

## MODELO DE PREVISÃO DE VAZÕES SAZONAIS E INTERANUAIS

Francisco de Assis de Souza Filho<sup>1,2</sup>; Upmanu Lall<sup>2</sup> & Rubem L.L. Porto<sup>3</sup>

**Resumo** - O artigo apresenta os resultados de um modelo de previsão de vazões sazonais e interanuais para um sistema de reservatórios no Estado do Ceará. As previsões terão um horizonte de até 18 meses, a previsão é realizada em julho.

O modelo é semiparamétrico. A metodologia apresentada utiliza a Análise dos Componentes Principais das vazões do sistema de reservatórios. Sob os componentes principais é realizada uma regressão linear sobre os índices climáticos. Com os parâmetros desta regressão aplica-se a metodologia dos k vizinhos mais próximos para o cálculo da distribuição de probabilidade das vazões afluentes. Esta metodologia é comparada com a apresentada por Souza Filho e Lall (2002). Os resultados obtidos foram satisfatórios, tendo a correlação da mediana da distribuição obtida com as observações o valor de 0,93. Pretende-se a partir da previsão realizada por esta metodologia operar um sistema de reservatórios com vistas ao suporte à decisão do processo de alocação de água.

**Abstract** - The article introduces the results of a forecast model of seasonal flows and interanuais for a reservoirs system in the State of the Ceará. The forecasts will have a horizon up to 18 months, the forecast is accomplished in July.

The model is semiparamétrico. The introduced methodology uses the Analysis of the Main Components of the flows of the reservoirs system. Under the main components is accomplished a lineal regression about the climatic indices. With the parameters of this regression it applies to methodology of the k nearer neighbors for the calculation of the probability distribution of the affluent flows. This methodology is compared with introduced her by Souza Filho and Lall (2002). The obtained results were satisfactory, having the correlation of the median of the distribution obtained with the observations the value of 0,93. It intends from the forecast accomplished by this methodology operate a reservoirs system with sights to the support to the decision of water allocation process.

---

<sup>1</sup> Fundação Cearense de Meteorologia e Recursos Hídricos. [assisfilho@secrel.com.br](mailto:assisfilho@secrel.com.br). (85)4331803

<sup>2</sup> Columbia University – International Research for Climate Prediction. [ula2@columbia.edu](mailto:ula2@columbia.edu)

<sup>3</sup> Escola Politécnica da Universidade de São Paulo - Departamento de Engenharia Hidráulica e Sanitária - Av. Prof. Almeida Prado, 271 – Trav. 3 - Cidade Universitária - São Paulo – Brasil - CEP: 05508-900 - Tel: (11) 3091 5403 [assisfilho@secrel.com.br](mailto:assisfilho@secrel.com.br), [rlporto@usp.br](mailto:rlporto@usp.br);

## INTRODUÇÃO

Regiões semi-áridas, como o Nordeste do Brasil, são particularmente vulneráveis a flutuações do clima e seus impactos nos suprimentos de água. Modelos de previsão de vazões em rios para horizonte de alguns meses ou ano tornam-se muito interessantes com vistas a operação mais eficiente dos suprimentos de água e o processo de alocação de água entre usos e usuários conflitantes.

A previsão de vazão pode ser realizada utilizando-se modelagem dinâmica ou estatística. Um procedimento para a realização de tais previsões é o uso de Modelos de Circulação Gerais (GCMs) do oceano e a atmosfera, seguindo-se por "downscaling" que usa Modelos regionais de previsão de clima ou aproximações estatísticas, seguido por modelos chuva-deflúvio concentrados ou distribuídos. Esta é uma direção de pesquisa útil. Neste momento, ainda estão sendo avaliados assuntos relacionados a propagação de incerteza ao longo da metodologia, processos estão sendo parametrizados e capacidade de previsão final esta sendo avaliada.

Uma alternativa a esta metodologia é o desenvolvimento direto de previsões estatísticas para os aportes de água e as demandas a partir de uma adequada seleção de preditores de clima. Um novo método de previsão estatística de escoamento em rios é apresentado neste texto.

Souza Filho e Lall (2002) apresentaram metodologia para a previsão de vazões sazonais e interanuais utilizando a "pull regression" na obtenção de regressão linear. No presente artigo será apresentada uma regressão após Análise dos Componentes Principais; avaliando-se os ganhos com relação à metodologia utilizada naquele trabalho. Objetivando um entendimento da abordagem realizada reproduziremos de forma atualizada a descrição dos estudos anteriores e da análise e diagnóstica descritos no referido trabalho.

O objetivo é desenvolver uma metodologia que fosse compatível com as necessidades de informação e métodos de análise de uma agência de água responsável para operar um sistema de reservatórios em um rio. Isto é, o método pretende prever vazões mensais mantendo a estrutura espacial e temporal da série de vazões históricas e em um período de operação de três meses a 18 meses. Muitos operadores de hidrossistemas tomam decisões simulando-o a partir de reamostragem de uma série de registros de vazões históricas e demandas projetadas. Desenvolveu-se neste trabalho um método de re-amostragem de vazões condicionada ao estado climático. Utilizaram-se técnicas de regressão multivariada para explorar a relação entre vazões anuais (ou sazonais) e um conjunto de potenciais preditores climáticos. Ponto chave nesta previsão é a potencial relação não linear e os resíduos desta regressão não apresentarem uma distribuição normal. A re-regressão linear é utilizada aqui para transformar o espaço das variáveis preditoras. Emprega-se, no espaço transformado das variáveis preditoras, uma aproximação não paramétrica para a densidade de probabilidade condicional da matriz de vazões mensais. Em todos os locais, para o horizonte de

interesse, é realizada simulações de Monte Carlos para as vazões futuras geradas pelo método não paramétrico k-Nearest Neighbor (Lall e Sharma, 1996). Desagregação para balanças de tempo de diário, usando métodos relacionados apresentados em Kumar et al (2000) também pode ser considerado.

