

DEMONSTRAÇÃO MATEMÁTICA DE POSSÍVEIS DIFERENÇAS ENTRE A CORRELAÇÃO DE DUAS VARIÁVEIS SIMPLES E A CORRELAÇÃO DE DUAS VARIÁVEIS DERIVADAS

Daniel Pires Bitencourt¹ e Prakki Satyamurty²
Centro de Previsão de Tempo e Estudos Climáticos - CPTEC/INPE
email: (1) daniel@cptec.inpe.br ; (2) saty@cptec.inpe.br

Abstract

Possible differences in the anomaly correlation coefficients between forecast and verification fields obtained with the primitive variables and derived variables are demonstrated through simple hypothetical cases. It is shown that errors in the prediction of the shortwaves phase is more problematic for the model's skill.

Palavras chave: Previsão / Avaliação / Erro Numérico

Introdução:

Nos dias de hoje a simulação da atmosfera através de modelos numéricos é uma ferramenta essencial para a prática de previsão de tempo. A partir dos resultados desses modelos é possível se fazer previsões mais precisas e com maior antecedência. Segundo Bonatti (1996), a previsão numérica viabiliza prognósticos de até cinco dias para a região extratropical do Hemisfério Sul. No entanto, é necessário certo cuidado quando utiliza-se de campos meteorológicos prognosticados por modelos numéricos. Para se fazer uma previsão de tempo é importante ser crítico, considerando basicamente três itens. Em primeiro lugar considera-se a climatologia da região e a época do ano. A previsão não pode apresentar situações atmosféricas muito diferentes do que normalmente é observado em cada região e em cada época do ano. Em segundo lugar considera-se o prazo de antecedência e a latitude. Segundo Moura (1984), o prazo de antecedência para previsão de tempo nos trópicos é menor do que nos extratropicos. Por último, chama-se a atenção para a variável que está sendo prevista. Segundo Bitencourt (1996), em geral, as variáveis derivadas apresentam menor desempenho do que as variáveis simples. Considera-se variável derivada aquela que requer cálculo para sua obtenção, como divergência, vorticidade, advecção térmica, etc. As variáveis simples são altura geopotencial, temperatura virtual, ventos zonal e meridional e umidade específica.

Esse trabalho sintetiza o desenvolvimento matemático feito por Bitencourt (1996) para demonstrar possíveis diferenças entre a correlação de duas variáveis simples e a correlação de duas variáveis derivadas. Utilizando-se de casos simples e hipotéticos, verifica-se que muitas vezes a correlação entre a previsão de uma variável simples e sua análise é diferente da correlação entre a previsão de uma variável derivada e sua análise.

Demonstração Matemática:

Faz-se uma demonstração matemática, para três casos simples e hipotéticos, a fim de analisar os resultados obtidos da correlação entre duas variáveis simples e da correlação entre duas variáveis derivadas.

i) Caso 1 - Seja a variável analisada V_A e a variável prevista V_P e considere que V_A e V_P são compostas de uma única onda senoidal, com $\bar{V}_A = 0$ e $\bar{V}_P = 0$, ou seja, $V'_A = V_A$ e $V'_P = V_P$. Assim $V_A = a \sin n \lambda$ e $V_P = b \sin n (\lambda + \varepsilon)$, onde a e b são as amplitudes das ondas V_A e V_P , respectivamente, n é o número de onda, λ é a longitude e ε é a defasagem da onda prevista em relação a onda analisada (ver Figura 1).

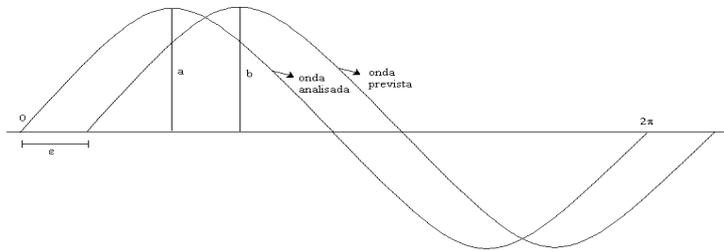


Fig. 1 - Esboço das ondas analisada e prevista, com $n=1$.

A covariância entre V_A e V_P é dada por $CV = \frac{1}{2} a b \cos n \varepsilon$. (equação 10)

Tomando as derivadas de V_A e V_P , obtém-se $V_A^\lambda = a n \cos n \lambda$ e $V_P^\lambda = b n \cos n (\lambda + \varepsilon)$, sendo $V_A^\lambda = \frac{d V_A}{d \lambda}$ e $V_P^\lambda = \frac{d V_P}{d \lambda}$.

