

A forma mais simples desse funcional é a que simula um quadrado Latino de 2 símbolos, no qual se observa uma correlação máxima entre o produto de duas variáveis e a terceira:  $cor(Y_1Y_2, Y_3)$ , a qual é nula se a distribuição conjunta for Gaussiana. Dado que as variáveis são normalizadas, aproximamos essa correlação pela média  $F(Y_1, Y_2, Y_3)[\mathbf{R}] = E(Y_1Y_2Y_3)$ .

O máximo é obtido pelo método Quase-Newton de descida contra o gradiente de  $-E(Y_1Y_2Y_3)$  em relação aos  $N_{ang}$  ângulos a partir de um primeiro *guess* desses vetores. Dada a existência de vários máximos locais de  $E(Y_1Y_2Y_3)$ , alguns deles triviais, devido às simetrias de  $F(Y_1, Y_2, Y_3)[\mathbf{R}]$  em relação às variáveis independentes  $Y_1, Y_2, Y_3$ , testaram-se vários primeiros *guesses* da maximização. Foi então tomado o máximo dos máximos locais de  $F(Y_1, Y_2, Y_3)[\mathbf{R}]$  como a melhor triada, a chamada triada dominante.

Na figura 6, mostram-se os valores de  $cor(Y_1Y_2, Y_3)$ ,  $I(Y_1, Y_2, Y_3)$ ,  $I_2(Y_1, Y_2, Y_3)$  e  $I_3(Y_1, Y_2, Y_3)$  para as triadas dominantes para cada valor de  $N_{rot}$  no intervalo [3,20]. Como esperado, à medida que aumenta a dimensão  $N_{rot}$  do espaço de procura, aumenta  $cor(Y_1Y_2, Y_3)$  e  $I_3(Y_1, Y_2, Y_3)$  e em consequência também  $I(Y_1, Y_2, Y_3)$ . Os valores de  $I_2(Y_1, Y_2, Y_3)$  não variam significativamente, podendo até pontualmente diminuir. Nota-se que quando são incluídas certas PCs no espaço de procura (e.g. PC<sub>7</sub> e PC<sub>8</sub>), o valor do funcional acusa uma variação significativa. Na figura 6,  $Y_1, Y_2, Y_3$  são representados por  $X, Y, Z$ .

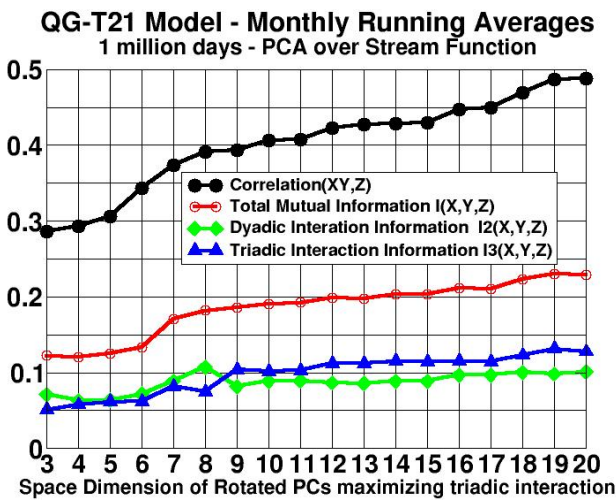


Figura 6 – Valores de  $cor(Y_1Y_2, Y_3)$ ,  $I(Y_1, Y_2, Y_3)$ ,  $I_2(Y_1, Y_2, Y_3)$  e  $I_3(Y_1, Y_2, Y_3)$  para as triadas dominantes para cada valor de  $N_{rot}$  no intervalo [3,20] para o modelo quase-geostrófico.

Tal como na Figura 1, mostra-se na Figura 7, a iso-superfície 0.001 da PDF conjunta referente à triada dominante para  $N_{rot}=8$  e  $N_{rot}=20$ . Nota-se uma semelhança com a forma da PDF da triada perfeita da Figura 1, consistindo esta numa maneira de visualizar uma triada através da PDF tridimensional.

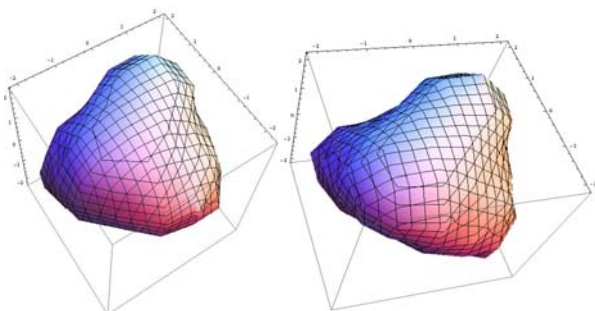


Figura 7 – Iso-superfície  $\rho=0.001$  da PDF correspondente às triadas dominantes em espaços de dimensão  $N_{rot}=8$  e  $N_{rot}=20$ , à esquerda e direita respetivamente.

