074
$r_{SOI<0}$	0.090	0.085	0.095

Como é possível constatar pela leitura da Tabela 5.1, Tabela 5.2 e Tabela 5.3, existe para as séries de dados utilizadas, uma correlação significativa entre os índices *NAO* e *SPI* e *AO* e *SPI*, sendo que, o índice *AO* está mais fortemente correlacionado com o *SPI* do que o índice *NAO* com o *SPI*. No caso do *SOI* os valores de correlação, são, de uma forma geral, bastante inferiores aos outros dois índices. É ainda visível que essa relação estatística vai baixando com o aumento dos meses de antecedência da *NAO* e *AO* em relação ao *SPI*. Contudo, isto não se verifica com o índice *SOI* em que quanto maior for o tempo de desfasamento entre este índice e o *SPI*, os valores de correlação aumentam, apesar de estes serem bastante baixos. A partir de lag=3, embora não esteja apresentado, os valores baixam significativamente. Mais importante ainda verificar, é a diferença significativa entre os valores da correlação das fases positiva e negativa dos índices *NAO* e *AO* com o *SPI*: a fase negativa em ambos os índices, ou seja, a fase que favorece maior precipitação em Portugal está mais fortemente relacionada com o *SPI* do que a fase positiva (Pires e Perdigão, 2007), ou seja, na utilização destes dois índices é melhorada ainda mais a previsão no caso de *AO-* e *NAO-*. No caso do índice *SOI*, a fase negativa está mais fortemente correlacionada com o *SPI* quando os meses são os mesmos, a partir do primeiro mês de antecedência entre os índices a fase negativa continua a estar mais correlacionada com o *SPI*, do que a fase positiva embora a diferença não seja tão grande. Assim, é aumentada a qualidade da previsão aquando de eventos de precipitação abundante (ex: cheias) face aos eventos de escassez de precipitação (ex: secas). Não é de estranhar deste modo, que no Inverno a correlação seja muito maior que no Verão ou ao longo de todo o ano (ver Tabela 5.4).

Tabela 5.4 Correlação entre o índice *NAO* e o índice *SPI-1* (lag=0), para todos os meses e para os meses de Outono-Inverno (Out-Mar) e Primavera-Verão (Abr-Set).

Correlação <i>r</i>	Ano Inteiro	Outono-Inverno	Primavera-Verão
<i>r</i>	-0.228	-0.465	0.013
$r_{NAO>0}$	0.000	-0.166	0.146
$r_{NAO<0}$	-0.182	-0.300	-0.077

Na Tabela 5.5 e Tabela 5.6, são mostrados os valores correspondentes às correlações entre os índices *NAO* e *AO* e o *SPI-12* anual, a vários "lag's". É visível que os valores de correlação são mais baixos do que no caso do *SPI-1* (mensal), uma vez que a série de *SPI-12* corresponde a valores médios anuais da precipitação acumulada. No caso do índice *SOI*, a correlação com o *SPI-12* obteve valores maiores do que no caso do *SPI-1*, embora apenas as correlações das fases positivas e negativas tenham valores significativos quando separadas (ver Tabela 5.7).

Tabela 5.5 Correlação entre o índice *NAO* e o índice *SPI-12* a vários meses de antecedência para a região do Alentejo.

Correlação <i>r</i>	Lag=0	Lag=1	Lag=2
<i>r</i>	-0.108	0.115	-0.093
$r_{NAO>0}$	0.020	0.019	0.050
$r_{NAO<0}$	-0.045	-0.102	-0.086

Tabela 5.6 Correlação entre o índice *AO* e o índice *SPI-12* a vários meses de antecedência para a região do Alentejo.

Correlação <i>r</i>	Lag=0	Lag=1	Lag=2
<i>r</i>	-0.137	-0.160	-0.151
$r_{AO>0}$	-0.020	-0.037	-0.024
$r_{AO<0}$	-0.154	-0.182	-0.165

Tabela 5.7 Correlação entre o índice *SOI* e o índice *SPI-12* a vários meses de antecedência para a região do Alentejo.

Correlação <i>r</i>	Lag=0	Lag=1	Lag=2
<i>r</i>	-0.015	-0.001	0.016
$r_{SOI>0}$	-0.143	-0.137	-0.122
$r_{SOI<0}$	0.137	0.152	0.179

### 5.3 Performance dos modelos Markov-SPI e Markov-SPI-ÍndiceCGA

Em seguida, apresentam-se várias tabelas comparando através de indicadores estatísticos, a qualidade das previsões das classes de seca, de 1 a 3 meses no modo Hindcast<sup>6</sup>, dos dois modelos com e sem índice de CGA, variando esse mesmo índice, nas versões homogêneas e não-homogêneas, no período anual (Janeiro –

<sup>6</sup> Tomando um mês no passado faz-se previsões para os meses seguintes também no passado, e posteriormente, faz-se a comparação dessas classes previstas com as classes observadas para os meses correspondentes (no passado).

Dezembro) e nos dois períodos sazonais (Outubro – Março e Abril – Setembro). MS, nas tabelas, representa MARKOV-SPI.

Para a avaliação da performance dos modelos nas suas variantes são usados os indicadores de precisão estatísticos HSS e PSS, apresentados na secção 3.3.1 e 3.3.2. Há que referir que o indicador PC refere-se à percentagem de acertos (classe prevista num determinado mês igual à classe observada para esse mesmo mês). Os indicadores HSS e PSS referem-se não só à percentagem de acertos mas também à utilização ou não, das várias classes de seca prevista, ou seja, se as previsões feitas pelo modelo compreendem todas as classes de seca ou não.

