

## CAOS E PREVISIBILIDADE DE SÉRIES DE TEMPO HIDROLÓGICAS

**Francisco Jácome Sarmiento**

*Colaborador da Universität Hannover  
Inst. für Wasserwirtschaft, Hydrologie und landwirtschaftl. Wasserbau  
Appelstrasse 9A  
30167 - Hannover - Deutschland / Fax: 0049 511 762 3731*

### RESUMO

*No presente trabalho é feita uma breve revisão dos principais conceitos relacionados com a identificação analítica de comportamento caótico para sistemas dinâmicos não-lineares. A abordagem está centrada nos sistemas dinâmicos para os quais não se dispõe de uma formulação matemática e sim apenas da série de tempo como representativa da evolução do sistema, caso dos fenômenos naturais de interesse na Engenharia de Recursos Hídricos.*

*No contexto da identificação do comportamento caótico são considerados a determinação da dimensão de Hausdorff, estimativa dos expoentes de Lyapunov e entropia de Kolmogorov. A aplicação destes procedimentos utilizando uma série sintética de índices pluviométricos estimada para uma bacia da região semi-árida do Nordeste do Brasil é apresentada e comentada em seus aspectos concernentes à eficiência na investigação da presença de um atrator estranho de baixa dimensão.*

*A relação entre a presença de um atrator estranho e a possibilidade de previsão de estados futuros de sistemas para os quais dispõe-se apenas da série de tempo como representativa de sua dinâmica é explanada.*

*A adequabilidade de um modelo de previsão baseado em aproximação local utilizando os vizinhos mais próximos no espaço de fase é analisada para uma série sintética de índices pluviométricos. Aspectos relacionados com a função representativa dos erros de previsão e estrutura da correlação entre valores previstos e reais para diferentes passos de previsão mensal são apresentados.*

*As previsões foram realizadas para passos de 1 e 2 meses.*

*Com base na premissa segundo a qual processos verdadeiramente estocásticos apresentam erros de previsão independentes do passo de tempo utilizado, o modelo mostra-se também útil na identificação de comportamento caótico de séries temporais.*

### CAOS NO CONTEXTO DA CIÊNCIA MODERNA

Em um passado relativamente recente pesquisadores de diversos ramos da ciência têm voltado sua atenção para os sistemas dinâmicos não lineares especificamente em decorrência de uma nova formulação dos modelos que os representam. Tal formulação está baseada na Teoria do Caos. Desde sua inicialização, a dinâmica não linear caótica tem se expandido rapidamente e muitas aplicações são esperadas no futuro. O impacto dessa teoria na ciência moderna se deve, entre outras razões, certamente também por envolver a possibilidade de previsão de sistemas dinâmicos até então possíveis de serem tratados apenas sobre bases probabilísticas e apresentarem técnicas capazes de identificar ordem onde outrora a ciência clássica encontrava apenas o aleatório, decorrendo daí a afirmação de que o Caos começa onde a ciência clássica pára.

Particularmente em Engenharia de Recursos Hídricos a análise de séries de tempo hidrológicas através de modelos baseados nesta nova técnica vem mostrando a possibilidade concreta de previsões de valores futuros, pelo menos a curto prazo. Esse conceito complementa a concepção do mundo formulada por Laplace no século XVIII

segundo a qual o estado momentâneo de um sistema natural é uma consequência do que ele foi no instante anterior e se imaginamos uma inteligência que conheça para um dado instante de tempo todas as relações entre as partes do Universo, ela poderia assim prever o local, o movimento e as relações gerais entre todas estas partes para todos os instantes de tempo no passado e no futuro.

Nesse contexto os fenômenos naturais que tais séries representam são vistos não mais como uma realização de um processo estocástico, mas sim como um sistema dinâmico altamente sensível às condições iniciais. Dependência sensível das condições iniciais significa na prática que pequenos erros no conhecimento do estado atual de um sistema crescem exponencialmente com o tempo, o que inviabiliza previsões a longo prazo.

### **ATRADORES ESTRANHOS - DIMENSÃO FRACTAL**

Por definição os atratores são parcelas compactas do espaço dimensional de evolução  $R^n$  de um sistema dinâmico para onde convergem as trajetórias representativas dos vários estados assumidos pelo sistema ao longo do tempo. Muitos sistemas dinâmicos podem apresentar atratores cuja topologia é irregular e sua configuração pode ser tal que ocupa uma dimensão intermediária, entre os valores inteiros considerados pela geometria euclidiana, uma dimensão fractal. Os atratores que possuem uma dimensão que não é inteira são denominados atratores estranhos. Lorentz (1963) foi o primeiro a constatar a existência de um atrator estranho em um modelo físico. O conceito de dimensão fractal foi introduzido por Mandelbrot (1982) e visto de uma forma prática, a dimensão fractal mede a eficiência de um objeto em ocupar o espaço (Gleick, 1987). Esse conceito mostra-se mais adequado à caracterização geométrica de objetos naturais irregulares como nuvens e campos de precipitações do que o conceito euclidiano de dimensão. Entre as diversas

formas de expressão da característica de dimensão, a chamada dimensão de capacidade é sem dúvida a mais simples (Froehling et al., 1981).

