

PREVISÕES DE TEMPERATURA DA SUPERFÍCIE DO MAR NO ATLÂNTICO TROPICAL USANDO ANÁLISES DE CORRELAÇÕES CANÔNICAS.

Rita Valéria Andreoli
Instituto Nacional de Pesquisas Espaciais
Av. dos Astronautas, 1758 - São José dos Campos, SP, 12227-010
valeria@cptec.inpe.br

Mary Toshie Kayano
Instituto Nacional de Pesquisas Espaciais
Av. dos Astronautas, 1758 - São José dos Campos, SP, 12227-010
mary@cptec.inpe.br

ABSTRACT

Statistical models to forecast the sea surface temperature anomalies (SSTA) in the Tropical Atlantic (TA) have been constructed using Canonical Correlation Analysis and the SSTA in the TA and/or Equatorial Pacific (EP) as predictors. The models with the predictors in the TA and EP have better skill in the Tropical North Atlantic (TNA) than in the Tropical South Atlantic (TSA) and Equatorial Atlantic (EA). The best 0-5 month forecasts for the TNA occurred from March to June. For the TSA and EA, only for the 0-month forecasts are good. Ours results suggested that monthly models are more appropriated to forecast the SSTA in the TA. The EP and TA as the predictor region should be for November to March months, and the TA for the others months.

1 - INTRODUÇÃO

Vários estudos mostram que a variabilidade sazonal de precipitação sobre o Nordeste do Brasil (NEB) é previsível com alguns meses de antecedência, usando uma combinação de anomalias de Temperatura da Superfície do Mar (TSM) das regiões do Pacífico e do Atlântico Tropical (AT) (Hastenrath, 1990; Hastenrath e Greishar, 1993). O desempenho dos modelos parece ser mais dependente das anomalias de TSM (ATSM) do AT do que das do Pacífico, o que aponta para a importância do AT na variabilidade climática do NEB. Apesar disso, pouca atenção tem sido dada à previsão das ATSM do AT.

O que tem sido usado atualmente para o AT são modelos que relacionam as variações de TSM do AT com as do Pacífico (Repelli e Nobre, 1998 e Peland e Matrosova, 1998). Peland e Matrosova (1998) usaram análise dos Padrões de Oscilação Principal (POP), para prever ATSM e mostraram que a previsibilidade de ATSM no AT Norte (ATN), para períodos de alguns meses a um ano, aumenta com os preditores sendo das ATSM da região tropical global em vez de ATSM somente do Atlântico. Esta melhor previsibilidade não foi notada sobre as regiões equatorial e do Atlântico Tropical Sul (ATS), onde a persistência foi um bom preditor.

Esses resultados concordam com os de Repelli e Nobre (1998) que utilizaram um modelo estatístico baseado em Análises de Correlações Canônicas (ACC), com uma pré-filtragem dos dados com Funções Ortogonais Empíricas (FOE). Eles mostraram que o modelo tem um bom desempenho para prever ATSM do AT para o período de Março-Abril-Maio (MAM), quando se utiliza os meses de setembro a fevereiro como preditores. O modelo teve melhor desempenho para prever ATSM no ATN e para o ATS, o desempenho foi melhor com ATSM persistidas. O modelo apresenta melhor desempenho para o ATS do que os campos persistidos somente para previsões com as condições iniciais de setembro.

As contribuições de forçantes externas ao AT na previsibilidade de TSM desta bacia foram estudadas por Ruiz de Elvira et al. (2000) com um modelo de previsão de TSM baseado na teoria "Bayesian" de previsão. Eles aplicaram o modelo para prever ATSM na região do Golfo da Guiné e do ATN. Os preditores foram ATSM das regiões do Pacífico (Niño 3) e Oceano Índico Tropical. Este modelo teve bom desempenho para previsões do ATN, para defasagens maiores que 4 meses com a região do Pacífico como preditora. Uma boa previsão para a região do Golfo da Guiné foi obtida com a inclusão das regiões do Oceano Índico Equatorial e Pacífico como preditoras.