**Tabela 5.8 Indicadores HSS e PSS para as previsões de 1 a 3 meses para os modelos com SPI-1 (mensal), com o índice da NAO, AO e SOI e sem índice de CGA nas várias opções de homogeneidade e não homogeneidade, na região do Alentejo.**

Modelo	Indicador	Homogéneo (Jan-Dez)				Não Homogéneo (Jan-Dez)				Não Homogéneo (Out-Mar)				Não Homogéneo (Abr-Set)			
		MS	MS-NAO	MS-AO	MS-SOI	MS	MS-NAO	MS-AO	MS-SOI	MS	MS-NAO	MS-AO	MS-SOI	MS	MS-NAO	MS-AO	MS-SOI
1 Mês	HSS	0.08	0.08	0.08	0.07	0.15	0.18	0.19	0.20	0.13	0.15	0.20	0.18	0.16	0.21	0.17	0.22
	PSS	0.06	0.06	0.06	0.05	0.12	0.16	0.16	0.17	0.11	0.13	0.18	0.16	0.14	0.18	0.14	0.19
2 Meses	HSS	0.00	0.00	0.00	0.00	0.06	0.06	0.05	0.06	0.00	0.00	0.01	0.02	0.12	0.13	0.08	0.12
	PSS	0.00	0.00	0.00	0.00	0.05	0.05	0.04	0.05	0.00	0.00	0.01	0.01	0.10	0.11	0.07	0.10
3 Meses	HSS	0.00	0.00	0.00	0.00	0.05	0.05	0.05	0.06	0.00	0.00	0.00	0.02	0.11	0.11	0.11	0.11
	PSS	0.00	0.00	0.00	0.00	0.04	0.04	0.04	0.05	0.00	0.00	0.00	0.01	0.09	0.09	0.09	0.09

**Tabela 5.9 Indicador PC para as previsões de 1 a 3 meses para os modelos com SPI-1 (mensal), com o índice da NAO, AO e SOI e sem índice de CGA nas várias opções de homogeneidade e não homogeneidade, na região do Alentejo.**

Modelo	Indicador	Homogéneo (Jan-Dez)				Não Homogéneo (Jan-Dez)				Não Homogéneo (Out-Mar)				Não Homogéneo (Abr-Set)			
		MS	MS-NAO	MS-AO	MS-SOI	MS	MS-NAO	MS-AO	MS-SOI	MS	MS-NAO	MS-AO	MS-SOI	MS	MS-NAO	MS-AO	MS-SOI
1 Mês	PC	0.54	0.54	0.54	0.54	0.57	0.57	0.58	0.59	0.56	0.56	0.58	0.58	0.58	0.59	0.57	0.60
2 Meses	PC	0.53	0.53	0.53	0.53	0.54	0.53	0.53	0.54	0.51	0.51	0.51	0.52	0.56	0.56	0.55	0.56
3 Meses	PC	0.53	0.53	0.53	0.53	0.54	0.54	0.54	0.54	0.51	0.51	0.51	0.52	0.56	0.56	0.56	0.56

Pela análise das duas tabelas acima é possível constatar que a formulação das cadeias de Markov não-homogénea tem uma performance bastante superior à formulação homogénea. A partir do segundo mês de previsão a qualidade decresce ligeiramente face à previsão para o primeiro mês. As performances dos modelos Markov-SPI-NAO, Markov-SPI-AO e MARKOV-SPI-SOI são ligeiramente maiores que a do modelo sem índice de CGA. No geral, para o período de Abril a Setembro e para todo o ano, o índice SOI produz melhores resultados que os índices da NAO e AO, mas no período de Outubro a Março, verifica-se que o índice da AO e do SOI têm os melhores resultados.

Tabela 5.10 Indicadores HSS e PSS para as previsões de 1 a 3 meses para os modelos com SPI-12 (anual), com o índice da NAO, AO e SOI e sem índice de CGA nas várias opções de homogeneidade e não homogeneidade, na região do Alentejo.

Modelo	Indicador	Homogéneo (Jan-Dez)				Não Homogéneo (Jan-Dez)				Não Homogéneo (Out-Mar)				Não Homogéneo (Abr-Set)			
		MS	MS-NAO	MS-AO	MS-SOI	MS	MS-NAO	MS-AO	MS-SOI	MS	MS-NAO	MS-AO	MS-SOI	MS	MS-NAO	MS-AO	MS-SOI
1 Mês	HSS	0.74	0.74	0.74	0.73	0.74	0.74	0.76	0.76	0.61	0.61	0.63	0.64	0.88	0.87	0.88	0.89
	PSS	0.74	0.74	0.74	0.71	0.73	0.72	0.74	0.75	0.58	0.58	0.60	0.61	0.87	0.85	0.88	0.88
2 Meses	HSS	0.59	0.59	0.59	0.60	0.62	0.59	0.63	0.64	0.51	0.44	0.52	0.51	0.74	0.73	0.74	0.77
	PSS	0.56	0.56	0.56	0.59	0.61	0.56	0.61	0.61	0.49	0.42	0.50	0.48	0.72	0.70	0.73	0.75
3 Meses	HSS	0.50	0.50	0.50	0.49	0.52	0.51	0.52	0.53	0.44	0.43	0.45	0.45	0.59	0.59	0.59	0.61
	PSS	0.48	0.48	0.48	0.47	0.49	0.48	0.49	0.50	0.41	0.40	0.42	0.43	0.57	0.56	0.57	0.58

Tabela 5.11 Indicador PC para as previsões de 1 a 3 meses para os modelos com SPI-12 (anual), com o índice da NAO, AO e SOI e sem índice de CGA nas várias opções de homogeneidade e não homogeneidade, na região do Alentejo.