### **COMPORTAMENTO CAÓTICO DE SÉRIES DE TEMPO**

Em sua totalidade os fenômenos físicos naturais de interesse na área de Engenharia Hidrológica são caracterizados através de medidas realizadas ao longo do tempo e discretizadas em um intervalo desejado. As séries de tempo provenientes de um sistema dinâmico podem representar realizações de um processo estocástico, neste caso a aleatoriedade do processo deriva de complicados impulsos envolvendo muitos graus de liberdade (Farmer e Sidorowich, 1987). Sob essa abordagem as interações desses impulsos são tomadas como uma estrutura probabilística randômica. Sob a ótica da Teoria do Caos é investigada a possibilidade de que a série de tempo contenha relações determinísticas simples que envolvam poucos graus de liberdade.

A dinâmica não linear que essas séries representam é capaz de produzir, tanto na dimensão do tempo como do espaço, valores não correlacionados, criando assim a aparência de um processo estocástico. As técnicas estatísticas convencionais são incapazes de distinguir entre essa aparência e o comportamento de um verdadeiro processo estocástico (Rodrigues-Iturbe et al., 1989). A busca de alguma ordem subjacente que permaneça no sistema é investigado como Caos Determinístico. A viabilidade do uso das técnicas associadas a esse novo ramo de análise de sistemas dinâmicos é evidenciada pela presença ou não de um atrator caótico que se constitui na própria ordem em meio a desordem.

### **Previsibilidade**

O comportamento caótico de atratores de dimensão fractal tem como característica

fundamental a divergência de trajetórias inicialmente vizinhas. Assim se tomarmos dois pontos em trajetórias vizinhas num atrator separados inicialmente de uma pequena distância  $\delta$  para  $t=0$  a divergência das trajetórias na evolução ao longo do tempo dá-se de forma exponencial (Grebogi et al., 1987). Essa propriedade limita a previsibilidade de tais sistemas a longo prazo. Mesmo em se tratando de sistemas que possuem descrição analítica via equações diferenciais, quando exibem tais atratores não podem irrestritamente serem tomados como determinísticos, pois mesmo independentes de perturbações exógenas, eles não são nem periódicos nem quasi-periódicos (Tsonis e Elsner, 1988).

De uma maneira prática essa característica de divergência das trajetórias é essencial na construção de um conceito de Caos, pois reflete o fenômeno do surgimento de aleatoriedade e imprevisibilidade até mesmo em sistemas totalmente determinísticos através do que Lorentz (1963) chamou de dependência sensível das condições iniciais, o que significa que a desconsideração de pequenas perturbações no sistema pode levar a grandes erros na previsão de estados futuros. No contexto de previsões meteorológicas tal comportamento foi denominado por Lorentz (1979) Efeito Borboleta.

A dependência linear entre valores sucessivos na série de tempo é medida pela função de autocorrelação. A forma com a qual essa função decai reflete o efeito de persistência, que pode ser constatado tanto para intervalos de discretização pequenos como grandes (por exemplo séries anuais) para séries de tempo dessa natureza. Essa propriedade foi denominada efeito José por Benoit Mandelbrot. O chamado efeito José tem, atuando em direção contrária, o efeito Noé, que significa descontinuidade do comportamento. Quando tais efeitos se propagam entre diferentes escalas de discretização, ou seja, as irregularidades são associadas entre as diferentes escalas, tem-se um decaimento mais lento da função

autocorrelação em relação à queda exponencial. Essa propriedade foi também reconhecida por Mandelbrot e é denominada Autosemelhança.

Em relação à propriedade de sensível dependência das condições iniciais de um sistema, a quantificação da dinâmica do atrator em termos de criação ou perda de informação pode ser feita tanto através do expoente de Lyapunov como também pela entropia de Kolmogorov. Em uma abordagem conjunta o expoente de Lyapunov reflete a instabilidade das trajetórias enquanto que a entropia de Kolmogorov mede a velocidade de perda de informação sobre um estado inicial ao longo do tempo (Kurths e Herzel, 1987).

### Espaço de fase e dimensão de inserção

Como a formulação matemática exata do sistema dinâmico que as séries hidrológicas de tempo representam é sempre desconhecida, o espaço de estado utilizado normalmente para a análise de sistemas de equações diferenciais é substituído pelo espaço de fase. Segundo essa concepção, a evolução do sistema não linear observado é representado por um conjunto de vetores  $p$ -dimensionais  $Y$  construídos a partir da série de tempo  $x(t)$ ,  $t=1,2,\dots,n$ , como sendo (Packard et al., 1980):

$$Y(t) = [x(t), x(t + \tau), x(t + 2\tau), \dots, x(t + (p-1)\tau)]$$

onde  $p$  é a dimensão do espaço de fase e  $\tau$  é o intervalo de tempo de defasagem, escolhido de tal forma que os eixos coordenados do espaço de fase tenham um substituto discreto equivalente às coordenadas usuais  $x(t)$ ,  $dx/dt, d^2x/dt^2, \dots, d^{(p-1)}x/dt^{(p-1)}$  (Abarbanel et al., 1990).