Finalmente, o modelo baseado em ACC de Landman e Mason (2001), para prever ATSM do Atlântico Equatorial geralmente não teve um bom desempenho, e a previsão realizada a partir de anomalias persistidas foram melhores em muitos casos.

O presente estudo investiga a previsibilidade de ATSM no AT. Serão investigados os desempenhos de modelos estatísticos lineares para previsão de ATSM do AT considerando um período de 48 anos. As previsões serão comparadas com as observações e com ATSM persistidas. Serão usados modelos mensais ao invés de um modelo considerando todos os meses do ano.

2 – DADOS E METODOLOGIA

Os dados utilizados são de TSM do "Comprehensive Ocean Atmosphere Data Set" (COADS). Esses dados encontram-se documentados em da Silva et al. (1994) e consistem de campos globais médios mensais com uma resolução horizontal de 1° em latitude por 1° em longitude, para o período de janeiro de 1945 a dezembro de 1993.

Modelos de previsão de ATSM são construídos com a técnica de ACC, freqüentemente usada como uma técnica de previsão (Barnett e Preisendofer, 1987; Barnston e Ropelewski, 1992, Landman e Mason, 2001). A ACC é uma metodologia de estatística multivariada para determinar, a partir de dois conjuntos de dados, combinações lineares, chamadas "variáveis canônicas", de forma que a correlação entre estas variáveis seja maximizada. A teoria de ACC encontra-se extensivamente descrita na literatura e ocupa o topo da hierarquia dos modelos de regressão múltipla (Barnett e Preisendofer, 1987).

2.1 - O PROCEDIMENTO PARA ACC

As ATSM na região do AT (30°N - 30°S; 60°W - 20°E), campos do preditando (Z), são previstas utilizando a técnica de ACC. Os campos do preditor (Y) são de ATSM em regiões selecionadas do Atlântico e Pacífico. Ambos os campos de Y e Z estão em uma resolução espacial de 4° em latitude por 4° em longitude, o que não compromete a representação dos padrões oceânicos de grande escala.

O primeiro passo é a pré-filtragem dos campos de Y e Z calculando-se as FOE destes campos separadamente. O modelo é construído usando um número de autovetores que retêm em torno de 80% da variância original de cada campo e que varia de 7 a 13, dependendo da área considerada. Essa pré-filtragem faz com que a ACC focalize-se nos modos de variabilidade dominantes dentro de cada conjunto de dados de entrada, e em princípio os ruídos de pequena escala são filtrados. Este procedimento é o mesmo utilizado por Repelli e Nobre (1998)

Em seguida, constroi-se uma matriz de correlação cruzada entre as séries de componentes principais de Y e Z. As dimensões dessa matriz são então reduzidas ao número de modos retidos pelo preditor e preditando. Da transposta desta matriz, são obtidos os autovetores e autovalores canônicos para Y e Z.

A função temporal canônica do preditor é encontrada através das combinações lineares entre os autovetores canônicos e a série de componentes principais do preditor (utilizada na construção da matriz) para cada modo. O número de modos utilizados neste caso, é o menor número de modos do preditor e preditando, fornecidos pela FOE. O último passo é derivar a equação de regressão expressando os modos canônicos em termos de variáveis originais. Assim, define-se um coeficiente de regressão para converter a função temporal canônica do preditor em uma função temporal canônica do preditando. Finalmente, uma equação preditiva que relaciona os campos preditor e preditando é construída e a previsão pode ser realizada.

Como o interesse do trabalho é estudar a previsibilidade das ATSM na região do AT, três experimentos foram realizados considerando em todos eles os campos de ATSM no AT como preditando e os campos de ATSM em diferentes áreas como preditor. No primeiro experimento, os campos de Y são as ATSM, na mesma área definida para o preditando. Num segundo experimento, o preditor Y, corresponde a ATSM do AT e do Pacífico Equatorial (PE) (157,5°E - 84,5°W; 5,5°N - 6,5°S), e num um terceiro experimento, Y corresponde aos campos de ATSM somente da região do PE. O uso de preditores em diferentes áreas permite investigar a importância das influências remotas e locais na previsibilidade das ATSM do AT. A Tabela (2.1) sumaria estes experimentos e fornece as dimensões espaciais dos preditores e preditandos.