Modelo	Indicador	Homogéneo (Jan-Dez)				Não Homogéneo (Jan-Dez)				Não Homogéneo (Out-Mar)				Não Homogéneo (Abr-Set)			
		MS	MS-NAO	MS-AO	MS-SOI	MS	MS-NAO	MS-AO	MS-SOI	MS	MS-NAO	MS-AO	MS-SOI	MS	MS-NAO	MS-AO	MS-SOI
1 Mês	PC	0.84	0.84	0.84	0.83	0.85	0.84	0.85	0.86	0.76	0.77	0.77	0.78	0.93	0.92	0.93	0.93
2 Meses	PC	0.75	0.75	0.75	0.76	0.77	0.75	0.78	0.78	0.70	0.67	0.71	0.70	0.84	0.84	0.85	0.86
3 Meses	PC	0.70	0.70	0.70	0.70	0.71	0.71	0.72	0.72	0.68	0.66	0.68	0.67	0.75	0.76	0.76	0.77

As duas tabelas acima apresentam os resultados dos indicadores para o SPI-12 (anual), e revelam um significativo aumento de qualidade face ao SPI-1 (mensal). Também se deve ao facto da série de SPI-12 ser menos sensível às alterações de precipitação de um determinado mês, já que é função da média das precipitações acumuladas dos 12 meses anteriores ao referente mês. O índice SPI-1 é muito sensível às alterações da quantidade de precipitação e assim mais difícil de prever. No caso do SPI-12 não há grandes melhorias na formulação não-homogénea face à homogénea, os resultados pioram para os meses de Outubro a Março e apenas melhoram nos meses de Abril a Setembro. Em geral, os índices SOI e AO produzem melhores resultados que o modelo sem índices de CGA, MARKOV SPI.

Não foram aqui apresentados os resultados para o caso do SPI-3 por uma questão de síntese de informação. Os resultados dos indicadores estatísticos obtidos para este caso foram melhores que no caso do SPI-1 e piores do que no caso do SPI-12.

## 5.4 Vector de equilíbrio

A longo prazo, a probabilidade da classe de seca é independente do estado inicial da rede de Markov. As tabelas seguintes, mostram as probabilidades de cada uma das classes de seca e do índice de CGA num futuro distante.

Tabela 5.12 Vector de equilíbrio do Índice de CGA para a região do Alentejo.

Modelo	Fase do índice	
	Positiva	Negativa
MARKOV-SPI-NAO	0.5084	0.4916
MARKOV-SPI-AO	0.4550	0.5450
MARKOV-SPI-SOI	0.5126	0.4874

A longo prazo, a probabilidade de estar numa fase positiva da NAO é cerca de 0.51, enquanto que a fase negativa é 0.49. No caso da AO a diferença já é maior, prevalecendo a fase negativa com uma probabilidade 0.54 para 0.46 da fase positiva. No caso do SOI, a fase positiva é de cerca de 0.51 e a fase negativa de cerca de 0.49.

Tabela 5.13 Vector de equilíbrio do SPI-1 e SPI-12 para a região do Alentejo.

Modelo	Classe de seca	Classe de seca			
		N	1	2	3
SPI-1	MARKOV-SPI	0.5269	0.3288	0.0805	0.0638
	MARKOV-SPI-NAO	0.5242	0.3314	0.0804	0.0640
	MARKOV-SPI-AO	0.5266	0.3289	0.0806	0.0639
	MARKOV-SPI-SOI	0.5279	0.3271	0.0816	0.0634
SPI-12	MARKOV-SPI	0.4706	0.3628	0.0851	0.0814
	MARKOV-SPI-NAO	0.4827	0.3696	0.0825	0.0653
	MARKOV-SPI-AO	0.4743	0.3667	0.0829	0.0762
	MARKOV-SPI-SOI	0.4715	0.3727	0.0794	0.0764

No caso do vector de equilíbrio do SPI não existem grandes diferenças entre os modelos com e sem índice de CGA como se pode ver na Tabela 5.13. No caso do SPI-1 a probabilidade de longo prazo para a classe de "Não-seca" é 0.52 (MARKOV-SPI-NAO) ou 0.53 (MARKOV-SPI e MARKOV-SPI-AO e MARKOV-SPI-SOI) e no caso do SPI-12 essa mesma classe de seca as probabilidades são também muito idênticas entre os 0.47 e 0.48. As diferenças mais acentuadas são entre o SPI nas diferentes escalas de tempo: para o SPI-1, a classe de "Não-seca" tem uma maior probabilidade de ocorrência que no caso do SPI-12 com uma diferença de cerca de 0.05, embora seja sempre esta, a classe mais provável. Para as restantes classes os valores não variam muito para as duas escalas de SPI apresentadas, sendo que, a classe 2 ("seca moderada") é a que apresenta menos diferenças de probabilidade entre as escalas mensal e anual do SPI.

## 5.5 Tempo médio de permanência ininterrupta no estado i

Na Tabela 5.14 é apresentado o tempo médio de permanência ininterrupta num determinado estado, ou tempo de residência esperado nesse mesmo estado, para os vários modelos e para as duas diferentes escalas de tempo do SPI.

Tabela 5.14 Tempo de residência esperado na classe i (meses) para a região do Alentejo.

Modelo		Classe de seca				
		N	1	2	3	
SPI-1	MARKOV-SPI	2.34	1.47	1.04	1.15	
	MARKOV-SPI-NAO	NAO+	1.80	1.49	1.04	1.23
		NAO-	3.29	1.44	1.05	1.10
	MARKOV-SPI-AO	AO+	1.73	1.65	1.10	1.18
		AO-	3.29	1.35	1.00	1.12
	MARKOV-SPI-SOI	SOI+	2.27	1.46	1.00	1.07
SOI-		2.41	1.48	1.11	1.22	
SPI-12	MARKOV-SPI	12.35	5.79	1.96	4.30	
	MARKOV-SPI-NAO	NAO+	11.25	5.09	1.93	6.75
		NAO-	13.55	6.75	2.00	2.67
	MARKOV-SPI-AO	AO+	9.92	5.45	2.00	4.75
		AO-	15.00	6.17	1.92	4.00
	MARKOV-SPI-SOI	SOI+	14.10	6.44	2.29	5.50
SOI-		11.00	5.20	1.55	3.79	