Em uma análise de capacidade de previsão dos principais modelos de Circulação Geral oceano-atmosfera para precipitação sazonal, Rajagopalan et al (2001), concluiu que o Nordeste do Brasil é uma das regiões do mundo onde se apresenta de forma consistente e estatisticamente significativa durante a estação chuvosa primária (março-maio) melhor capacidade de previsão. Seca é uma preocupação permanente no estado de Ceará. E o sistema de reservatórios freqüentemente é levado a seus níveis mínimos mesmo sendo projetados para regularização plurianual com volumes aproximadamente duas vezes a vazão média afluyente. Dado o potencial para previsões de horizonte longo duração, e a alta utilidade de tal informação, o Ceará provê um estudo de caso importante para os métodos desenvolvidos aqui. A metodologia de previsão desenvolvida é aplicada para o sistema de 6 reservatórios nas Bacias do Jaguaribe e Metropolitanas no Ceará.

Em próximos trabalhos mostrar-se-á como a previsão devera ser utilizada para a construção de curvas guia e operação "ótima" de hidrossistemas.

## **ESTUDOS ANTERIORES**

Inicialmente é analisada a informação sobre as secas no Ceara e as suas decorrências no suprimento de água e na demanda e a variação da vazão afluyente aos reservatórios. Os Esforços anteriores de previsões estatísticas de vazões ou precipitação na região serão analisados no contexto da identificação de um conjunto satisfatório de predictores climático para o Ceará.

O sistema de água de interesse é o Hidrossistema Jaguaribe-Metropolitano (JMH) no Ceará. Este é o sistema de água mais importante no Estado. Seis reservatórios principais (veja Tabela 1 - o primeiro três estão na Bacia de Jaguaribe e o restante na Bacia Metropolitana) provêm as áreas irrigadas do Jaguaribe e a área metropolitana (Fortaleza). Os primeiros quatro reservatórios estão na região semi-árida do estado. Precipitação nas bacias dos rios Pacoti-Riachao e Cocó são influenciadas por chuvas orográficas da serra de Baturité. Os registros de Chuva nas bacias estão disponível desde aproximadamente 1911. O Registro de vazão nos locais tem diferentes datas de início variando de 1912 a 1970. Por conseguinte, calibrou um modelo chuva-deflúvio para reconstruir a vazão em cada um dos reservatório. Espera-se que a qualidade dos dados de vazão seja melhor para o reservatório de Oros, e pior para o Pacoti-Riachão. Observa-se certa quantidade de anos com vazão anual zero.

A vazão anual em todos os locais é altamente variável e assimétrica (Tabela 1). Noventa por cento das vazões tem sua ocorrência nos meses de janeiro a junho. O Plano Estadual de Recursos Hídricos (SRH, 1991) e Plano de Bacia de Hidrograficas (COGERH,2000\_a; COGERH,2000\_b) prover as projeções de demanda para JMH. A Bacia do Jaguaribe a demanda de água é 80% Irrigação e 20% urbano. A demanda de água das Bacias Metropolitanas é predominantemente para uso Urbano e Industrial. Por conseguinte, são distribuídas relativamente uniformemente as demandas na bacia Metropolitana durante o ano, enquanto esses na bacia de Jaguaribe se concentram na estação de irrigação (agosto a novembro).

Alguns esforços recentes de identificação dos mecanismos e causas climáticas do regime de chuva e vazão no Nordeste do Brasil são discutidos em Uvo et al (2000), Uvo e Graham (1998), Uvo et al (1998), Marengo et al (1998), e Xavier et al (2000). A chuva na região é altamente variável no espaço, dentro da estação chuvosa e durante anos (Kousky, 1979). A sazonalidade da chuva regional, e conseqüentemente das vazões, é governado em grande parte pela migração norte/sul da zona de convergência Intertropical (ITCZ). Uvo et al (1998) sintetize uma descrição da conexão entre chuva e o ITCZ baseado em pesquisas passada. Eles indicam que a estação chuvosa principal é iniciada entre fevereiro e março, quando a ITCZ sobre o Oceano Atlântico tropical alcança sua posição mais ao sul. A migração da ITCZ para o norte sinaliza o fim da estação chuvosa. O início e o fim deste processo tem datas com alta variabilidade, e este fato afeta significativamente o total de chuva sazonal. A chuva de janeiro-fevereiro é afetada por frentes frias ou por mecanismos a elas associados (Kousky, 1979). Conexões entre o Pacífico Oriental e o comportamento de ITCZ no Atlântico Tropical foi estudado por Nobre e Shukla (1996), Saravanan e Chang (2000), e Chiang et al (2000) com duas hipóteses contrastantes. Nobre e Shukla explicam a conexão entre um ENSO maduro no inverno boreal e a parte do norte do SST Atlântico tropical no inverno e a primavera seguinte em termos de uma "ponte atmosférica". Saravanan e Chang apontam ao papel de Circulação de Walker nesta conexão. Chiang et al oferecem apoio observational para este mecanismo e analisam suas variações interdecadais (estes relacionam diretamente à frequência variável de El Nino e La Nina com eventos de 21 anos em janelas móveis).