Não há no atual nível de conhecimento um critério para a escolha do intervalo de defasagem  $\tau$ . Em estudos meteorológicos adota-se  $\tau$  como sendo o valor para o qual o decaimento exponencial da função autocorrelação atinge um valor menor do que  $1/e$  (Tsonis e Elsner, 1988). Um outro critério

arbitrário proposto por Abarbanel et al. (1990) é adotar  $\tau$  entre 1/10 e 1/20 do tempo no qual ocorre o primeiro mínimo local na função de autocorrelação.

### Dimensão de Hausdorff e a determinação de p

A dimensão do espaço de inserção p deve ser tal que as trajetórias em  $R^p$  não se interceptem, o que seria uma contradição em se tratando de um modelo determinístico (Loistl e Betz, 1993). Por definição a projeção de um atrator estranho em um espaço de dimensão infinito possui dimensão finita. Então, se o sistema possui um atrator, a dimensão p deve ser suficientemente grande para envolver todas as informações geométricas sobre o atrator.

A dimensão característica do atrator necessária para a definição de p é quantitativamente representada pela dimensão Hausdorff. O método proposto por Grassberger e Procaccia (1983) para distinguir comportamento estocástico de caos determinístico é normalmente usado na definição do valor de saturação de p, ou seja, o valor mínimo da dimensão do espaço de inserção no qual o atrator (se ele existir) pode ser inserido sem perda de informação. O método é baseado na correlação integral  $C(r)$ , onde r é o raio de um pequeno círculo (uma esfera para  $p=3$  ou hiperesfera para  $p \geq 4$ ) com centro em um ponto arbitrário do conjunto de vetores Y no espaço de fase. Para um dado número de valores da série n e uma dimensão de inserção p tem-se:

$$C(r) = \frac{2}{n(n-1)} \sum_{i=1}^n \sum_{j=1}^n \theta(r - \|Y(j) - Y(i)\|)$$

onde  $\|$  representa a norma euclidiana dos vetores  $Y(i)$  e  $Y(j)$  e  $\theta(h)$  é a função Heaviside, que assume valor nulo se  $h \leq 0$  e valor unitário se  $h > 0$ . Para cada ponto de centragem, a função  $\theta(h)$  conta simplesmente o número de pontos da trajetória no espaço de fase contidos dentro do raio r. No caso

bidimensional o número de pontos separados por uma distância r e  $C(r)$  serão proporcionais a  $r^2$ , no caso tridimensional a  $r^3$  e para o caso geral tem-se  $C(r) \propto r^v$ , onde  $v=p$  é a chamada dimensão de correlação (Grassberger e Procaccia, 1983) dada por:

$$v = \lim_{r \rightarrow 0} \frac{\log C(r)}{\log r}$$

No caso de um processo estocástico o limite acima tende para infinito indicando a ausência de uma estrutura de correlação. A presença de uma dinâmica determinística em uma série de tempo manifesta-se através do paralelismo entre as plotagens  $\log(C(r))$  x  $\log(r)$  a partir de uma determinada dimensão do espaço de fase, o que equivale a atingir um valor de saturação no gráfico  $v$  x  $p$ .

### Dependência das condições iniciais

O grau de sensibilidade de um sistema às condições iniciais é medido pelo chamado expoente de Lyapunov. O expoente é definido a partir da consideração do deslocamento de uma p-esfera infinitesimal no espaço de fase, que devido ao fluxo de fase irá se deformar tornando-se um p-elipsóide. O comprimento assumido pelos eixos principais do elipsóide  $l_i(t)$  define o primeiro expoente de Lyapunov unidimensional:

$$\lambda_1 = \lim_{t \rightarrow \infty} \log_2 \frac{l_1(t)}{l_1(0)}$$

Os expoentes de Lyapunov  $\lambda_i$  descrevem a convergência exponencial ( $\lambda < 0$ ), a neutralidade ( $\lambda = 0$ ) ou a divergência ( $\lambda > 0$ ) de duas trajetórias inicialmente vizinhas pertencentes a um sistema dinâmico (Loistl e Betz, 1993). A presença de uma dinâmica caótica é evidenciada, portanto, pela existência de um expoente de Lyapunov positivo.