Neste caso, os valores são completamente distintos entre as diferentes escalas de tempo do SPI. Para o SPI-1, as diferenças entre os tempos de residência esperados para as 3 classes de seca são muito semelhantes. O tempo de residência esperado para a classe de "Não-seca" é pouco maior que nas outras 3 classes de seca, estando entre o dobro e o triplo dos tempos dessas mesmas classes. Isto deve-se ao facto do SPI-1 não ter em conta o agravamento de uma seca e a recuperação da mesma, uma vez que apenas reage às mudanças mensais da precipitação. Ainda assim, é de notar que para o caso da classe de "Não-seca", os tempos de residência são menores nos casos das fases positivas da AO e NAO (devido à escassez de precipitação típica destas fases) e maiores aquando das fases negativas das mesmas. No caso do SOI a diferença entre os tempos de residência é mínima, uma vez que o fenómeno do El Niño não é tão preponderante na PI. Sem índice de CGA (MARKOV-SPI) o valor é intermédio aos modelos com índice de CGA. Para as restantes classes de severidade de seca os valores são muito semelhantes entre si.

No caso do SPI-12 os valores já são bem distintos entre as várias classes de categoria de seca. Mais uma vez, é notória a diminuição do tempo de residência na classe de "Não-seca" aquando das fases positivas da NAO e AO e o contrário para as fases negativas. No caso do índice SOI o tempo de residência é maior para a fase positiva do índice. Para a classe de "Não-seca" (1), os valores são mais extremos nos casos das fases positiva e negativa da AO do que no caso da NAO, enquanto que, se verifica o contrário para a classe de "seca severa ou extrema" (3), ou seja, os valores do tempo de residência são mais extremos nas fases positiva e negativa da NAO face às mesmas fases da AO. Assim sendo, a fase da AO negativa é preciso esperar mais tempo para deixarmos de estar em condições de Não-seca do que na fase negativa da NAO. Quando uma seca extrema ou severa está instalada é preciso esperar mais tempo para sair dela no caso da fase positiva da NAO do que no caso da fase positiva da AO.

## 5.6 Tempo médio para a primeira visita ao estado j

Na Tabela 5.15 e Tabela 5.16 são apresentados, respectivamente, os tempos de recorrência e os tempos médios de chegada à classe de "Não-seca", partindo, num dado mês da qualquer uma das classes de seca.

Tabela 5.15 Tempo de recorrência (meses) para as várias classes de seca nos casos de SPI-1 e SPI-12, para a região do Alentejo.

Modelo \ Classe de seca		Classe de seca			
		N	1	2	3
SPI-1	MARKOV-SPI	1.90	3.04	12.43	15.68
	MARKOV-SPI-NAO	1.91	3.02	12.44	15.63
	MARKOV-SPI-AO	1.90	3.04	12.41	15.64
	MARKOV-SPI-SOI	1.89	3.06	12.26	15.77
SPI-12	MARKOV-SPI	2.13	2.76	11.75	12.28
	MARKOV-SPI-NAO	2.07	2.71	12.13	15.32
	MARKOV-SPI-AO	2.11	2.73	12.07	13.13
	MARKOV-SPI-SOI	2.12	2.68	12.59	13.09

Os tempos de recorrência para o caso do SPI-1 vão aumentando com o aumento da severidade de cada classe e não diferem muito para cada um dos modelos. As maiores diferenças são entre as escalas de tempo do SPI, para as classes 1 e 3. Para o SPI-12, comparativamente ao SPI-1 e de uma forma geral, é preciso esperar menos tempo para se voltar a estar numa classe de seca (1, 2 ou 3), nomeadamente no caso da classe "seca severa ou extrema" em que a diferença é aproximadamente de 3 meses, exceptuando no caso da NAO em que os valores são idênticos. No caso do SPI-12, a classe de "Não-seca" demora mais tempo a ser visitada do que no caso do SPI-1. Como seria de esperar em ambas as escalas de tempo a classe de "Não-seca" tem o menor valor de recorrência que todas as outras classes.

Tabela 5.16 Tempo médio de chegada à classe de "Não-seca" (N), partindo num dado mês de uma classe de seca i (1, 2 ou 3), nos casos de SPI-1 e SPI-12, para a região do Alentejo.

Modelo \ Classe de seca		Classe de seca			
		1	2	3	
SPI-1	MARKOV-SPI	1.99	2.23	2.52	
	MARKOV-SPI-NAO	NAO+	2.28	2.55	3.20
		NAO-	1.76	1.97	2.11
	MARKOV-SPI-AO	AO+	2.54	2.73	3.15
		AO-	1.69	1.93	2.12
	MARKOV-SPI-SOI	SOI+	2.01	2.21	2.33
SOI-		1.97	2.25	2.67	
SPI-12	MARKOV-SPI	13.75	17.31	21.25	
	MARKOV-SPI-NAO	NAO+	23.32	35.20	41.95
		NAO-	10.87	10.97	13.62
	MARKOV-SPI-AO	AO+	27.15	38.60	43.35
		AO-	9.84	11.01	14.82
	MARKOV-SPI-SOI	SOI+	12.86	19.26	24.76
SOI-		15.53	18.17	21.57	

Da tabela anterior é possível notar de imediato que os tempos médios de chegada à classe de "Não-seca" a partir de qualquer uma das classes de seca 1, 2 e 3, para os casos de SPI-1 e SPI-12, são completamente diferentes. Isto prende-se com o facto já mencionado das escalas de tempo do SPI. O SPI-12 (anual) tem informação da quantidade de precipitação dos últimos 12 meses e portanto para este caso é preciso muito mais tempo para recuperar de uma situação de seca do que no caso do SPI-1 (mensal) e daí a diferença de valores ser tão acentuada. Neste caso, o SPI-12 dá-nos uma informação mais verdadeira do tempo que é preciso esperar

para a recuperação de uma seca meteorológica. Como seria de esperar em ambas as escalas de tempo do SPI, o tempo de chegada à classe de “Não-seca” aumenta consoante a severidade da seca presente, sendo por isso que no caso da “Seca severa ou extrema” (3), o tempo de recuperação chega a ser de cerca 43 e 41 meses (respectivamente, os casos das fases positivas da AO permanente e da NAO permanente). Para os casos da NAO e AO é verifica-se para todas as 3 classes de severidade, que é preciso muito mais tempo para se chegar à classe de “Não-seca” aquando das suas fases positivas, do que nas suas fases negativas. Para o caso da fase positiva do SOI é necessário esperar menos tempo para chegar à classe de “Não-seca” partindo da classe 1, do que no caso da sua fase negativa. É ainda notória a grande diferença de valores de tempos, entre o modelo MARKOV-SPI e os restantes modelos com a inclusão do índice de CGA.