A variabilidade de chuva foi relacionada a variações em Temperaturas de Superfície de Mar (Markham e McLain (1977), Moura e Shukla (1981), Hastenrath et al (1984), Hastenrath (1990), Ward et all (1991), Ward et al (1988), e Ward et al (1993)). Ward e Folland(1991) achou que é melhor usar o EOFs só das temperaturas de Superfície de Mar Atlânticas tropicais como predictors da chuva de Nordeste. As Anomalias de Pacífico associadas com ENSO representam um papel mais fraco. Ward et al (1993) indicou que os padrões espaciais de EOF no Atlântico não são freqüentemente robustos com respeito ao período de análise e especulam várias razões para este fato. Eles demonstram que as previsões são estatisticamente significativas de chuva sazonal no

Nordeste sazonal com de 0 a 2 meses de antecipação usando regressão linear Múltipla e análise de discriminante linear. O trabalho de Uvo et al (1998) considerando uma modelagem multivariada mais detalhada espacialmente (105 estações) e tempo (mensal e sazonal) analisa a precipitação no Nordeste e sua relação com as SSTs. Os resultados indicam que as anomalias de SST no Atlântico equatorial Meridional é associado com uma migração mais cedo do ITCZ e conduz a aumento de chuva em partes do Nordeste inclusive no Ceará. De acordo com estudos prévios identificam que a ITCZ em abril e maio e conseqüentemente o fim da estação chuvosa no Norte do nordeste é em grande parte determinada por um gradiente de Norte-Sul na SST do Atlântico equatorial. A correlação com um índice de ENSO-Pacífico durante este período também é significativa. Baseado nos resultados da análise de multivariada, Uvo et al constroem índices de SST para o Pacífico Central, o Atlântico Equatorial Norte e Sul, e a diferença entre o atlântico Norte e Sul. Observaram que o índice do "dipolo do Atlântico" equatorial é um preditor das precipitações do Nordeste com 1 a 3 meses de antecedência, e confirma trabalho anterior que reconhece a utilidade de tal um índice. Avaliaram que os meses de abril e maio eram os mais importantes nas variações de interanuais das precipitação de Nordeste, e que o índice de dipole Atlântico é altamente correlatado com estas flutuações. O ENSO tem importância menor, mas estatisticamente significativo, e é associado com precipitação em Janeiro e Fevereiro, e de abril-maio. Um contexto para estas observações é provido pela análise de Chiang et al (2000) que também enfocou a dependência das chuvas no Ceará com os índice ENSO e o gradiente de SST entre o atlântico equatorial Norte e sul. Chiang observa que aumentos no NINO3 a amplitude e média das chuvas no Ceará tendem a decrescer. A interpretação sugerida por aqueles autores é que quando há pequena convecção no Oceano Pacífico equatorial leste (La Nina), a ITCZ do Atlântico Tropical é influenciado por outros fatores, principalmente o gradiente de SST entre o atlântico tropical norte e sul. Como aumentos da convecção sobre o Pacífico equatorial oriental, subsidência anômala sobre o Atlântico tropical reduz a chuva no nordeste do Brasil e sua variação. A relação não linear entre a SSTs do Pacífico e a convecção e sua influência na SST do Atlântico Tropical e as chuva no Ceará é identificada como um fator nas aparente mudança nas correlação entre NINO3 e suas teleconecções sobre o Atlântico.

A previsão de vazões com uma ou duas estação de antecedencia para a região Amazônica foi realizada por Uvo et al (2000) utilizando regressão através de redes Neurais utilizando a SSTs so Pacífico equatorial e do Atlântico Tropical para o período de 1946-1992. Obteve-se correlações entre observado e previsto das vazões no rio em 9 localidades que variam de 0.35 a 0.76. Estes resultados foram melhorados com a utilização de um modelo de regressão linear canônico implantado por Uvo e Graham (1998). A região Amazônica tem clima bastante diferente da região Nordeste, e conseqüentemente não está claro se tal resultado pode ser extrapolado para o Ceará. Não é claro porem a possibilidade do método de redes neurais ser eficaz na geração de cenários

probabilísticos para diversos locais ao mesmo tempo mantendo a mesma estrutura nos diversos locais e sendo de fácil comunicação para o operador de reservatórios.

No contexto das previsões de vazões, as conexões do ENSO no início da estação chuvosa pode ser importante para prescrever a umidade do solo inicial e conseqüentemente potencial de deflúvio. Considerando que se esta interessado nas previsões de vazões no meio do ano anterior ao que se pretende prever, esta-se interessado em na previsibilidade potencial dos índices do Atlântico ou do pacífico para os meses que antecedem a estação chuvosa. Uma investigação completa da dependência da vazão com relação a SST com diferentes defasagens em toda a região de interesse. Esta análise não será realizada neste momento, utilizar-se-a os preditores identificados por outros pesquisadores. Neste trabalho procura solucionar assuntos relacionados a não linearidade e não normalidade das relações estatísticas, e as modificação necessária nos métodos clássicos de análise linear e multivariada. As duas séries temporais selecionada como preditores são o NINO3 (definida como a anomalia de temperatura da superfície do mar na região do Pacífico equatorial entre 150 graus W e 90 graus W e 5 graus S e 5 graus N) e um Gradiente de SST do Atlântico (EAD) série definida como a diferença na média mensal da anomalia de SST na região do Atlântico Norte(5-20N, 60-30W) e a média mensal da região do Atlântico Sul (0-20S, 30W-10E). A série de tempo mensal para estes índices foi derivada da malha de dados de SST desenvolvidos por Kaplan et al (1996) e disponível em [http://ingrid.ldeo.columbia.edu/SOURCES/KAPLAN/EXTENDED /](http://ingrid.ldeo.columbia.edu/SOURCES/KAPLAN/EXTENDED/)).