TABELA 2.1 - ÁREAS DOS CAMPOS DO PREDITOR E PREDITANDO, PARA OS EXPERIMENTOS 1, 2 E 3, RESPECTIVAMENTE.

	Preditor	Observações	Preditando	Observações
Exp1	AT	218	AT	218
Exp2	PE e AT	340	AT	218
Exp3	PE	120	AT	218

Para cada experimento são considerados os campos de ATSM dos meses de janeiro a dezembro, individualmente, como preditores e os campos predictandos são selecionados até os seis (6) meses seguintes. Isso permite identificar a(s) melhor(es) área(s) preditora(s) e a máxima antecedência com a qual se pode conseguir as melhores previsões com os modelos estatísticos. A definição da defasagem deste trabalho é exemplificada na Figura 2.1 para previsão de março de 1986. Nesta figura, os meses sublinhados representam os meses escolhidos como preditor e predictando.

Preditor		Preditando	
1985	1986	1986	Defasagem
set out nov dez	Jan <u>fev</u>	<u>Mar</u>	0 mês
set out nov dez	<u>Jan</u> fev	<u>Mar</u>	1 mês
set out nov <u>dez</u>	Jan fev	<u>Mar</u>	2 meses
set out <u>nov</u> dez	Jan fev	<u>Mar</u>	3 meses
set <u>out</u> nov dez	Jan fev	<u>mar</u>	4 meses
<u>set</u> out nov dez	Jan fev	<u>mar</u>	5 meses

Figura 2.1 - Esquema da defasagem para previsão.

2.2 - DESEMPENHO DOS MODELOS

O procedimento escolhido para testar o desempenho dos modelos estatísticos é o método de validação cruzada (Barnett e Preisendorfer, 1987). Para cada tempo η , onde $1 \leq \eta \leq n$, o η -ésimo par de campos do preditor e predictando é excluído, e os $n-1$ campos restantes, são usados para construir um modelo. A previsão do predictando no tempo η é então realizada com este modelo. Este procedimento é repetido n vezes, seqüencialmente omitindo-se um único par de campos do preditor e predictando. Isto resulta em séries de n previsões. No caso em questão, com 49 anos de observações, de 1945 a 1993, o ano de 1946 é excluído e a ACC é usada para desenvolver um modelo de previsão com os 48 anos restantes e assim, sucessivamente. O modelo excluindo-se o primeiro ano não é construído em vista das defasagens usadas nas previsões. Os dados do preditor, previamente reconstruídos com base nos autovetores retidos para a construção de cada modelo, para cada ano excluído, são então utilizados para gerar os campos dos predictandos para os respectivos anos excluídos.

Vale salientar que os campos previstos pelo modelo contêm 80% da variância total dos dados e que nenhum fator de ajuste é utilizado para a correção da magnitude das anomalias previstas. Assim, opta-se, em fazer a comparação considerando os campos observados reconstruídos a partir dos autovetores e autovalores utilizados na construção dos modelos. Para o conjunto de 48 previsões de um certo mês, para cada preditor e predictando, a verificação é feita calculando-se o coeficiente de correlação linear em cada ponto de grade entre as séries observada e prevista. Os resultados são apresentados como médias do coeficiente de correlação sobre as áreas do ATN (5,5°N, 21,5°N, 50°W e 20°W), ATS (20,5°S, 4,5°S, 20°W e 10°E) e Atlântico Equatorial (AE) (50°W - 0 e 5,5°N - 5,5°S).

Um teste de significância estatística de hipótese nula (t-student) é aplicado para calcular o nível de significância dos coeficientes de correlação. Nestes experimentos, o grau de liberdade é o número de anos utilizados para verificação (48 anos). Assim, somente correlações superiores a 0,3 e 0,4 são significantes ao nível de 95% e 99%, respectivamente.

Para finalizar, as previsões obtidas com os modelos estatísticos são comparadas com previsões obtidas pela persistência das ATSM no AT.