O tempo médio de chegada à classe de “Não-seca” partindo de uma qualquer classe de seca, pode ser interpretado como o tempo de recuperação da seca.

## 5.7 Previsões a curto prazo

A Tabela 5.17 e a Tabela 5.18 mostram as previsões mais prováveis dentro do prazo de 1 a 3 meses e a respectiva probabilidade de ocorrência para o SPI-1 e SPI-12, respectivamente.

Tabela 5.17 Previsões a curto prazo nos casos de SPI-1, nos vários modelos, para a região do Alentejo.

Modelo	Estado Inicial	Classe mais provável 1 mês depois		Classe mais provável 2 meses depois		Classe mais provável 3 meses depois	
		Estado	Probabilidade	Estado	Probabilidade	Estado	Probabilidade
MARKOV-SPI	N	N	0.5719	N	0.5322	N	0.5275
	1	N	0.5279	N	0.5277	N	0.5272
	2	1	0.4375	N	0.5168	N	0.5254
	3	1	0.4211	N	0.4917	N	0.5226
MARKOV-SPI-NAO	N	N	0.4430	N	0.4250	N	0.4222
	1	N	0.4811	N	0.4271	N	0.4222
	2	1	0.4364	N	0.4112	N	0.4214
	3	1	0.5625	N	0.3972	N	0.4192
MARKOV-SPI-NAO	N	N	0.6959	N	0.6515	N	0.6443
	1	N	0.5824	N	0.6337	N	0.6417
	2	N	0.4634	N	0.6208	N	0.6392
	3	N	0.4091	N	0.5915	N	0.6348
MARKOV-SPI-AO	N	N	0.4225	1	0.4145	1	0.4155
	1	N	0.4186	1	0.4148	1	0.4153
	2	1	0.4783	1	0.4138	1	0.4155
	3	1	0.5000	1	0.4247	1	0.4161
MARKOV-SPI-AO	N	N	0.6959	N	0.6578	N	0.6513
	1	N	0.6126	N	0.6436	N	0.6491
	2	N	0.4800	N	0.6257	N	0.6450
	3	N	0.3889	N	0.5980	N	0.6411
MARKOV-SPI-SOI	N	N	0.5594	N	0.5259	N	0.5220
	1	N	0.5171	N	0.5192	N	0.5221
	2	1	0.5185	N	0.5238	N	0.5202
	3	N	0.3750	N	0.4966	N	0.5214
MARKOV-SPI-SOI	N	N	0.5850	N	0.5405	N	0.5341
	1	N	0.5397	N	0.5382	N	0.5336
	2	N	0.4286	N	0.5039	N	0.5281
	3	1	0.5000	N	0.4831	N	0.5256

Pela análise da Tabela 5.17 é possível verificar que na grande maioria dos casos partindo de uma qualquer classe de severidade de seca é possível recuperar no espaço de 2 ou 3 meses ao estado de "Não-seca", com excepção de estarmos na presença de uma fase positiva da AO, na qual apenas se consegue chegar à classe (1), "seca ligeira". Isto significa que uma fase positiva permanente da AO é de todas as situações aqui presentes, a mais difícil para a recuperação de uma seca meteorológica. Para o espaço de tempo de 1 mês é possível chegar ao estado de "Não-seca" nas fases negativas da AO e NAO, a partir de qualquer estado de severidade de seca inicial. Para o caso das fases positivas da AO e NAO e no caso do modelo sem índice de CGA, se o estado inicial for 2 ou 3 (respectivamente "seca moderada" e "seca severa ou extrema"), apenas é possível recuperar para o estado 1 ("seca ligeira"). Todos estes resultados são baseados no SPI-1 mensal, que é bastante sensível às alterações de precipitação. Na prática a recuperação de uma seca leva na maior parte dos casos um tempo muito maior de recuperação, dependendo claro da severidade de uma dada seca. No geral, as probabilidades de se atingir uma determinada classe de severidade de seca para o SPI-1, não são muito elevadas (entre 0.38 e os 0.70) quando comparadas com as do SPI-12, da tabela seguinte.

**Tabela 5.18 Previsões a curto prazo nos casos de SPI-12, nos vários modelos, para a região do Alentejo.**

Modelo	Estado Inicial	Classe mais provável 1 mês depois		Classe mais provável 2 meses depois		Classe mais provável 3 meses depois		
		Estado	Probabilidade	Estado	Probabilidade	Estado	Probabilidade	
MARKOV-SPI	N	N	0.9190	N	0.8521	N	0.7965	
	1	1	0.8273	1	0.7115	1	0.6286	
	2	2	0.4898	1	0.3837	1	0.4146	
	3	3	0.7674	3	0.6276	3	0.5312	
MARKOV-SPI-NAO	NAO+	N	N	0.9111	N	0.8381	N	0.7775
		1	1	0.8036	1	0.6731	1	0.5804
		2	2	0.4828	3	0.3725	3	0.4038
	3	3	0.8519	3	0.7665	3	0.7081	
NAO-	N	N	0.9262	N	0.8646	N	0.8134	
	1	1	0.8519	1	0.7519	1	0.6796	
	2	2	0.5000	1	0.4837	1	0.5282	
	3	3	0.6250	3	0.4062	1	0.3242	
MARKOV-SPI-AO	AO+	N	N	0.8992	N	0.8161	N	0.7471
		1	1	0.8165	1	0.6915	1	0.6013
		2	2	0.5000	3	0.3799	3	0.4008
	3	3	0.7895	3	0.6847	3	0.6205	
AO-	N	N	0.9333	N	0.8783	N	0.8328	
	1	1	0.8378	1	0.7286	1	0.6494	
	2	2	0.4800	1	0.4831	1	0.5149	
	3	3	0.7500	3	0.5792	3	0.4551	
MARKOV-SPI-SOI	SOI+	N	N	0.9291	N	0.8702	N	0.8212
		1	1	0.8448	1	0.7284	1	0.6388
		2	2	0.5625	2	0.3647	3	0.3269
	3	3	0.8182	3	0.7092	3	0.6352	
SOI-	N	N	0.9091	N	0.8343	N	0.7720	
	1	1	0.8077	1	0.7068	1	0.6413	
	2	1	0.5294	1	0.6189	1	0.6136	
	3	3	0.7358	3	0.5688	3	0.4514	