### **Desenvolvimento do Modelo de Previsão**

O procedimento utilizado para desenvolver as previsões para os seis locais no Ceará utilizando a série de EAD e NINO3 é descrita aqui. Os resultados das previsões serão analisados. A descrição geral do algoritmo utilizado é descrita a seguir.

As idéias principais são:

i) As vazões nos seis locais de Ceará tem alta correlação espacial e são aparentemente influenciadas pelo clima de uma maneira semelhante, este fato sugere a possibilidade da utilização de um mesmo modelo para todas as localidades.

ii) enquanto os índices de clima são autocorrelacionados, e tendo distribuição aproximadamente normal, as vazões anuais não parecem ser correlacionadas(pelo menos na na análise autoregressiva habitual), sugerindo que um procedimento de previsão apropriado pode ser obtido pela regressão das vazões anuais/mensais e uma seqüência de valores de índices de climáticos;

iii) o uso do de métodos estatísticos tradicionais, métodos paramétricos para construir um modelo comum de regressão linear para construir um modelo de previsão na escala anual e posterior desagregação mensal podem ser de difícil elaboração devido a grande assimetria das

distribuições anuais e mensais, o grande número de vazões zero, e a relação de não linear entre vazão e os índices do clima, e entre vazões mensais e anuais;

iii) os métodos não paramétricos para a regressão e estimativa da densidade de probabilidade podem também ter um sucesso limitado em uma aplicação direta, devido a grande amplitude dos dados de vazão, as dimensões do problema multivariado no espaço, e a quantidade de dados disponível;

iv) por conseguinte, desenvolveu-se uma aproximação semi-paramétrica decompondo o problema de estimação em três partes: (1) a transformação das vazões anuais em cada local, com vistas a garantir na regressão destas vazões com os índices um resíduo normal e com variância aproximadamente constante; (2) um modelo de regressão de vetor (por exemplo, regressão agrupada, regressão de componente principal, ou regressão canônica) com redução de dimensão com vistas a desenvolver um modelo comum de previsão das vazões anuais utilizando como preditores um conjunto de índices climáticos com certa antecedência; e (3) Utilizar este modelo de previsão para a reamostragem não paramétrica do conjunto de dados anuais da série histórica (e conseqüentemente um conjunto de vazões mensais/anuais em todos os locais), reamostragem esta contruída apartir de valores atuais dos índices climáticos.

A previsão por reamostragem semi-paramétrica utilizará uma série histórica de dados de 1914-2000 para as vazões e para os índices climáticos. Reservou-se um bloco contínuo de 5 a 10 anos para a validação e verificação do ajuste do modelo.

O preditores considerados para a previsão de julho das vazões anuais (Jan-Dec) para o ano seguinte são os valores dos índices do NINO3 e EAD de abril-maio-junho (AMJ). A Tabela 3, mostrada anteriormente, indicam que estes índices não são correlacionados. Utilizou-se uma transformação de potência nas vazões anuais (Jan. para Dec.) para cada um dos seis locais de interesse. A transformação raiz cubica proporcionou uma distribuição de probabilidade aproximadamente simétrica para as vazões anuais em cada local, e os resíduos das regressão nos dois índices tiveram variância aproximadamente constante, mas estas distribuições tiveram caudas mais grossas que as da distribuição Normal.

Avaliou-se duas metodologias alternativas para a regressão. A primeira a proposta por Souza Filho e Lall(2002), que consiste em uma regressão agrupada. A segunda realiza a regressão da primeira componente principal com os índices climáticos.

A primeira metodologia. Os coeficientes lineares para os índices foram obtidos por um procedimento de regressão de stepwise em todos os locais. Em seguida, calculou-se uma regressão agrupada (Dillon et al, 1984) da transformação das vazões anuais  $\mathbf{q}_s = (\mathbf{a}_s^{1/3} - \text{media}(\mathbf{a}_s^{1/3})) / \text{stdev}(\mathbf{a}_s^{1/3})$ , na série de AMJ de EAD e NINO3. Aqui,  $\mathbf{a}_s$  é o vetor que contém a série temporal de

vazões anuais no local  $s$ . A regressão agrupada ( $\mathbf{q}=\mathbf{X}\boldsymbol{\beta}+\boldsymbol{\varepsilon}$ ) não foi achado para ser diferente das seis regressões (veja Tabela 4) para cada local com um nível de significância de 95%.

Como uma alternativa para a regressão agrupada, foram consideradas também regressões dos componentes principais. Obteve-se os componentes principais das variáveis  $\mathbf{q}_s$ . Sob o primeiro componente principal realizou-se a regressão dos índices climáticos.