3 – RESULTADOS

Os resultados da validação cruzada para previsão de ATSM na região do AT, considerando os três experimentos e os coeficientes de correlação médios do ATN, ATS e AE para defasagens de 0 a 5 meses, são apresentados nas figuras abaixo.

Para o ATN (Figura 3.1), o exp2 fornece melhores previsões que o exp1 para os meses de março a junho e defasagens maiores que 1 mês, com máximas correlações em abril ou maio, dependendo da defasagem. Para abril e maio, é notável o melhor desempenho dos modelos estatísticos do exp2 em relação à persistência, para defasagens maiores que 2 meses. Isto não ocorre quando se considera os meses de abril e maio e o exp1, onde, a persistência apresenta correlações comparável aos dos modelos. As melhores previsões para os meses de março a junho sugerem que os meses de novembro a março são bons preditores. Boas previsões para os meses de abril a junho também são encontradas para o exp1 com março como preditor. Por outro lado, as piores previsões no exp2 não ocorrem em um período definido mas com o mês de outubro como preditor e com os meses de novembro, dezembro, janeiro e fevereiro como preditandos. Para o exp1, as piores previsões aparecem com o mês de dezembro como preditor e defasagens de 1 a 3 meses. Neste caso os menores valores de correlação ocorrem nos meses de fevereiro a abril. A importância da região do Pacífico Equatorial na previsibilidade do ATN é confirmada quando se utiliza somente essa região como preditora. Os coeficientes de correlação da validação cruzada para o exp3 são significantes somente para os meses de março a junho, exceto para defasagens de 1 mês.