Comparando as duas tabelas anteriores constata-se que os valores das probabilidades de se atingir uma determinada classe de severidade de seca são bastante mais elevadas que no caso do SPI-1. Verifica-se ainda que é impossível para qualquer dos modelos, atingir num espaço de 3 meses um estado de "Não-seca" partindo de qualquer um dos estados de seca 1, 2 ou 3. No caso do modelo MARKOV-SPI e das fases negativas dos índices da AO e NAO, para o primeiro mês não há alteração das classes mais prováveis atingidas e para o segundo e terceiro mês não há alteração das classes à excepção da classe "seca moderada" (2) que é mitigada para "seca ligeira" (1). A situação é idêntica no caso da fase negativa do SOI, com a diferença de haver recuperação da classe de seca 2 para a 1 em todos os 3 meses partindo da classe de seca 2. No caso das fases positivas permanentes da AO e NAO não há recuperação de seca em nenhum caso, havendo possibilidade de agravamento da classe 2 para a 3, dentro do espaço de tempo de dois a três meses. No caso da fase positiva do SOI só é provável haver agravamento da classe 2 para a 3 após o terceiro mês.

Nesta secção foram apenas apresentados os resultados da formulação homogénea para o vector de equilíbrio, tempo médio de permanência ininterrupta no estado  $i$ , tempo médio para a primeira visita ao estado  $j$ , e previsões a curto prazo, uma vez que há muitas transições de estado que não se verificam quando analisadas mensalmente (formulação não-homogénea). Assim, as matrizes correspondentes às transições têm várias entradas nulas, o que impede de haver um resultado final, devido ao facto destas matrizes corresponderem um sistema de equações possível mas indeterminado. Os dados manipulados provenientes das reanálises do NCEP não foram apresentados, uma vez que os resultados têm uma qualidade bastante inferior aos resultados apresentados, obtidos a partir dos dados do INAG. Os dados das reanálises foram trabalhados a partir de um único ponto localizado na intersecção das fronteiras entre Portugal e Espanha e entre Alentejo e Algarve (ver Figura 4.2), e por isso os campos das várias variáveis são suavizados devido à baixa resolução dos mesmos ( $2.5^\circ$  lat x  $2.5^\circ$  long).

## 6. Conclusões

A modelação de secas através de cadeias de Markov torna-se bastante prática usando o índice SPI, já que este, além de sua flexibilidade que lhe é inerente por poder ser calculado para várias escalas temporais, depende apenas de dados de precipitação acumulada. Este índice, permite não só monitorizar a precipitação, bem como detectar as suas anomalias, e o início e estabelecimento de uma seca, que por ser a maior parte das vezes lento, é de difícil percepção. Permite também, monitorizar o desenvolvimento e dissipação da seca. O facto de poder ser calculado para várias escalas temporais permite uma leitura adequada a cada situação e ao objectivo da sua implementação.

Existem algumas dificuldades na aplicação das cadeias de Markov ao fenómeno da seca no que diz respeito à determinação do número de estados a utilizar e à atribuição de valores a esses intervalos. Contudo, foi utilizada a discretização de classes de SPI da Tabela 4.2, à semelhança do que foi usado em vários trabalhos de Paulo e Pereira (2005, 2006 e 2007) e Paulo (2004).

Na secção dos resultados é possível verificar que existe uma correlação significativa entre o SPI e todos os índices de CGA utilizados, correlação essa que depende de caso para caso, não só para certos períodos do ano, bem como em relação às fases positivas ou negativas desses mesmos índices. Para as séries de dados usadas a AO mostrou ter a correlação mais forte com o SPI, em relação aos outros dois índices de CGA usados. Foi possível verificar, por exemplo, que a NAO tem uma correlação bastante considerável com o SPI durante os meses de Inverno, pelo que é preciso estudar caso a caso, os períodos do ano em que cada preditor poderá ser mais útil e ter mais performance preditiva. As fases positivas da NAO e AO e a fase negativa do SOI favorecem a escassez de precipitação no Alentejo (e na Península Ibérica), e deste modo propiciam situações de seca, enquanto que as fases negativas da AO e NAO e a fase positiva do SOI favorecem abundância de precipitação no Alentejo, e por isso, menor probabilidade de seca. As matrizes de transição dos modelos de Markov estão de acordo com o esperado, para cada fase dos índices de CGA e mês dos índices SPI e de CGA.

É possível verificar que, de uma forma geral, a utilização dos índices de CGA melhoram um pouco a qualidade das previsões das classes de seca de 1 a 3 meses, e que para as séries dos índices de CGA usadas o SOI e a AO obtiveram as melhores performances. Há grandes diferenças aquando do estudo com um SPI-1 (mensal) ou SPI-12, havendo melhores resultados no caso do último, devido à memória dos 12 meses anteriores ao referente mês.