## 5 TRIADAS DOMINANTES NO CAMPO DA VARIABILIDADE INTERANUAL DA TEMPERATURA DO OCEANO GLOBAL

No sentido de encontrar teleconexões triádicas no oceano global, procedeu-se à utilização do método explicado na secção 4 e aplicou-se ao campo da temperatura da superfície do mar (SST). Para tal recorreu-se à base dados mensais da SST do *Goddard Institute for Space Studies*: NASA GISS no período 1880-2012. Consideraram-se apenas as séries com a totalidade de dados sem falhas neste período. De seguida calcularam-se médias móveis com um ano de amplitude e retirou-se a tendência linear de todas as séries pontuais. Executou-se uma PCA e calcularam-se PCs. A fração acumulada de variância explicada surge na Figura 8 em que as primeiras 10 PCs explicam cerca de 65% da variância total. O primeiro modo (EOF) é fortemente projetado no ‘El Niño’, os modos seguintes até ao décimo projetam-se nas bacias Pacífica Norte e Sul e Atlântica Norte e Sul, sendo visíveis assinaturas do padrão da Oscilação Multidecadal do Atlântico e Pacífico (não mostrado).

GISS SST Annual Running Averages  
1880-2012 PCA over Global Ocean

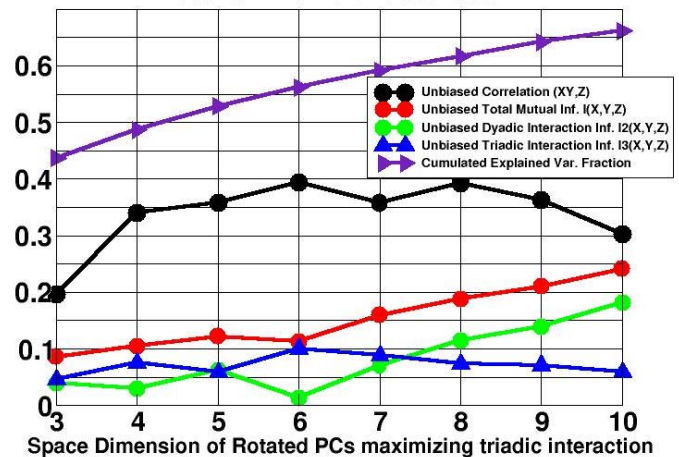


Figura 8 – Valores de  $cor(Y_1Y_2, Y_3)$ ,  $I(Y_1, Y_2, Y_3)$ ,  $I_2(Y_1, Y_2, Y_3)$  e  $I_3(Y_1, Y_2, Y_3)$  para as triadas dominantes para cada valor de  $N_{rot}$  no intervalo [3,10] para o campo da SST global.

Neste caso, o tamanho das séries é da ordem de 650, mas com um número de graus de liberdade temporais da ordem ou inferior a  $N_{dof}=150$ , devido à auto-correlação serial, o que é muito menor que no caso das amostras no modelo quase-geostrófico. Por isso, todas as estatísticas tiveram de ser ajustadas com vista a ser-lhes retirados os vieses positivos devidos ao reduzido tamanho relativo das séries e ao efeito de ‘overfitting’ ou sobre-ajustamento devido à liberdade de escolha dos parâmetros, neste caso, ângulos de Euler. Foi considerado um valor equivalente  $N_{dof} = N_{dof} - N_{ang}$  de graus de liberdade.

Foram calculados, tal como para o modelo quase-geostrófico,  $cor(Y_1Y_2, Y_3)$ ,  $I(Y_1, Y_2, Y_3)$ ,  $I_2(Y_1, Y_2, Y_3)$  e  $I_3(Y_1, Y_2, Y_3)$  para as triadas dominantes no espaço das  $N_{rot}$  PCs normalizadas da SST, até  $N_{rot}=10$ , com resultados sintetizados na Figura 8. Os resultados são desviesados admitindo  $N_{dof}$  graus de liberdade. A triada estatisticamente mais robusta e que evidencia maior valor de  $cor(Y_1Y_2, Y_3)$  ajustado (~0.40), corresponde a  $N_{rot}=6$ . As séries temporais das correspondentes PCs rodadas vêm no gráfico da Figura 9. Notam-se como esperado as seguintes situações predominantes em termos de oscilações: 1) todas as variáveis na fase positiva ou 2) duas das variáveis em fase negativa e uma em oposição na fase positiva.