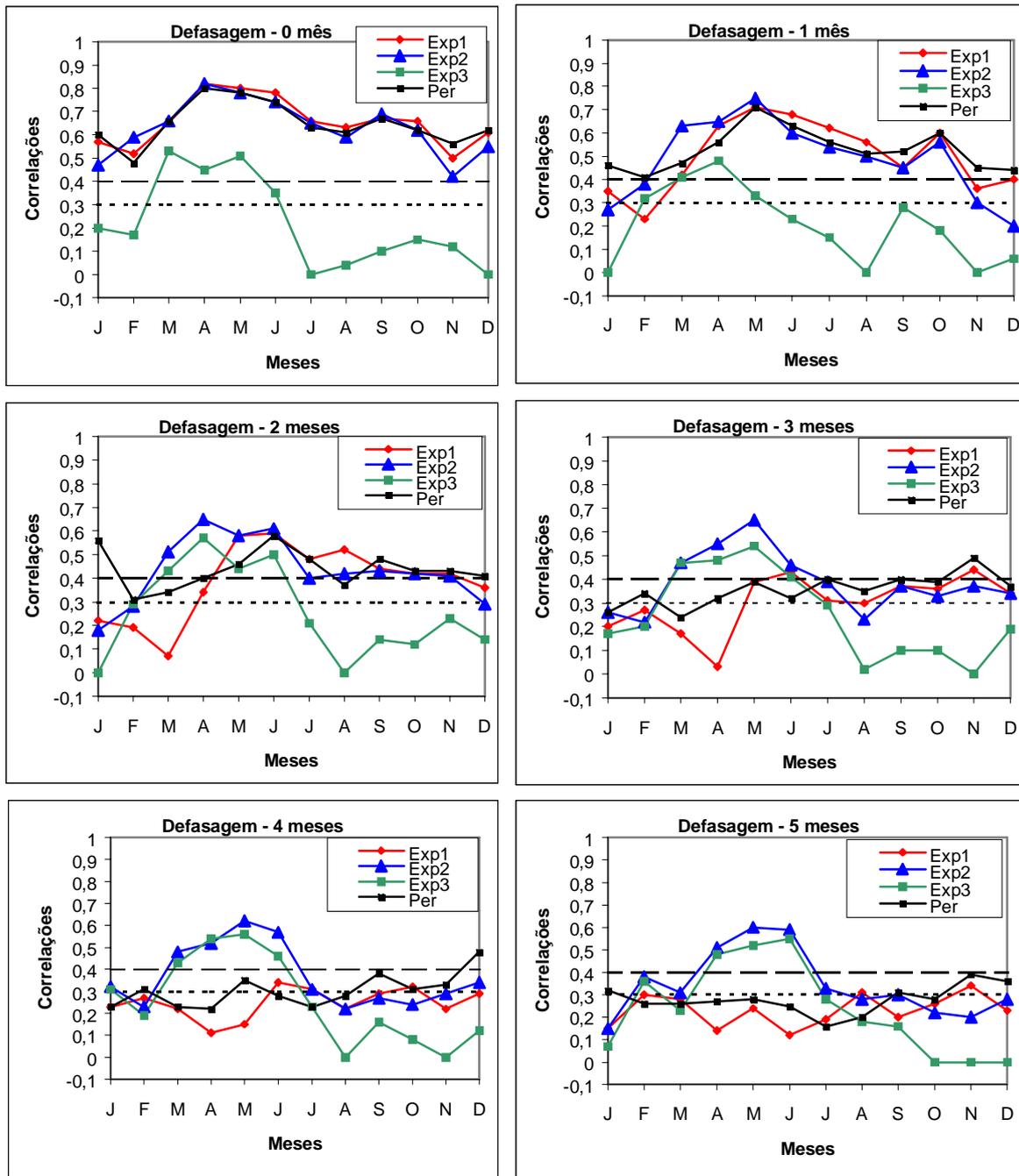


Fig. 3.1 – Correlações entre ATSM previstas pela ACC e observadas para os experimentos 1, 2 e 3, e ATSM persistidas e observadas no ATN, para defasagens de 0 a 5 meses. As linhas horizontais tracejadas com traços menores e maiores indicam níveis de significância de 95% e 99%, respectivamente.

Os coeficientes de correlação médios para o ATS para os exp1, exp2, exp3 e da persistência são apresentados na Figura 3.2. As maiores correlações ocorrem para defasagem zero com máximos em maio de 0,77 e 0,74 para os exp1 e exp2, respectivamente. Para defasagens acima de zero há uma redução nos valores das correlações. Em geral, os coeficientes de correlação do exp1 são maiores que os do exp2, para todas as defasagens. Em relação à persistência, os modelos do exp1 são ligeiramente melhores em alguns casos. As correlações do exp3 não são significantes. Esse resultado indica que a previsibilidade do ATS pode estar relacionada mais com efeitos locais do que com influências remotas.