O tempo de residência esperado numa dada classe de seca é um bom indicador da persistência dessa classe de categoria de seca e é bastante afectado pelas fases positivas e negativas dos índices de CGA e também pela escala temporal do índice SPI.

O tempo médio de chegada à classe N ("Não-seca") poderá ser interpretado como o tempo esperado de recuperação da uma qualquer classe de severidade de seca.

Tendo em conta os resultados das previsões a curto prazo (1 a 3 meses), durante as fases positivas da AO, NAO e SOI permanentes há maior tendência a agravar a seca do a mitigá-la. AO invés, as fases negativas destes índices mostraram alguma contribuição na recuperação da seca. Estes resultados mostram também que não é provável que haja recuperação de uma seca com uma determinada severidade (1,

2 ou 3) num espaço de três meses. A recuperação é assim bastante mais lenta como se pode ver pelos resultados do tempo médio de chegada à classe N ("Não-seca").

A utilização da formulação não-homogénea é mais vantajosa em relação à formulação homogénea, já que a primeira depende do ciclo anual da precipitação, dando assim melhores resultados consoante o mês em estudo.

Como trabalho futuro há que se estudar em maior detalhe os períodos do ano em que cada preditor está mais correlacionado com o SPI do local em estudo, no sentido de tirar o melhor partido destes preditores. Assim, seria possível usar mais do que um preditor ao mesmo tempo no mesmo modelo. É possível ainda, testar se há melhorias nos resultados e na performance da previsão, ao ser aumentado o número de estado para os índices de CGA, o que também implica maiores recursos computacionais, devido ao aumento de matrizes de transição do índice SPI condicionada a cada intervalo de valores dos preditores.

## **7. Referências**

- AMS, (1997): *Meteorological drought - Policy statement*. Bulletin of the American Meteorological Society, **78**, 847-849.
- Banik P., Mandal A., Rahman M. S., (2002): *Markov chain analysis of weekly rainfall data in determining drought-proneness*. Discrete Dynamics in Nature and Society. Volume 7, Issue 4, pp. 231-239
- Barnston, A.G., R.E. Livezey, e M.S. Halpert, (1991): *Modulation of Southern Oscillation-Northern Hemisphere mid-winter climate relationships by the QBO*. J. of Climate. **4**, 203-217.
- Cancelliere, A., Salas J. D., (2004): *Drought length properties for periodic-stochastic hydrologic data*. Water Resour. Res., **40**, WO2503, 13 pp.
- Chelliah M., Bell G. D., (2005): *Leading Tropical Modes Associated with Interannual and Multidecadal Fluctuations in North Atlantic Hurricane Activity*, Journal of Climate, **19**, pp. 590-612.
- Çinlar, E., (1975): *Introduction to Stochastic Processes*. Prentice-Hall, Inc., New Jersey.
- Doob, J. L. (1953): *Stochastic Processes*. New York: John Wiley and Sons. ISBN 0-471-52369-0.
- Dutra E., (2006): *Hidrologia da Península Ibérica: Algumas aplicações*, Trabalho de estágio da Licenciatura em Meteorologia Oceanografia e Geofísica (variante Meteorologia).
- Ferreira, E. M. N. O., (2004): *Aplicação de modelos Log-Lineares ao estudo das probabilidades de transição entre classes de seca na região do Alentejo*. Tese de Mestrado, Instituto Superior de Agronomia, Universidade Técnica de Lisboa.
- Fonseca, M. J., Coelho, C. A., (1998): *Aplicação de modelos Log-Lineares ao estudo de Tabelas de Contingência de Estrutura Diagonal*. In: *Estatística: a diversidade na unidade*. Actas do V Congresso Anual da Sociedade Portuguesa de Estatística, pp. 437-446.
- GPPAA (2000): *Agricultura no Alentejo: principais indicadores*. Min. Da Agricultura, Desenvolvimento Rural e Pescas.
- Haan, C. T. (1997): *Statistical Methods in Hydrology*. Iowa State Univ. Press.
- Hurrell, J.W., (1995): *Decadal Trends in the North Atlantic Oscillation: Regional Temperatures and Precipitation*. Science: Vol. **269**, pp.676-679.
- Jolliffe I. T., Stephenson D.B., (2003): *Forecast Verification – A Practitioner’s Guide in Atmospheric Science*, Wiley, pp. 45 – 84.
- Kondrashov D., Ghil M., Ide K., Todling R., (2002): *Data Assimilation and Weather Regimes in a Three-Level Quasi-Geostrophic Model*.