Alguns algoritmos para estimativa dos expoentes de Lyapunov são apresentados e

discutidos em Abarbanel et al. (1990). O primeiro e até então mais conhecido algoritmo com a finalidade de estimar o expoente de Lyapunov de uma série de tempo foi proposto por Wolf et al. (1985). O maior expoente de Lyapunov é obtido para um dado instante de tempo  $k$  através da seguinte relação:

$$\hat{\lambda} = \frac{1}{t_k - t_0} \sum_{k=1}^N \log_2 \frac{d(t_k)}{d(t_{k-1})}$$

### Entropia de Kolmogorov

Um outro método de análise de comportamento caótico é a chamada entropia de Kolmogorov. A estimativa dessa grandeza para sistemas analiticamente definíveis é relativamente fácil em comparação a sua estimativa a partir de uma série de tempo. A entropia de Kolmogorov também conhecida como entropia  $K$  é definida para um sistema dinâmico com  $p$  graus de liberdade através das probabilidades de que as trajetórias  $x(t)$  pertencentes à zona de influência de um atrator visitem regiões do espaço de fase. Para o estabelecimento de tais probabilidades o espaço de fase  $p$ -dimensional é dividido em caixas ou hipercaixas de dimensão  $\varepsilon^p$ . A evolução das trajetórias no espaço de fase é agora observada ao longo de intervalos de tempo discretos  $\Delta t$ . A entropia de Kolmogorov  $K$  é dada por:

$$K = - \lim_{\varepsilon \rightarrow 0} \lim_{\Delta t \rightarrow 0} \lim_{N \rightarrow \infty} \frac{1}{d \Delta t} \sum_{i_1, \dots, i_d} w(i_1, \dots, i_d) \ln w(i_1, \dots, i_d)$$

onde  $w(i_1, \dots, i_d)$  representa a probabilidade conjunta de que para  $t = \Delta t$  a trajetória  $x(t)$  esteja na caixa  $i_1$ , para  $t = 2\Delta t$  a trajetória  $x(t)$  esteja na caixa  $i_2, \dots$  e para  $t = d\Delta t$  a trajetória  $x(t)$  esteja na caixa  $i_d$ . Atratores estranhos como suscitadores de entropia criam imprevisibilidade; esta está diretamente relacionada com a capacidade que os sistemas dinâmicos dessa natureza apresentam de "criarem" informação.

No caso da disponibilidade apenas de uma série de tempo, Grassberger e Procaccia (1983) definiram uma entropia de Kolmogorov de segunda ordem  $K_2$  com as seguintes propriedades: a)  $K_2 > 0$ , b)  $K_2 < K$ , c) para sistemas randômicos  $K_2$  é infinito e d)  $K_2 \neq 0$  para sistemas caóticos. O algoritmo proposto é essencialmente o mesmo usado na definição da dimensão de correlação.

### MODELOS DE PREVISÃO

A previsão para séries de tempo caóticas tem sido proposta através de modelos determinísticos não lineares desde 1987 (Farmer e Sidorowich, 1987; Crutchfield e McNamara, 1987; Casdagli, 1989; Sugihara e May, 1990). A abordagem mais utilizada tem sido os chamados métodos de aproximação local onde a relação funcional (equação 1) entre um estado presente do sistema e o futuro é aproximada por um número  $k$  de variáveis vizinhas.

$$x_{i+k} = f^k(x_i, x_{i-1}, \dots, x_{i-m-i}) \quad (1)$$

O procedimento de adoção de polinômios de ordem  $k$  no espaço de fase  $R^p$  para ajustamento do conjunto de dados tem sido descartado por vários autores por se mostrar desvantajoso na previsão, pois os erros de aproximação crescem exponencialmente (Farmer e Sidorowich, 1987). Uma dificuldade adicional surge para o caso da dimensão  $p$  ser grande, pois o número de parâmetros a serem ajustados pode se tornar intratável.

O modelo de Casdagli (1991) foi proposto como um algoritmo de previsão capaz de estabelecer uma ponte entre abordagens de natureza determinística e estocástica. O modelo aproxima a equação (1) por  $k$  vizinhos mais próximos, onde valores pequenos de  $k$  correspondem a uma modelagem determinística enquanto que valores elevados levam a um modelo estocástico linear. O algoritmo é também utilizado na identificação de comportamento

caótico de baixa dimensão, partindo-se da premissa de que sistemas com dinâmica próxima do extremo determinístico devem apresentar a curto prazo previsões mais precisas traduzidas por uma função de erros integrada aos passos do algoritmo resumidos a seguir (Casdagli, 1991):

- a. A série de tempo disponível deve ser dividida em duas partes: um conjunto de ajustamento  $x_1, \dots, x_{N_f}$  e um conjunto de teste  $x_{N_f+1}, \dots, x_{N_f+N_t}$ ;
- b. Adotar uma dimensão de inserção  $p$ , um tempo de retardo  $\tau$  e um tempo de previsão  $T$ ;
- c. Escolher um vetor de retardo  $Y_i$  com  $i \geq N_f$  para um teste de precisão com  $T$ ;
- d. Calcular a distância  $d_{i,j}$  do vetor de teste  $Y_i$  ao vetor de retardo  $Y_j$  com  $1 + (p-1)\tau \leq j \leq N_f - T$  no conjunto de ajustamento;
- e. Ordenar as distâncias  $d_{i,j}$  e selecionar os  $k$  vizinhos mais próximos  $Y_{j(1)}, \dots, Y_{j(k)}$  de  $Y_i$  e ajustar um modelo com a seguinte forma:

$$x_{j(i)+T} \approx \alpha_0 + \sum_{n=1}^k \alpha_n x_{j(i)-(n-1)\tau} \quad \ell = 1, \dots, k \quad (2)$$

onde o número de vizinhos mais próximos  $k$  deve variar para valores representativos entre  $2(p+1)$  e  $N_f - T - (m-1)\tau$