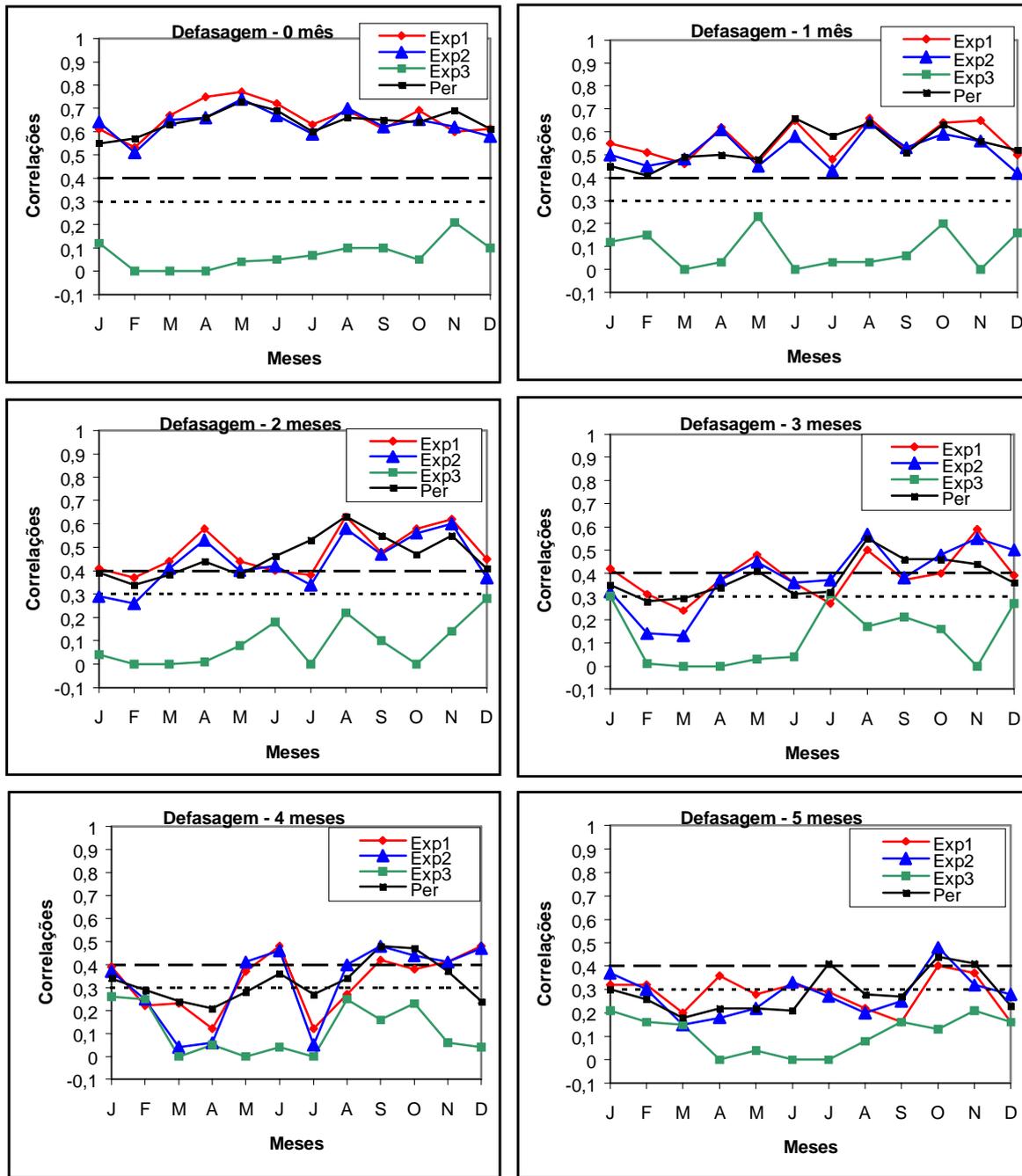


Fig. 3.2 – Correlações entre ATSM previstas pela ACC e observadas para os experimentos 1, 2 e 3, e ATSM persistidas e observadas no ATS, para defasagens de 0 a 5 meses. As linhas horizontais tracejadas com traços menores e maiores indicam níveis de significância de 95% e 99%, respectivamente.

Uma outra região importante é o AE, visto que para certos períodos o padrão de variabilidade equatorial é dominante (Zebiak, 1993). A Figura 3.3 apresenta os valores médios do coeficiente de correlação para esta área. Para defasagem zero, os coeficientes de correlação do exp1 atingem valores maiores que 0,6, exceto em fevereiro, e as melhores previsões ocorrem nos meses de abril a julho. Para essa mesma defasagem, as correlações médias para as previsões de junho a dezembro do exp2 são maiores do que as do exp1. Os coeficientes de correlação médios dos modelos estatísticos do exp2 apresentam valores maiores que os do exp1 em muitos casos, para todas as defasagens. Em relação à persistência, os modelos

estatísticos dos exp1 e 2 apresentam melhor desempenho em muitos casos. Como para o ATS, as previsões realizadas com o exp3 não apresentam valores significantes em todas as defasagens.