Considerando que interessa-se em gerar uma previsão de vazões anuais e mensais, o modelo de regressão não foi usado de forma direta na previsão. Para gerar uma previsão de vazões anuais, precisa-se de estimativas da distribuições de probabilidade condicional  $f(a_{t+f,s}|x_t)$  do vetor de vazões anuais do local  $s$ ,  $a_{t+f,s}$ . Assume-se um modelo para a distribuição de probabilidade dos resíduos da regressão, e realiza-se a transformação da variável retornando ao espaço inicial das vazões. Este procedimento gera vazões negativas, caso seja imposto para estas vazões negativas serem iguais a zero a funções densidade de probabilidade não se parecerá com a função densidade de probabilidade dos dados originais. Outra dificuldade neste processo de geração é a preservação da estrutura de correlação espacial entre os locais após a transformação inversa. Para os valores de vazões mensais em cada local, seria necessário calcular a distribuições de probabilidade condicional  $f(m_{t+f} | a_{t+f}, x_t)$ ,  $f(m_{t+f} | a_{t+f})$  ou  $f(m_{t+f} | x_t)$ , como apropriado para os dados. A desagregação de vazões anuais para a estimativa das vazões mensais (por exemplo, Bras e Rodriguez-Iturbe, 1984, segundo 3.5) com vistas a preservar a consistência espacial e temporal pode ser considerada. Objetivando a geração de previsão realizada por um conjunto de valores obtida de uma amostragem dos dados sa série histórica apresentar conveniência pratica.

Utiliza-se uma metodologia não paramétrica neste ponto. O procedimento  $k$ -nearest neighbor de estimativa da função de densidade descritos em Lall e Sharma (1996) e Karlsson e Yakowitz (1987) é adaptado ao problema atual.

Dado um vetor com dados atuais  $1 \times p$  de preditores  $x^*$ , realiza-se uma amostragem das vazões anuais  $a_{t+f}$  e o vetor correspondente  $m_{t+f}$  das vazões mensais que implicitamente reflete a distribuição de probabilidade condicional  $f(M|x)$ . A estratégia básica é selecionar os  $k$  vizinhos mais próximos de  $x^*$  no conjunto de dados histórico  $X$ , estima-se os pesos apropriados ou probabilidades para nomear a cada um destes vizinhos, e então reamostra-se o vector das vazões( no local  $s$  correspondente)  $M$ , estimando as probabilidade para cada um dos valores reamostrados. Por exemplo, suponha que o único predictor era o NINO3, desejamos emitir uma previsão para o próximo ano em janeiro que usa os dados para NINO3 os valores de OND. Digamos que o mais recente OND estimam de NINO3 é 2.5. Então localizaria na série histórica  $k$  (por exemplo, 30) dos valores mais próximos deste valor do NINO3. A probabilidade de cada um destes  $k$  anos selecionados sera proporcional ao quão próximo o valor do NINO3 estará do valor 2.5. Estes anos selecionados serão utilizados para a operação do sistema de reservatórios, ao invés de uma série



sortiada ao acaso. Os parâmetros chaves do algoritmo são o número de vizinhos,  $k$ , usar, a seleção da métrica para definir a "proximidade" nos espaço dos predictores, particularmente no contexto de análise multivariada, e os pesos da probabilidade. No contexto deste trabalho o procedimento utilizado é o descrito a seguir:

1. Cálculo da distancia  $d_i$  entre o valor dos vetor de predictores atual  $x^*$  e o valor dos predictores em sua série histórica,  $x_i$ , na forma:

$$d_i^2 = \sum_{j=1}^p \left\{ (x_j^* - x_{i,j}) \gamma_j \right\}^2 \quad (1)$$

onde  $x^*$  é um vetor  $1 \times p$ ,  $x_i$  é um vetor  $1 \times p$  de predictores durante o ano de  $i$ th usado no modelo ajustando, e  $\gamma$  é um vetor  $p \times 1$ . Para a regressão com dois predictores,  $\gamma = [\beta_1 \ \beta_2]^T$ , onde  $\beta_i$  é o  $i^o$  coeficiente da regressão da variável das vazões  $q$  padronizada e transformada e os índices climáticos EAD e NINO3, respectivamente, no caso da regressão agrupada ou da regressão do primeiro componente principal com os índices climáticos.

As distâncias registram a semelhança da condição de predictor atual com cada uma das condições passadas. A seleção das variáveis paramétricas, transformação e procedimento de regressão usados no passo preliminar são usados aqui para desenvolver "pesos" para cada componente da matriz dos predictores. Por conseguinte, a distância Euclidiana é usada para definir semelhança dos  $k$ -vizinhos mais próximos, e "transfere" o conhecimento da regressão paramétrica multivariada das vazões anuais.

2. usando o vetor de distância  $d$  calculado no passo anterior, identifica-se o conjunto de valores mais próximos identificando a posição em um ordenamento realizado a partir do índice  $J$  dos vizinhos mais próximos. O elemento de  $j$ th deste conjunto registra o ano  $t$  associou a  $j$ th com  $x_i$  mais próximo de  $x^*$ . O núcleo da reamostragem é definido no passo 3 baseado na ordem de elementos em  $J$ . Onde várias observações tem a mesma distância, o ordenamento original dos dados podem influenciar o ordenando em  $J$ . Para evitar tais problemas, nós copiamos o índices do vetor tempo  $t$  em um vetor temporário e proporcionamos uma perturbação randomica no ordenamento deste vetor antes de realizar o ordenamento do vetor pelas distâncias.