- f. O modelo traduzido pela equação (2) é agora utilizado para estimar com passo  $T$  o valor  $Y_{i+1}(k)$  para o vetor de teste  $Y_i$ , calculando-se em seguida o erro  $e_i(k)$ , dado pela diferença absoluta entre a valor previsto e o real;
- g. Repetir os passos de c) a f) para todo  $i$  no conjunto de teste e calcular a raiz quadrada do erro de previsão normalizada por:

$$E_w(k) = \frac{\left\{ \sum_i e_i^2(k) \right\}^{1/2}}{\sigma} \quad (3)$$

onde  $\sigma$  é o desvio padrão da série de tempo.

A função de erros (Equação 3) pode ser usada como índice direcionador para a fixação de valores ótimos para parâmetros como  $\tau$ , desta forma o procedimento consiste em tentativa e erro.

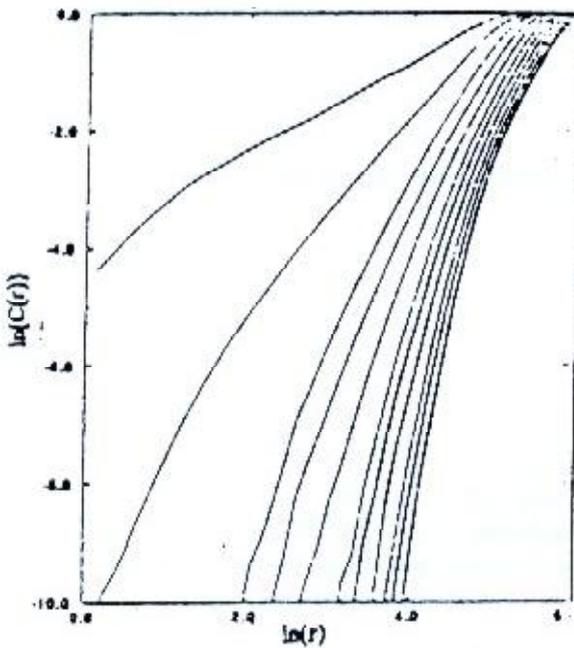
## APLICAÇÃO

Os procedimentos acima descritos foram aplicados a uma série de precipitações sintéticas (924 valores mensais) denominada Vetor Regional (Hiez e Rancan, 1987). A relação  $\ln(C(r)) \times \ln(r)$  estimada sobre o vetor regional é mostrada na Figura 1a para os espaços de inserção variando de 2 a 22. O "delay" de tempo utilizado foi  $\tau=2$ , valor para o qual a função autocorrelação da série atinge valores menores do que  $1/e$ . A derivada da relação  $\ln(C(r)) \times \ln(r)$  para a série filtrada por ajuste polinomial é mostrada na Figura 1b na qual observa-se uma leve tendência de saturação entre as dimensões 2 e 3, o que indicaria a possibilidade de existência de um atrator estranho de dimensão fractal neste intervalo. Convém salientar que a série investigada inclui os ruídos associados aos erros de medição e imprecisão de 63 séries (bacia do Acaraú-Ceará-Brasil) utilizadas na sua estimativa.

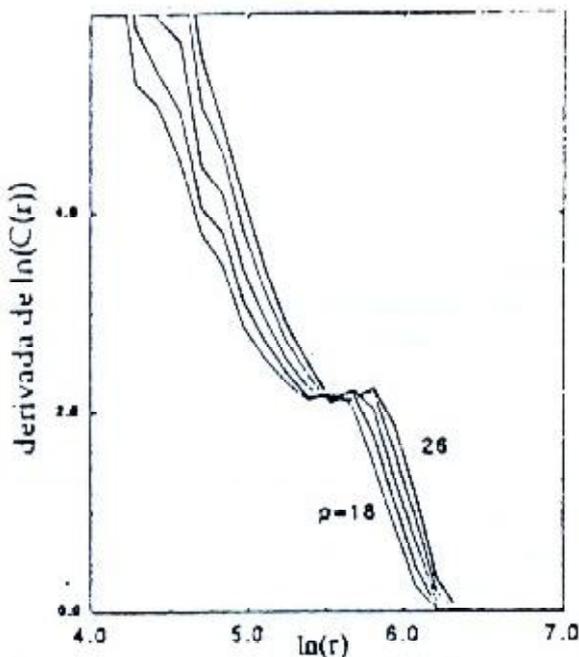
O efeito do acentuado percentual de ruídos na série formada pelo vetor regional tem certamente influência na divergência das derivadas após as mesmas parcialmente tenderem a estabilização na dimensão aproximadamente igual a 2,2 (utilizando a série filtrada). A estimativa da entropia de Kolmogorov de segunda ordem conforme proposta por Grassberger e Procaccia (1983), por ser baseada nas curvas  $\ln(C(r)) \times \ln(r)$ , não leva a resultados conclusivos.