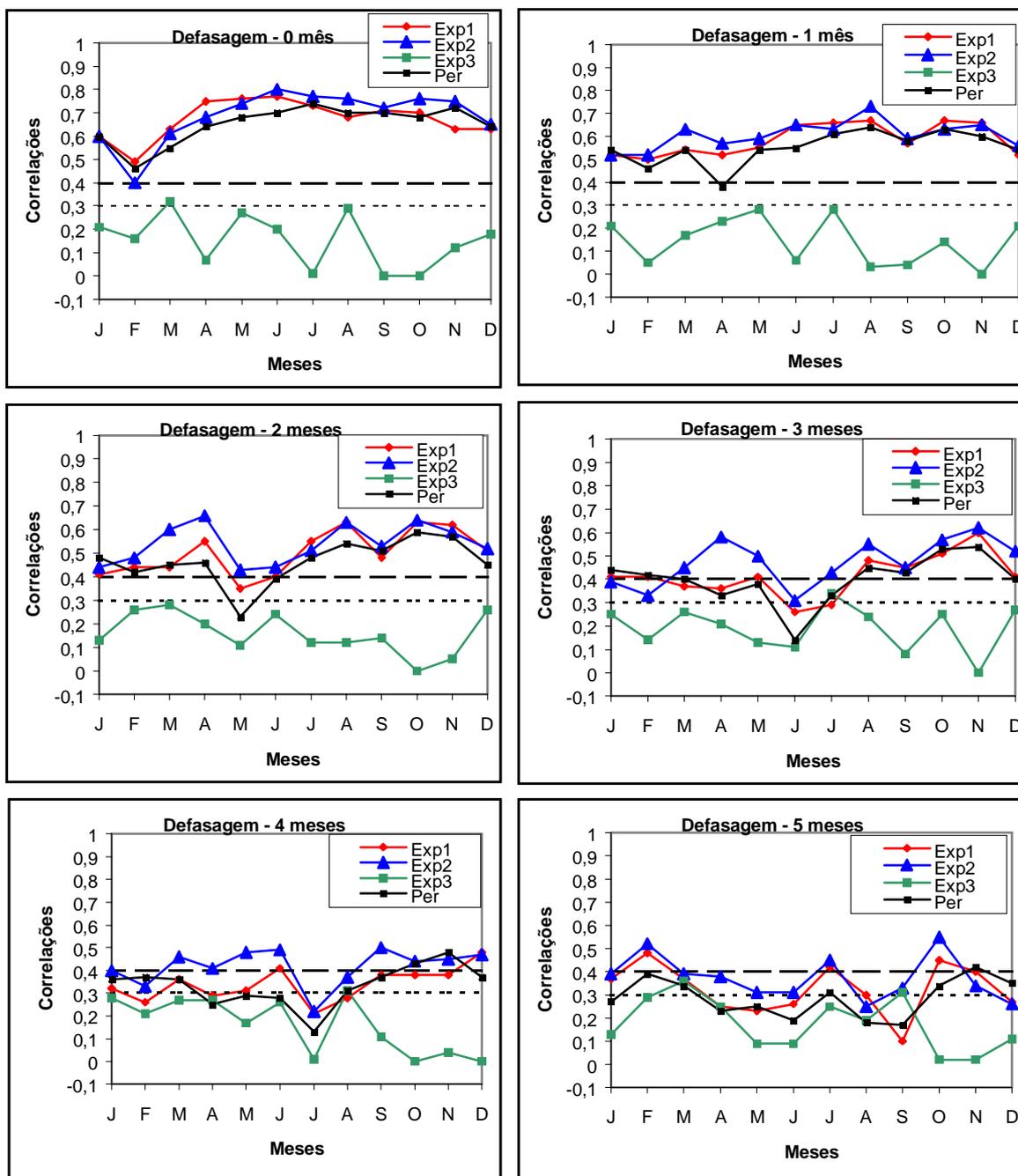


Fig. 3.3 – Correlações entre ATSM previstas pela ACC e observadas para os experimentos 1, 2 e 3, e ATSM persistidas e observadas no AE, para defasagens de 0 a 5 meses. As linhas horizontais tracejadas com traços menores e maiores indicam níveis de significância de 95% e 99%, respectivamente.

4 – DISCUSSÕES E CONCLUSÕES

Modelos estatísticos para a previsão de ATSM do AT foram construídos usando análises por correlações canônicas e campos de ATSM do AT e/ou PE como preditores. O desempenho destes modelos é investigado considerando todos os meses de janeiro a dezembro como preditores para defasagens de 0 a 5 meses. Comparações das previsões são feitas com as ATSM observadas e persistidas.