3. Seleciona-se o número de vizinhos a serem utilizados ( $k$ ) e o núcleo da reamostragem ou função de peso  $K(j)$ . A escolha núcleo da reamostragem segundo Lall e Sharma (1996):

$$K(j) = \frac{1/j}{\sum_{i=1}^k 1/i}$$

Uniforme:  $K(j)=1/k$

Combinações diferentes de  $k$  e  $K(j)$  podem dar resultados semelhantes, propriedades de núcleo são discutidas em Hardle (1991).

4. a matriz de previsão de vazões é então reamostrada utilizando o núcleo  $K(j)$ . Se o elemento de  $j$ th é tirado do núcleo, o ano correspondente é identificado de  $J$ , e a previsão é o conjunto de vazões ocorridas naquele ano e nos meses deste ano em todas as localidades. Este processo é repetido para gerar o número desejado de previsões de conjunto de  $a_{t+f}$  e  $m_{t+f}$ .

### **Resultados com o cálculo dos coeficientes utilizando a regressão agrupada**

A metodologia utilizada para o cálculo dos coeficientes utilizando Regressão Agrupada foi desenvolvida e apresentada em Souza Filho e Lall (2003).

Testou-se os valores de  $k$  que variam de 10 a 30, e os primeiros dois núcleos indicados anteriormente. As diferenças por tipo de núcleo utilizada são secundárias, e a previsão mediana é bem parecido com 10 ou 30 vizinhos. Observou-se que a mediana da previsão é sempre mais próxima dos valores observados que a mediana climatológica, e que a variação nos quantis da previsão são menores que os da climatologia. A correlação da previsão mediana com as observações é 0,91 durante os 8 anos de validação do modelo 1993-2000.

Os anos secos de 1993 e 1998 tiveram uma boa sinalização pelo modelo, enquanto a previsão durante os anos 1994 e 2000 foi mais difusa.

### **AGRADECIMENTO**

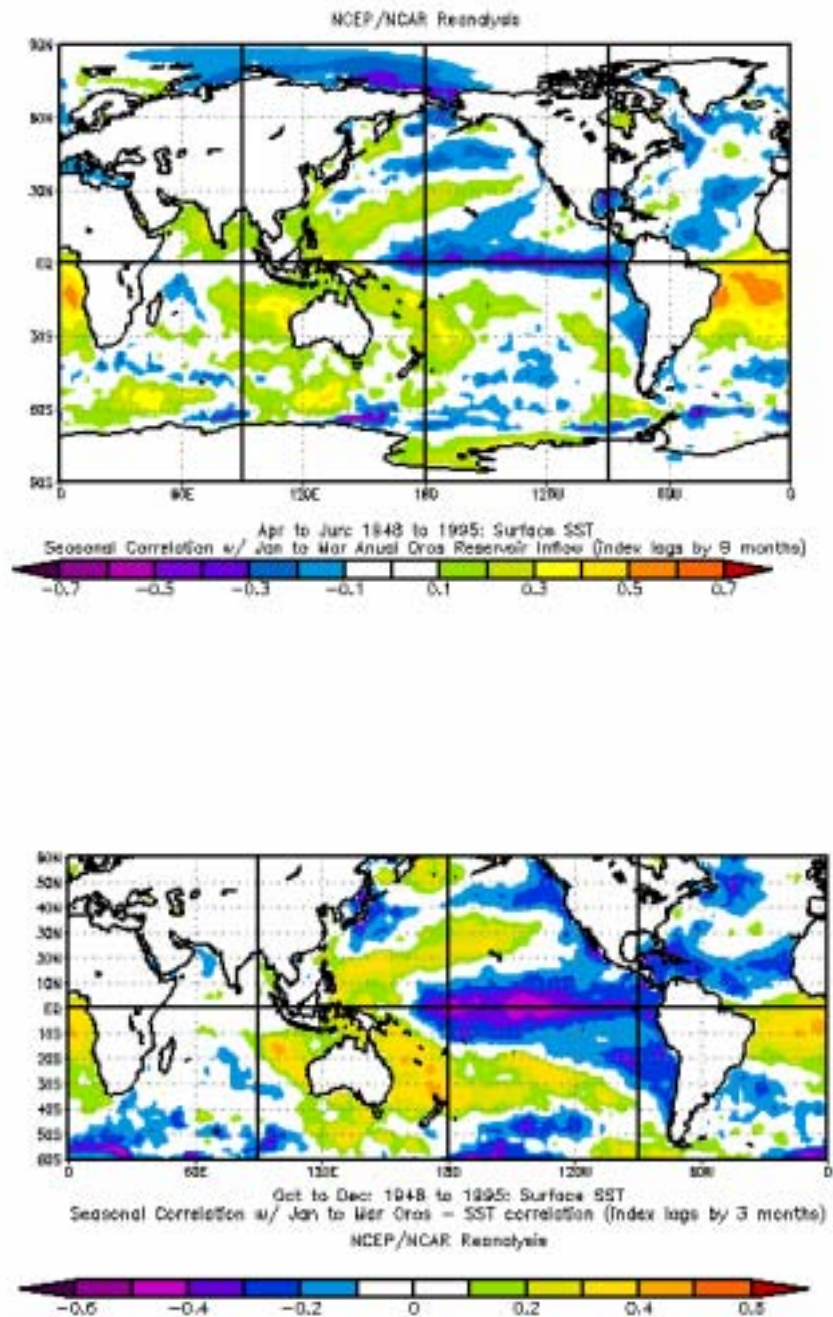
Agradecemos ao CNPQ pelo auxílio dado na elaboração das pesquisas que possibilitaram este trabalho.