Admitindo-se uma dimensão de inserção igual a 3, conforme apontado pela Dimensão de Hausdorff, o expoente de Lyapunov máximo foi avaliado para diversos valores do parâmetro representativo do passo de evolução usado no algoritmo de pesquisa e

traduzido pela variável EVOLV no programa FORTRAN fornecido pelo autor.



(a)



(b)

Figura 1. a) Relação  $\ln(C(r)) \times \ln(r)$ , b) derivada da relação  $\ln(C(r)) \times \ln(r)$  e derivada filtrada da relação  $\ln(C(r)) \times \ln(r)$ .

A Figura 2 mostra a convergência do expoente máximo em função do número de valores considerados e com  $EVOLV=2$ . O limite de previsibilidade do sistema dado por  $\lambda^{-1}$  seria  $1/0,62 \approx 1,7$  meses.

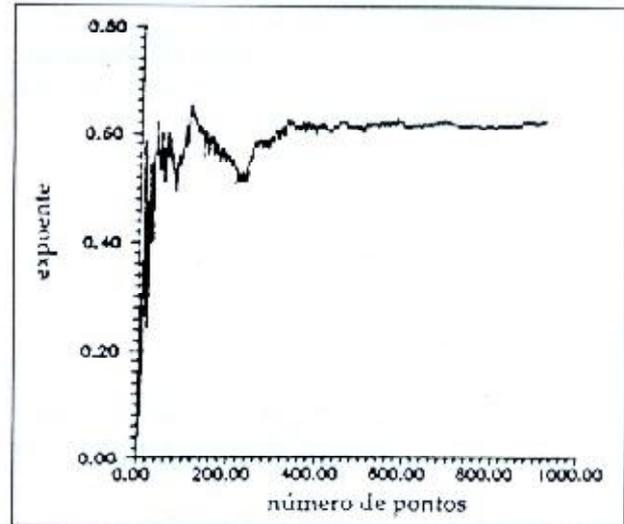


Figura 2. Expoente de Lyapunov em função do número de vizinhos. Dimensão de inserção 3 e  $\tau=2$ .

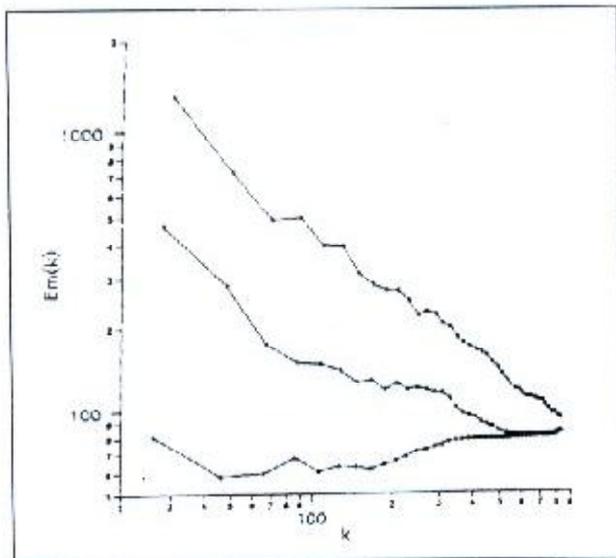
O modelo de previsão descrito foi aplicado a série sintética de índices pluviométricos já mencionada. A série disponível foi dividida em 3 partes, a primeira com 70 anos (840 meses) constituiu o grupo de ajustamento, a segunda 4 anos (48 meses) corresponde ao conjunto de teste e finalmente os três últimos anos são usados na validação das previsões.

O erro de previsão normalizado  $E_m(k)$  para as dimensões de inserção 2,3 e 4 é mostrado na Figura 3a. Observa-se que para um número de vizinhos  $k > 500$  os erros obtidos com as dimensões 2 e 3 se equivalem. Conforme assinalado por Casdagli (1991) a dimensão da série utilizada seria adequada caso a mesma fosse livre de ruídos (dimensão da série  $\approx 10^9$ ).

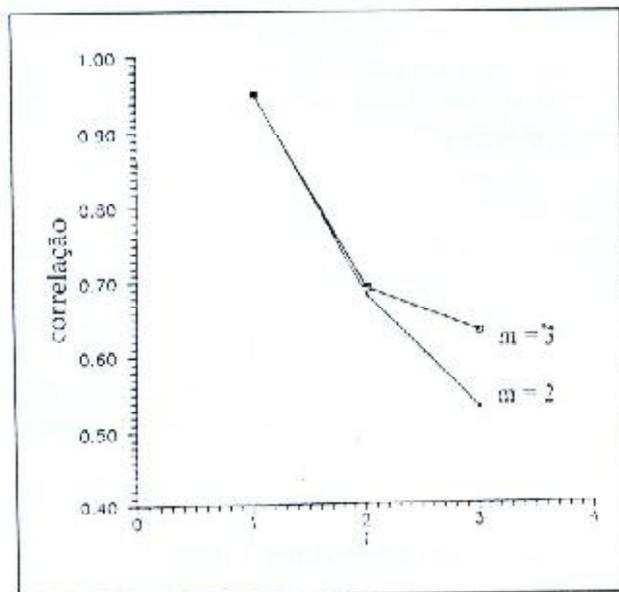
### Previsão com passo de 1 mês

O comportamento caótico da série analisada é evidenciado através do uso de modelos de previsão partindo-se da premissa

de que previsões em passo de tempo pequeno devem apresentar uma correlação melhor com o valor real do que previsões com passo de tempo mais longo. Se a série é aleatória o erro de previsão não depende do passo de tempo utilizado. A Figura 3b mostra os coeficientes de correlação obtidos na previsão de três anos com passo de tempo de 1, 2 e 3 meses utilizando 750 vizinhos no espaço de inserção. Conforme se pode