É notável que para alguns meses de antecedência, as previsões para o ATN são melhores que as obtidas para o ATS e AE quando se considera as ATSM do AT e PE como preditoras. Nesta análise não é considerada a natureza da influência, entretanto, esse resultado é consistente com trabalhos anteriores. A tendência do ATN ser melhor previsto ao se incluir a região do PE além do AT como preditor (exp2), também foi notada por Peland e Matrosova (1998) e Ruiz de Elvira et al. (2000). Isso sugere que os fenômenos como o El Niño – Oscilação Sul (ENOS) que modulam as variações de TSM do PE têm maior influência nas variações da TSM do ATN do que nas do AE e ATS. Por outro lado, a previsibilidade do ATN aumenta quando se considera os meses de março a junho como preditandos e defasagens de 2 a 5 meses. Isso implica que os preditores são de novembro a março, que corresponde ao período em que geralmente ocorre a fase madura dos extremos do ENOS no PE, confirmando que o ENOS tem um papel importante nas variações de ATSM do ATN. Para previsões feitas para os meses de julho a janeiro, a inclusão do PE não melhora o desempenho dos modelos.

Para as regiões do ATS e AE as maiores correlações ocorrem para defasagem zero para os experimentos 1 e 2. No caso do ATS, em geral as correlações do exp1 (somente o AT como preditor) são maiores do que as do exp2 (AT e PE como preditores), para todas as defasagens. Por outro lado, para o AE essa relação é inversa. Estes resultados sugerem que em relação à antecedência para se prever ATSM, no ATS e AE, é relativamente menor do que aquela do ATN. Além disso, foi notado que as ATSM somente do PE como preditores produz as piores previsões para esses dois setores. Para o ATN, correlações significantes para os meses de abril, maio e junho para defasagens de 2 a 5 meses são encontradas quando se considera somente o PE como preditor.

Assim, os resultados encontrados sugerem que para se obter previsões de ATSM do AT mais precisas deveriam ser usados modelos mensais (um para cada mês do ano). Além disso, a inclusão do PE além do AT como a região dos preditores deve ser restrita aos meses de novembro a março, considerando defasagens de 0 a 5 meses. Enquanto para os demais meses, somente o AT deve ser considerado a região dos preditores.

5 - REFERÊNCIAS BIBLIOGRÁFICAS

- Barnett, T.P; R. Preisendofer: Origins and levels of monthly and seasonal forecast skill for United States surface air temperatures determined by canonical correlation analysis. **Monthly Weather Review**, 115: 1825-1850, 1987.
- Barnston, A. G.; C. F. Ropelewski: prediction of ENSO episodes using canonical correlation analysis. **Journal of Climate**, 5(11), 1316-1345, 1992.
- Da Silva, A.A.M., C.C. Young e S. Levitus. Atlas of surface marine data. Vol. 1: Algorithms and procedures, **NOAA ATLAS NESDIS 6**, U. S. Department of Commerce, NOAA, NESDIS, 83pp, 1994.
- Hastenrath, S. Prediction of Northeast Brazil rainfall anomalies. **Journal of Climate**, 3:893-904, 1990.
- Hastenrath, S.; L. Greischar. Further work on the prediction of northeast Brazil rainfall anomalies. **Journal of Climate**, 6: 743-758, 1993.
- Landman, W. A, S. J. Mason. Forecasts of near-global sea surface temperatures using canonical correlation analysis. **Journal of Climate**, 15: 3819-3833, 2001.
- Penland C.; L. Matrosova: Prediction of Tropical Atlantic sea surface temperatures using linear inverse modeling. **Journal of Climate**, 11; 483-496, 1998.
- Repelli, C. A., P. Nobre. Statistical prediction of sea surface temperature over the Tropical Atlantic. [CD ROM] **X Congresso Brasileiro de Meteorologia**, Brasília –DF, Brasil, 1998.
- Ruiz de Elvira, A. M. J. Ortiz Beviá e W. Cabos Narcaez. Empirical forecasts of tropical Atlantic sea surface temperature anomalies. **Quartely Journal Royal Meteorological Society**., 126: 2199-2210, 2000.
- Zebiak, S.E.: Air-sea interaction in the equatorial Atlantic region. **Journal Climate**, 6, 1567-1586, 1993.