Calculando agora a covariância das variáveis derivadas V_A^λ e V_P^λ obtém-se $CV^\lambda = \frac{1}{2} a b n^2 \cos n \varepsilon$. (equação 11)

O desvio padrão das variáveis analisada e prevista e de suas derivadas são

$$\sigma_{V_A} = \frac{a}{\sqrt{2}} \quad ; \quad \sigma_{V_P} = \frac{b}{\sqrt{2}} \quad \text{(equação 12)}$$

$$\sigma_{V_A^\lambda} = \frac{a n}{\sqrt{2}} \quad ; \quad \sigma_{V_P^\lambda} = \frac{b n}{\sqrt{2}} \quad \text{(equação 13)}$$

Usando os resultados das equações 10 e 12 obtém-se a correlação entre as

variáveis simples analisada e prevista, dada por

$$CC = \cos n \varepsilon \quad (\text{equação 14})$$

Com os resultados das equações 11 e 13 obtém-se a correlação entre as variáveis derivadas analisada e prevista, dada por

$$CC^\lambda = \cos n \varepsilon \quad (\text{equação 15})$$

Assim, verifica-se que quando as variáveis correlacionadas são compostas de uma única onda n , a correlação entre as variáveis analisada e prevista (CC) e a correlação entre as variáveis derivadas analisada e prevista (CC^λ) são iguais. É importante lembrar que, para

$$\begin{aligned} n \varepsilon = 180^\circ & \Rightarrow CC = CC^\lambda = -\cos n \varepsilon = -1 & (\text{correlação inversa}) \\ n \varepsilon = 90^\circ \text{ ou } 270^\circ & \Rightarrow CC = CC^\lambda = 0 & (\text{correlação nula}) \\ n \varepsilon = 0, 2\pi & \Rightarrow CC = CC^\lambda = 1 & (\text{correlação perfeita}) \end{aligned}$$

Os resultados obtidos acima são válidos tanto para ondas de amplitudes iguais ($a = b$) como para ondas de amplitudes diferentes ($a \neq b$).

ii) Caso 2 - Quando as variáveis analisada e prevista são compostas de ondas com número de onda diferentes ($n \neq m$) da seguinte forma: $V_A = a \sin n \lambda$ e $V_P = b \sin m \lambda$, sendo suas derivadas dadas por $V_A^\lambda = an \cos n \lambda$ e $V_P^\lambda = bm \cos m \lambda$, as correlações são nulas, com $CC = CC^\lambda = 0$.

Isso demonstra que o campo previsto deve, pelo menos, indicar o mesmo número de onda que o campo analisado para que estes tenham alguma correlação.

iii) Caso 3 - Supondo-se agora que as variáveis a serem correlacionadas sejam compostas de duas ondas diferentes, tem-se: $V_A = a_m \sin m \lambda + a_n \sin n \lambda$ e $V_P = b_m \sin m (\lambda + \varepsilon_m) + b_n \sin n (\lambda + \varepsilon_n)$, suas derivadas são $V_A^\lambda = ma_m \cos m \lambda + na_n \cos n \lambda$ e $V_P^\lambda = mb_m \cos m (\lambda + \varepsilon_m) + nb_n \cos n (\lambda + \varepsilon_n)$.

As covariâncias de V_A e V_P e de suas derivadas são, respectivamente

$$CV = \frac{1}{2} (a_m b_m \cos m \varepsilon_m + a_n b_n \cos n \varepsilon_n) \quad (\text{equação 16})$$

$$CV^\lambda = \frac{1}{2} (m^2 a_m b_m \cos m \varepsilon_m + n^2 a_n b_n \cos n \varepsilon_n) \quad (\text{equação 17})$$

O desvio padrão de V_A , V_P , V_A^λ e V_P^λ são, respectivamente

$$\sigma_{V_A} = \left(\frac{a_m^2 + a_n^2}{2} \right)^{1/2} ; \quad \sigma_{V_P} = \left(\frac{b_m^2 + b_n^2}{2} \right)^{1/2} \quad (\text{equação 18})$$

$$\sigma_{V_A^\lambda} = \left(\frac{m^2 a_m^2 + n^2 a_n^2}{2} \right)^{1/2} ; \quad \sigma_{V_P^\lambda} = \left(\frac{m^2 b_m^2 + n^2 b_n^2}{2} \right)^{1/2} \quad (\text{equação 19})$$