### **REFERÊNCIAS BIBLIOGRÁFICAS**

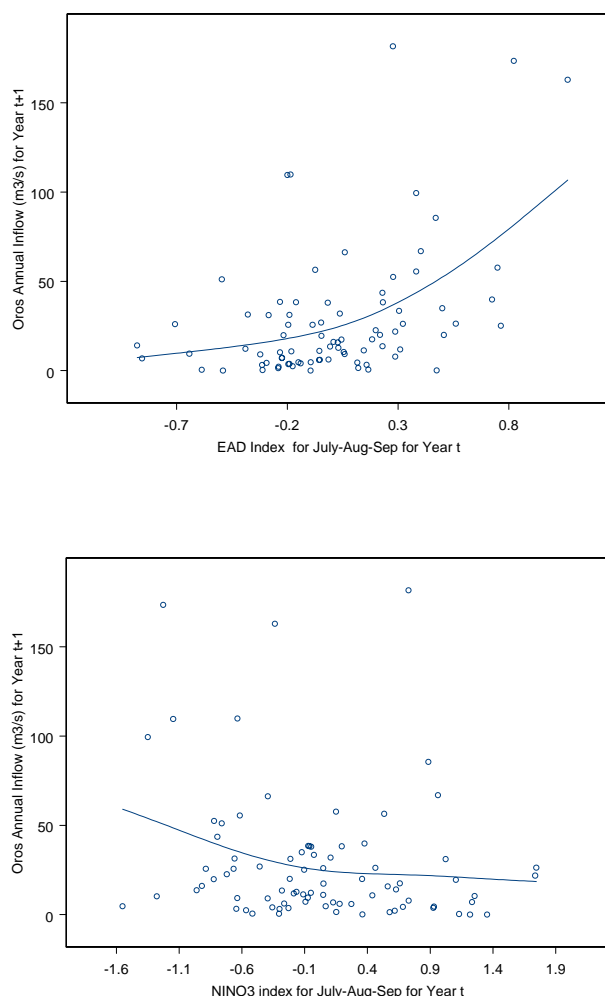
Chiang j. C. H., y. Kushnir e e. Zebiak, interdecadal changes in eastern pacific itcz variability and its influence on the atlantic itcz, geophysical research letters, 27 (22): 3687-3690 nov 15 2000.

- Cogerh, plano de gerenciamento da bacia do jaguaribe, companhia de gestão dos recursos hídricos, fortaleza. Ceará. 1999a.
- Cogerh, plano de gerenciamento das bacias metropolitanas, companhia de gestão dos recursos hídricos, fortaleza. Ceará. 1999b.
- Dillon, w.r e goldstein, m. Multivariate analyses. Pag 246-248. John wiley & sons. 1984.
- Hastenrath, s., predictability of northeast brazil droughts, nature, 307 (5951): 531-533, 1984.
- Hastenrath, s., prediction of northeast rainfall anomalies, j. Atmos. Sci., 35, 2222-2231, 1990.
- Kendall, d.r., e j.a. dracup, a comparison of index-sequential and ar(1) generated hydrological sequences, j. Hydrol., 122, 335-352, 1991.
- Kousky, v.e., frontal influences on northeast brazil, mon. Weather rev., 107, 1140-1153, 1979
- Kumar, d.n., u. Lall, and m. Peterson, multi-site disaggregation of monthly to daily streamflow, water resources research, 36(7), 1823-1834, 2000.
- Lall, u. E a. Sharma, a nearest neighbor bootstrap for resampling hydrologic time series, water resources research, 32(3), 679-693, 1996.
- Marengo, j. A., j. Tomasella, e c. R. Uvo, trends in streamflow and rainfall in tropical south america: amazonia, eastern brazil, and northwestern peru, j. Of geophysical res., 103(d2), 1775-1783, 1998.
- Markham, c. G., e d. R. Mclain, sea surface temperatures related to rain in ceara, northeast brazil, nature, 265, 320-323, 1977.
- Moura, a. D., e j. Shukla, on the dynamics of droughts in northeast brazil: observation, theory and numerical experiments with a general circulation model, j. Atmos. Sci., 38, 2653-2675, 1981.
- Nobre p. E j. Shukla, variations of sea surface temperature, wind stress, and rainfall over the tropical atlantic and south america, j. Climate, 9 (10), 2464-2479, 1996.
- Saravanan r. E p. Chang, interaction between tropical atlantic variability and el nino-southern oscillation, j. Climate, 13 (13), 2177-2194, 2000.
- Souza filho, f. A. And u. Lall, 2002, multi-scale, multi-site, seasonal to interannual probabilistic ensemble streamflow forecasts for ceara, brazil: applications of a multivariate, semi-parametric algorithm, water resources research, in press.
- Srh, plano estadual de recursos hídricos, secretaria de recursos hídricos do estado do ceará, fortaleza, ceará. 1991.
- Uvo, c. B. E n. E. Graham, seasonal runoff forecast for northern aouth america: a statistical model, water resources research, 34(12), 3515-3524, 1998.
- Uvo. C. B. , c. A. Repelli, s. E. Zebiak, e y. Kushnir, the relationship between tropical pacific and atlantic sst and northeast brazil monthly precipitation, j. Of climate, 11(4), 551-562, 1998.