(a)



(b)

Figura 3. a) Erro normalizado (dimensões 2, 3 e 4) e b) Correlações.

observar, em termos de correlação as dimensões de inserção 2 e 3 fornecem para o passo de previsão de 1 e 2 meses praticamente a mesma correlação (0,95). A Figura 4a mostra o ajustamento entre os valores real e previstos para as dimensões de inserção 3.

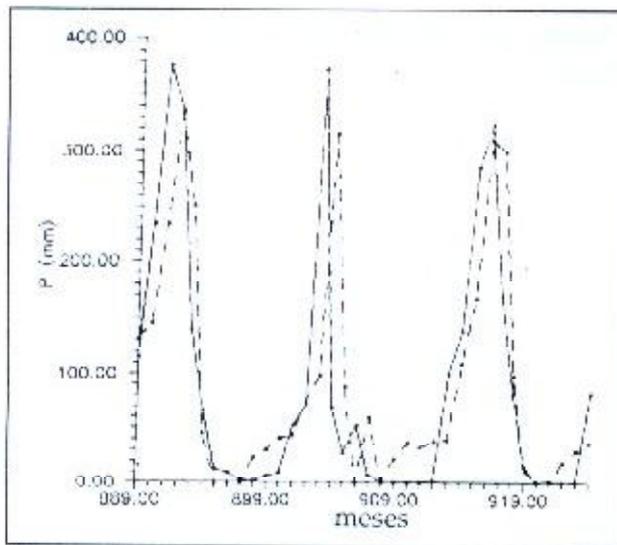
### Previsão com passo de 2 meses

A previsão de valores para um intervalo de 2 meses foi realizada em duas etapas, prevendo-se inicialmente o primeiro mês e utilizando-se esse valor na previsão do segundo mês. Esse procedimento difere daquele que conduziu a correlação mostrada na Figura 3b (0,69), através do qual a previsão do segundo mês futuro é feita diretamente. A correlação utilizando o procedimento em duas etapas atingiu 0,73. Porém, em termos de ajustamento os resultados deixam bastante a desejar, conforme mostra a Figura 4b. Essa constatação coincide com a previsibilidade indicada pelo expoente de Lyapunov estimada para esta série.

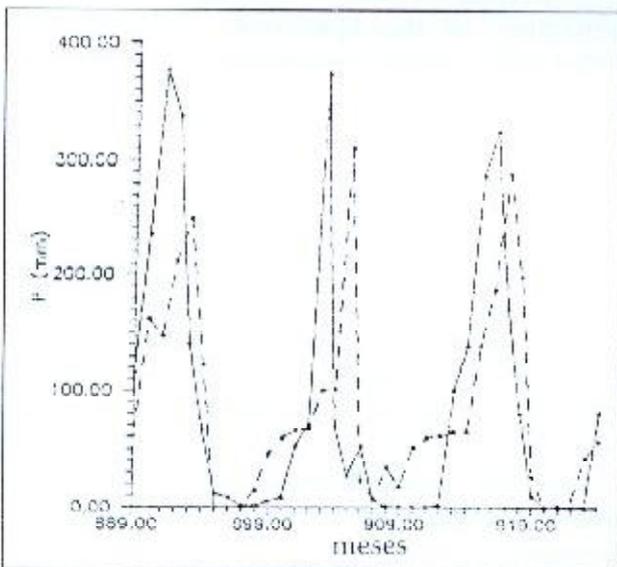
### CONCLUSÕES

As técnicas mais comuns de análise de comportamento caótico determinístico (dimensão de Hausdorff, Entropia de Kolmogorov e expoente de Lyapunov) quando aplicadas a série de tempo hidrológica analisada têm sua eficiência limitada pelo nível de ruído incorporado a mesma. Na aplicação aqui realizada, a análise dos resultados provenientes desses procedimentos fornece uma indicação não conclusiva sobre presença de um atrator estranho de baixa dimensão no sistema dinâmico representado pelo vetor regional. Embora a constatação de um expoente de Lyapunov máximo positivo represente um indicador da presença de uma dinâmica caótica no sistema, estabelecendo seu limite de previsibilidade em menos de 2 meses, sujeito aos erros de previsão expressos pelos

coeficientes de correlação comentados na seção anterior.



(a)



(b)

Figura 4. a) Previsões com passo de 1 mês e b) 2 meses (dimensão 2).