- Kondrashov D., Ide K., Ghil M., (2003): *Weather Regimes and Preferred Transition Paths in a Three-Level Quasigeostrophic Model*, J. Atmos. Sci., **61**, 568 – 587.
- Kondrashov D., Shen J., Berk R., Andrea F., Ghil M., (2007): *Predicting Weather Regime Transitions in Northern Hemisphere Datasets*, Clim. Dyn., **29**, 535 – 551.
- Lamb, P., R. Pepler, (1991): *West Africa. Chapter 5. In Teleconnections: Linkages between ENSO, Worldwide Climate Anomalies, and Societal Impacts*. M.H. Glantz. R.W. Katz, N. Nicholls, Eds., Cambridge University Press, 121-189.
- Lohani V. K., Loganathan, G. V. (1997): *An Early Warning System for Drought Management Using The Palmer Drought Index*, J. Amer. Water Resour. Assoc., (JAWRA), 33(6), 1375-1386.
- Lohani, V. K., Loganathan. G. V.; Mostaghimi, S., (1998): *Long-term Analysis and Short-Term Forecasting of Dry Spells by Palmer Drought Severity Index*. Nordic Hydrology, 29(1), pp. 21-40.
- Marshall J., Kushnir Y., Battisti D., Chang P., Czaja A., Dickson R., Hurrell J., McCartney M., Saravanan R., Visbeck M., (2001): *North Atlantic Climate Variability: Phenomena, Impacts and Mechanisms*, Int. J. Climatol. **21**: 1863–1898
- Mckee, T. B., Doesken N. J., Kleist J., 1993: The relationship of drought frequency and duration to time scales. *Eight Conference on Applied Climatology*, Boston, pp. 179 – 184.
- Mo, K. C., Livezey R. E., (1986): *Tropical-extratropical geopotential height teleconnections during the Northern Hemisphere winter*. Mon. Wea. Rev., **114**, 2488-2515.
- Ochola, W.O., Kerkides, P., (2003): *A markov chain simulation model for predicting the critical wet and dry spells in Kenya: analyzing rainfall events in Kano plains*, Irrigation and Drainage, 52 (4), 327-342.
- Paulo, A.A., (2004): *Metodologias de Caracterização e Previsão das Secas – Modelação Espacial e Temporal*, Doutoramento em Engenharia Agrónómica, Univer. Técnica de Lisboa – Instituto Superior de Agronomia, Lisboa, 2004.
- Paulo, A.A., Ferreira E., Coelho, C., Pereira, L. S., (2005): *Drought Class Transition Analysis Through Markov and Loglinear models, an Approach to Early Warnings*, Agricultural Water Management, pp. 59 – 81.
- Paulo, A. A., L. S. Pereira, e P. G. Matias, (2003): *Analysis of local and regional droughts in southern Portugal using the theory of runs and standardized precipitation index*. Tools for Drought Mitigation in Mediterranean Regions, Kluwer, 55-78 pp.
- Paulo A.A., Pereira L. S., (2006): *Prediction of SPI Drought Class Transition Using Markov Chains*, Water Resour Manag, 1813 – 1827.
- Paulo A.A., Pereira L. S., (2007): *Stochastic Prediction of Drought Class Transition*, Water Resour Manag, **22**, 1277 – 1296.
- Pereira L. S., Cordery I., Iacovides I., (2002): *Coping with Water Scarcity*, UNESCO IHP VI, Technical Documents in Hydrology No. **58**, UNESCO, Paris, 269 p. (disponível online: <http://unesdoc.unesco.org/images/0012/001278/127846e.pdf>).

Pires C., Perdigão R., (2007): *Non-Gaussianity and Asymmetry of the Winter Monthly Precipitation Estimation from the NAO*, *Monthly Weather Review*, **135**, pp. 430–448.

Pires, V., (2003). *Frequência e Intensidade de Fenómenos meteorológicos extremos associados a precipitação*. Dissertação para a obtenção do grau de Mestre em Ciências e Engenharia da Terra, Lisboa, pp.98.

Rocha A., (1999): *Low-frequency variability of seasonal rainfall over the Iberian Peninsula and ENSO*, *Int. J. Climatol.*, **19**, 889 – 901.

Rogers, J.C. and H. van Loon, 1979: *The Seesaw in Winter Temperatures Between Greenland and Northern Europe. Part II: Some Oceanic and Atmospheric Effects in Middle and High Latitudes*. *Mon. Wea. Rev.*, **107**, 509-519.

Steinemann, A. (2003): *Drought indicators and triggers: a stochastic approach to evaluation*. *J. Amer. Water Resour. Assoc. (JAWRA)* 39(5), pp. 1217 – 1233.

Stephenson D. B., Pavan V., Bojariu R., (2000): *Is the Atlantic Oscillation a random walk?*, *Int. J. Climatol.*, **20**, 1 – 18.

Thompson D. W. J., Wallace J. M., (1998): *The Arctic Oscillation signature in the wintertime geopotential height and temperature fields*. *Geophysics Research Letters*, **25**: 1297–1300.

Thompson D. W. J., Wallace J. M., Hegerl G. C., (2000a): *Annular modes in the extratropical circulation: part I: month to month variability*. *Journal of Climate*, **13**: 1000–1016.

Thompson D. W. J., Wallace J. M., Hegerl G. C., (2000b): *Annular modes in the extratropical circulation: part II: trends*. *Journal of Climate*, **13**: 1018–1036.

Van Loon, H., J.C. Rogers, (1978): *The Seesaw in Winter Temperatures Between Greenland and Northern Europe. Part I: General Description*. *Mon. Wea. Rev.*, **106**, 296-310.

Vicente-Serrano, S. M. (2005): *El Niño and La Niña influence on droughts at different timescales in the Iberian Peninsula*, *Water Resour. Res.*, **41**, W12415, doi: 10.1029/2004WR003908.

Walker, G. T., E. W. Bliss, (1932): *World Weather*, V. *Mem. Roy. Meteor. Soc.*, 4, No. **36**, 53-84.

Wallace, J. M., Gutzler D. S., (1981): *Teleconnections in the geopotential height field during the Northern Hemisphere Winter*. *Mon. Wea. Rev.*, **109**, 784-812.

Wallace J. M., (2000): *North Atlantic Oscillation/Annular Mode: Two paradigms? One Phenomenon*. *Quart., J. Royal Met. Soc.*, **126**: 791–805.

Wilks, D. S., (2006): *Statistical Methods in the Atmospheric Sciences*, Second Edition, Academic Press.

Vlachos, E., James, L.D., (1993): *Drought Impacts*. In: V. Yevjevich, L. V. Cunha, E. Vlachos (Eds.) *Coping With Droughts*. Water Resources Publications, Littleton, CO, pp. 44-73.

WMO, (1975): *Drought and agriculture*, 127 pp.

<http://www.cpc.ncep.noaa.gov/>

<http://www.drought.unl.edu/>

<http://jisao.washington.edu/>

<http://www.meteo.pt/>

<http://nsidc.org/index.html> (National and Snow Data Center)

<http://www.noaa.gov/>

<http://www.esrl.noaa.gov/psd/people/cathy.smith/best/details.html>

[http://gcmd.nasa.gov/records/GCMD\\_NOAA\\_NWS\\_CPC\\_BEST.html](http://gcmd.nasa.gov/records/GCMD_NOAA_NWS_CPC_BEST.html)