Usando os resultados das equações 16 e 18 obtém-se a correlação entre as variáveis analisada e prevista, dada por

$$CC = \frac{a_m b_m \cos m \epsilon_m + a_n b_n \cos n \epsilon_n}{\left(a_m^2 + a_n^2 \right)^{1/2} \left(b_m^2 + b_n^2 \right)^{1/2}} \quad (\text{equação 20})$$

Com os resultados das equações 17 e 19 obtém-se a correlação entre as variáveis derivadas analisada e prevista, dada por

$$CC^\lambda = \frac{m^2 a_m b_m \cos m \epsilon_m + n^2 a_n b_n \cos n \epsilon_n}{\left(m^2 a_m^2 + n^2 a_n^2 \right)^{1/2} \left(m^2 b_m^2 + n^2 b_n^2 \right)^{1/2}} \quad (\text{equação 21})$$

Portanto, quando as variáveis correlacionadas são compostas de duas (ou mais) ondas com números de onda m e n distintas, a correlação entre as variáveis analisada e prevista é diferente da correlação entre as suas derivadas, ou seja, $CC \neq CC^\lambda$. Para o caso de $m = n$ ou $m \approx n$, volta-se ao **Caso 1** onde as variáveis são compostas de uma única onda, fazendo com que as correlações sejam iguais ou próximas ($CC = CC^\lambda$ ou $CC \approx CC^\lambda$).

Com o objetivo de facilitar a interpretação dos resultados obtidos nas equações 20 e 21, considera-se que as amplitudes das ondas são todas iguais e unitárias, $a_m = a_n = b_m = b_n = 1$. Nesse caso obtém-se

$$CC = \frac{\cos m \epsilon_m + \cos n \epsilon_n}{2} \quad (\text{equação 22})$$

$$CC^\lambda = \frac{m^2 \cos m \epsilon_m + n^2 \cos n \epsilon_n}{m^2 + n^2} \quad (\text{equação 23})$$

Para apreciar a diferença entre as equações 22 e 23 assumo o exemplo de $n = 4m$. Tem-se

$$CC = \frac{\cos m \varepsilon_m + \cos 4m \varepsilon_n}{2} \quad (\text{equação 24})$$

$$CC^\lambda = \frac{\cos m \varepsilon_m + 16 \cos 4m \varepsilon_n}{17} \quad (\text{equação 25})$$

É fácil notar, através dos resultados expressos pelas equações 24 e 25, que na CC^λ o que domina é a onda curta (isto é, onda número $4m$), entretanto na CC as duas ondas contribuem com pesos iguais. Assim, qualquer erro que ocorra nas previsões de ondas curtas ($\cos 4m \varepsilon_n$) será mais sensível aos resultados de correlação entre variáveis derivadas do que entre variáveis simples. O erro ocorrido nas ondas curtas é mais prejudicial às previsões de variáveis derivadas.

Discussões Finais:

Quando as variáveis prevista e analisada são compostas de uma única onda número n , a correlação entre variáveis simples tem o mesmo valor da correlação entre variáveis derivadas, sendo que este resultado não torna-se diferente até quando as ondas previstas e analisadas possuem amplitudes diferentes. No entanto, quando a onda prevista possui número de onda diferente da onda analisada, tanto as correlações entre variáveis simples como as correlações entre variáveis derivadas são nulas. Assim, o campo previsto deve pelo menos ter o mesmo número de onda do campo analisado para que esses possuam alguma correlação.

Na hipótese das variáveis analisadas e previstas serem compostas de duas (ou mais) ondas diferentes, com números de onda m e n , a correlação entre variáveis simples e a correlação entre variáveis derivadas são diferentes. Considerando que as amplitudes de todas as ondas possuem valores iguais e unitários e que estas ondas possuem um número de onda quatro vezes maior que o outro ($n = 4m$), nota-se que o valor da correlação entre duas variáveis derivadas é 16 vezes mais influenciado pela onda curta (onda número $4m$) e que o valor da correlação entre duas variáveis simples é influenciado igualmente pelas duas ondas. Assim, um erro ocorrido nas ondas curtas é mais prejudicial às previsões de variáveis derivadas.

Referências Bibliográficas:

- Bitencourt, D. P. Desempenho das previsões de variáveis simples e derivadas obtidas pelo modelo global do CPTEC/COLA para alguns casos significativos ocorridos sobre o Centro-Sul do Brasil. (INPE - 6467 - TDI/617). SJC, 1996. 138 p.
- Bonatti, J. P. Verificação Estatística do Modelo Global do CPTEC. Comunicação Pessoal, 1996.
- Moura, A. D. The ocean and the dynamics of the tropical atmosphere; its possible predictability. São José dos Campos, INPE, 1984. 8 p. (INPE - 2998 - PRE/450).