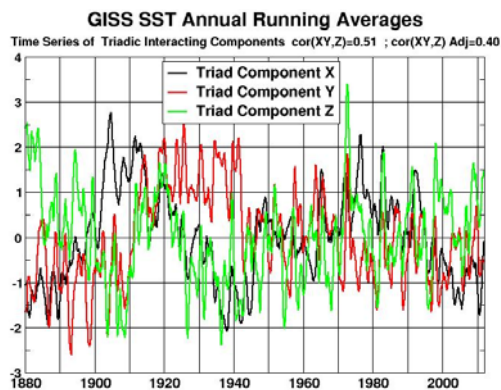


Figura 9 – Séries temporais das componentes da tríada dominante do campo da SST, mais robusta estatisticamente.

Os campos dos padrões espaciais que multiplicam as variáveis  $Y_1, Y_2, Y_3$  vêm mostrados na figura 10 e projetam-se especialmente em zonas de grandes correntes, principalmente as das correntes frias. Nos gráficos das Figuras 9 e 10, X,Y,Z correspondem a  $Y_1, Y_2, Y_3$  no texto.

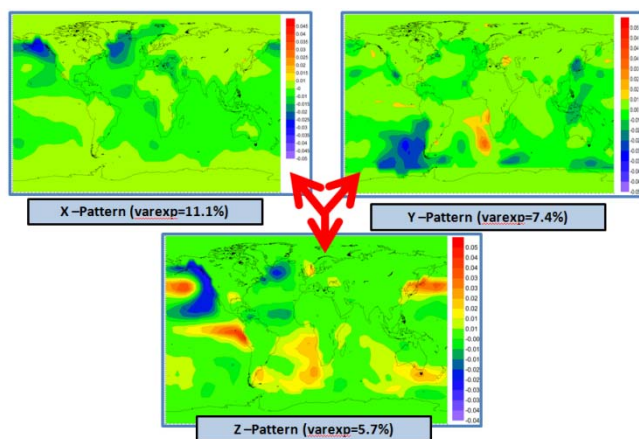


Figura 10 – Padrões espaciais das componentes da tríada dominante do campo da SST, mais robusta estatisticamente.

Os 3 padrões explicam a totalidade de 24% da variância, o que é significativo em termos da projeção da interação triádica na variabilidade da SST global. O padrão espacial da série  $Y_1$  está especialmente projetada nas zonas da corrente do Alaska e do Labrador; na zona das correntes frias de Benguela e Perú para o padrão da série  $Y_2$  e nas zonas da corrente da Califórnia e da corrente Equatorial Sul (zona do El Niño) para o padrão da série  $Y_3$ . Os padrões dominantes de fase temporal, associados à correlação positiva entre  $Y_1 Y_2$  e  $Y_3$  são: (+++), (+--), (-+-) e (-+-), o que é visível no gráfico das 3 séries (Figura 9). A explicação de um mecanismo possível desta interação resta ainda um pouco especulativo e sai fora do âmbito deste trabalho.

## 6 CONCLUSÕES

As variáveis que caracterizam a variabilidade lenta da atmosfera e oceano exibem distribuições de probabilidade multivariadas com valores baixos de não-Gaussianidade, quando comparados as de sistemas caóticos de baixo número de variáveis. A não-Gaussianidade manifesta-se especialmente nas escalas planetárias e temporais longas superiores a 1 mês.

A organização da variabilidade em distribuições não-Gaussianas é muito mais variada e ‘exótica’ que nas distribuições de probabilidade Gaussianas. Tal permite extrapolar o conceito de teleconexão para situações em que variáveis estão descorrelacionadas linearmente embora o estejam não linearmente, o que acontece nas chamadas díadas não-Gaussianas. Outra possibilidade é o surgimento de tríadas não-Gaussianas em que três variáveis descorrelacionadas ou mesmo

independentes duas a duas, podem ser dependentes quando vistas as três em simultâneo. Um exemplo é o da correlação entre o - produto de duas variáveis - e uma terceira variável, o que corresponde numa versão discretizada a uma relação do tipo Quadrado Latino de 2 símbolos. Este tipo de interatividade ou sinergia global é mesmo generalizável a mais variáveis (quartetos e quintetos) e é quantificável pela informação de interação que elimina na sua definição a contribuição das interações entre subconjuntos de variáveis.