- Uvo c. B., u. Tolle, e r. Berndtsson, forecasting discharge in amazonia using artificial neural networks, *int. J. Climatology*, 20 (12), 1495-1507, 2000.
- Wahba, g., spline methods for observational data, 169 pp., siam, philadelphia, 1990.
- Ward, m. N., s. Brooks, c. K. Folland, predictability of the seasonal rainfall in the northern nordeste region of brazil, in *recent climate change*, ed., s. Gregory, belhaven, london, uk, 237pp., 1988.
- Ward, m. N., e c. K. Folland, prediction of seasonal rainfall in the north nordeste of brazil using eigenvectors of sea-surface temperatures, *int. J. Of climatology*, 11, 711-743, 1991.



**Figure 1.** Correlações da Temperatura da Superfície do Mar (TSM) com as vazões do Oros com Oros no período de Janeiro a Junho com 9 e 3 meses de lags.



**Figure 2.** Correlação não linear entre as vazões anuais afluentes ao Oros e os Índices de Temperatura da Superfície do Mar(TSM). Relações similares são encontradas em outras estações. A curva é um spline suavizado ajustado aos dados.

**Tabela 1:** Dados Básicos e Estatísticos da afluência anual aos reservatórios ( $m^3/s$ ) baseados nos dados de 1913-1990

<b>RESERVATÓRIO</b>	<i>Oros</i>	<i>Banabuiu</i>	<i>Pedras Branca</i>	<i>Pacajus</i>	<i>Pacoti Riachao</i>	<i>Gaviao</i>
Rio	Jagaribe	Banabuiú	Sitiá	Choró	Pacoti-Riachão	Cocó
Area da Bacia(km <sup>2</sup> )	24563	14931	1787	4060	1108	95
Armazenamento (hm <sup>3</sup> )	1956	1800	434	148	420	54
1 <sup>st</sup> Quartio	6.9	4.5	0.7	2.9	2.8	0.4
Median	18.5	15.1	2.1	17.3	7.1	0.8
3 <sup>rd</sup> Quartio	37.2	34.3	5.9	32.7	12.0	1.5
MÉDIA	30.0	26.6	5.2	24.6	8.5	1.2
Std Dev	37.8	31.8	8.0	29.5	7.8	1.1
CV	1.3	1.2	1.5	1.2	0.9	0.9

**Tabela 2:** Correlação anual das vazões nos locais

	<i>Oros</i>	<i>Banabuiu</i>	<i>Pedras Branca</i>	<i>Pacajus</i>	<i>Pacoti Riachao</i>	<i>Gaviao</i>
<i>Oros</i>	1.00	0.76	0.78	0.78	0.65	0.64
<i>Banabuiu</i>	0.76	1.00	0.83	0.73	0.63	0.56
<i>Pedras Branca</i>	0.78	0.83	1.00	0.83	0.73	0.67
<i>Pacajus</i>	0.78	0.73	0.83	1.00	0.84	0.82
<i>Pacoti Riachao</i>	0.65	0.63	0.73	0.84	1.00	0.94
<i>Gaviao</i>	0.64	0.56	0.67	0.82	0.94	1.00

**Tabela 3.** Correlação entre a vazão anual do reservatório e os índices sazonais do clima. Encontra-se em negrito as correlações com nível de significancia superior a 95%.

	<i>EAD OND</i>	<i>Nino3 OND</i>	<i>EAD JAS</i>	<i>Nino3 JAS</i>	<i>EAD AMJ</i>	<i>Nino3 AMJ</i>
<i>EAD OND</i>	1					
<i>Nino3 OND</i>	0.08	1				
<i>EAD JAS</i>	<b>0.76</b>	0	1			
<i>Nino3 JAS</i>	0.05	<b>0.90</b>	-0.02	1		
<i>EAD AMJ</i>	<b>0.54</b>	0.01	<b>0.83</b>	0.02	1	
<i>Nino3 AMJ</i>	-0.10	<b>0.64</b>	-0.19	<b>0.74</b>	<b>-0.20</b>	1
<b>OROS</b>	<b>0.51</b>	<b>-0.21</b>	<b>0.47</b>	<b>-0.20</b>	<b>0.33</b>	<b>-0.23</b>

**Tabela 4.** Coeficientes da Regressão linear das variaveis transformadas e padronizadas com relação ao índice sazonal NINO3 and EAD.

<b>Data da Pevisão→</b>	<b>Janeiro</b>		<b>Outubro</b>		<b>Julho</b>	
	<b>Coef. EAD</b>	<b>Coef. Nino3</b>	<b>Coef. EAD</b>	<b>Coef. Nino3</b>	<b>Coef. EAD</b>	<b>Coef. Nino3</b>
Pacajus	1.14	-0.43	0.76	-0.33	0.39	-0.35
Pacoti-Riachão	1.28	-0.38	0.82	-0.26	0.48	-0.25
Gavião	1.46	-0.39	0.98	-0.30	0.59	-0.31
Pedras Branca	1.18	-0.32	0.87	-0.21	0.45	-0.23
Banabuiu	1.03	-0.19	0.92	-0.13	0.61	-0.16
Oros	1.32	-0.37	1.22	-0.37	0.62	-0.40
<i>Pooled</i>	1.23	-0.35	0.93	-0.27	0.52	-0.28