No entanto a manifestação de sinergias com valores significativos em distribuições quase-Gaussianas de elevada dimensão exige métodos de otimização não linear no espaço das projeções (ou das rotações).

Desenvolveu-se por isso um método de otimização no espaço dos ângulos de Euler para maximizar uma função simuladora da informação de interação, neste caso a média do produto de 3 variáveis descorrelacionadas entre si.

O método foi aplicado a um modelo quase-geostrófico usado na simulação da variabilidade lenta da atmosfera no Inverno setentrional com escalas de tempo superiores a um mês ou mais bem como à variabilidade interanual da SST do oceano global. Registraram-se neste caso, tríadas dominantes com correlação triádica de ~40%. Os padrões de cada variável interveniente projetam-se em zonas, especialmente das correntes frias, o que indica a presença de variabilidade climática não-Gaussianas e não linear nessas zonas.

## 7 REFERÊNCIAS

- Bocquet, M.; Pires, C.; Lin, W. Beyond Gaussian Statistical Modeling in Geophysical Data Assimilation. *Mon. Wea. Rev.* 2010, 138, 2997-3023. DOI: 10.1175/2010MWR3164.
- Cover, T.M.; Thomas, J.A. Elements of Information Theory; John Wiley & Sons, Inc.: New York, NY, USA, 1991
- Donges, J.F.; Zou, Y.; Marwan N.; Kurths J. 2009. Complex networks in climate dynamics. *The European Physical Journal Special Topics.* Volume 174, Issue 1, pp 157-179
- Hyvärinen, A.; Oja, E. Independent Component Analysis: Algorithms and Application, *Neural Networks*, 2000, 13(4-5),411-430.
- Jakulin, A.; Bratko, I.; 2004. Quantifying and Visualizing Attribute Interactions: An Approach Based on Entropy. arXiv:cs/0308002v3 [cs.AI], 2 Mar 2004.
- Marshall J, Molteni F, 1993: “ Toward a Dynamical understanding of Planetary-Scale Flow Regimes”. *J. Atmos. Sci.*, 50, 1792-1818.
- McGilly WJ, 1954, “Multivariate information transmission”. *Psychometrika*, 19, 97-116.
- Peters, J. M.; Kravtsov S.; Schwartz, T. Predictability associated with nonlinear regimes in an atmospheric model. *J. Atmos. Sci.*, 2012, DOI: http://dx.doi.org/10.1175/JAS-D-11-0168.1
- Pires CA, Perdigão RAP, (2007): "Non-Gaussianity and asymmetry of the winter monthly precipitation estimation from the NAO". *Mon. Wea. Rev.* 135(2): 430-448.
- Pires, C.A.L.; Perdigão, R.A.P. Minimum Mutual Information and Non-Gaussianity Through the Maximum Entropy Method: Theory and Properties. *Entropy* 2012, 14, 1103-1126. doi:10.3390/e14061103 (http://www.mdpi.com/1099-4300/14/6/1103)
- Pires C.A.L (2013) Compreender o Clima, uma aventura pelos Paradigmas da Modelação in Matemática do Planeta Terra, Fernando Pestana da Costa, João Teixeira Pinto e Jorge Buescu (eds), IST Press, ISBN: 978-989-8481-26-9, 562 pp, 99-139
- Silverman, B.W. (1998). *Density Estimation for Statistics and Data Analysis*. London: Chapman & Hall/CRC. p. 48. ISBN 0-412-24620-1
- Stephenson, D. B.; Hannachi, A.; O'Neill, A. On the existence of multiple climate regimes. *Quart. J. Roy. Meteor. Soc.*, 2004, 130, 583–605.
- Strouine, K.; Kravtsov, S.; Kondrashov, D.; Ghil, M. Reduced models of atmospheric low-frequency variability: Parameter estimation and comparative performance. *Physica D* 2010, 29, Issue 3-4. 145-166 DOI: 10.1016/j.physd.2009.10.013.
- Sura P.; Newman, M.; Penland C.; Sardeshmuck, P. Multiplicative Noise and Non-Gaussianity: A Paradigm for Atmospheric Regimes? *J. Atmos. Sci.*, 2005, 62, 1391–1406.
- Stephenson, D. B.; Hannachi, A.; O'Neill, A. On the existence of multiple climate regimes. *Quart. J. Roy. Meteor. Soc.*, 2004, 130, 583–605.

Trabalho executado no âmbito do projeto FCT PTDC/GEO-MET/3476/2012: “Avaliação da Predictabilidade e hibridação de Previsões sazonais de seca na Europa Ocidental – PHDROUGHT”