O modelo de previsão baseado nos vizinhos próximos mostra-se também ser um indicador do comportamento caótico de sistemas, considerando a premissa de que em processos verdadeiramente estocásticos os erros de previsão não apresentam nenhuma dependência do passo de tempo, porém para a série aqui analisada deve-se considerar um possível efeito da periodicidade. Para todas as

dimensões de inserção estudadas a função de correlação entre os valores reais e previstos apresentou-se decrescente para passos de previsão crescentes, refletindo a velocidade de perda de informação do sistema.

Dentro das limitações impostas pelo comprimento da série e a presença de ruídos os resultados obtidos na previsão em passo de tempo de 1 e 2 meses são compatíveis com o limite de previsibilidade indicada pelo expoente de Lyapunov. O coeficiente de correlação entre a subsérie de validação e os valores previstos com espaço de inserção 3 e passo de previsão 1 embora seja surpreendentemente elevada (0,95) deve ser visto com reservas com relação ao seu uso prático, pois trata-se de técnicas muito recentes cujas bases teóricas ainda possuem muitas questões em aberto.

Sob o ponto de vista da prática da Engenharia Hidrológica valores previstos para o vetor regional representam na realidade uma avaliação paralela de valores futuros para os  $m'$  postos de medição que entraram na estimativa do vetor. Em uma análise qualitativa, modelos de previsão que venham a fornecer previsões mais seguras para os meses chuvosos podem por exemplo levar a conclusões sobre a ocorrência ou não de seca em regiões semi-áridas como o Nordeste do Brasil onde mais de 70% das chuva anual ocorre em apenas 3 meses.

## REFERÊNCIAS

- ABARBANEL, H. D. I., BROWN, R., KADTKE, J. B. (1990) Prediction in chaotic nonlinear systems: Methods for time series with broadband Fourier spectra. *Phys. Rev. A*, 41(4), p.1782-1807.
- CASDAGLI, M. (1989) Nonlinear prediction of chaotic time series. *Physica D*, 35, p. 335-356.
- CASDAGLI, M. (1991) Chaos and Deterministic versus stochastic Non-linear Modelling. *J. R. Statist. Soc. B*, 54, No. 2, p.303-328.
- CRUTCHFIELD, J. P., McNAMARA, B. S. (1987) Equations of motion from a data series. *Complex Syst.*, 1, p. 417-452.

- FARMER, J. D., SIDOROWICH, J. J. (1987) Predicting chaotic time series. *Phys. Rev. Lett.*, 59(8), p. 845-848.
- FROEHLING, H., CRUTCHFIELD, J. P., FARMER, J. D., PACKARD, N. H., SHAW, R. (1981) On determining the Dimension of chaotic flows. *Physica* 3D, p. 605-617.
- GLEICK, J. (1987) *Chaos, Making a new Science*. Viking: New York.
- GRASSBERGER, P., PROCACCIA I. (1983) Estimation of the Kolmogorov entropy from a chaotic signal. *Phys. Rev. A*, 28, p. 2591-2593.
- GRASSBERGER, P., PROCACCIA I. (1983) Measuring the strangeness of strange attractors. *Physica*, 9D, p.189-208.
- GREBOGI, C., OTT, E., YORKE, J. A. (1987) Chaos, strange Attractors, and fractal basin Boundaries in Nonlinear Dynamics. *Science*, vol 238, p. 632-638.
- HIEZ, G. L. G., RANCAN, L. (1983) Aplicação do Vetor Regional no Brasil. In: V SIMPÓSIO BRASILEIRO DE RECURSOS HÍDRICOS, Anais, Vol. 2.
- KURTHS, J., HERZEL H. (1987) An attractor in a solar time series, *Physica*, 25D, p. 165-172.
- LOISTL, O., BETZ, I. (1993) *Chaosteorie: zur Theorie nichtlinearer dynamischer Systeme*. R Oldenbourg Verlag Munchen Wien.
- LORENZ, E. N. (1963) Deterministic non-periodic flow. *J. Atmos. Sci.*, 20, p. 130-141.
- LORENZ, E. N. (1979) Predictability: Does the flap of a butterfly's wings in Brazil of a tornado in Texas? In: Annual Meeting, Am. Assoc. for the Adv. of Sci., Washington, D.C., December 29.
- MANDELBROT, B. B. (1982) *The Fractal geometry of Nature*. Freeman, Cooper, San Francisco, Calif.
- PACKARD, N. H., CRUTCHFIELD, J. P., FARMER, J. D., SHAW, R. S. (1980) Geometry from a time Series. *Phys. Rev. Lett.*, 45, p. 712-716.
- RODRIGUEZ-ITURBE, I., POWER, B. F., SHARIFI, M. B., GEORGAKAKOS, K. P. (1989) Chaos in rainfall. *Wat. Resour. Res.*, 25(7), p. 1667-1675.
- SUGIHARA, G., MAY, R. M. (1990) Nonlinear forecasting as a way of distinguishing chaos from measurement error in time series. *Nature*, 344, p. 734-741.
- TSONIS, A. A. und ELSNER, J. B. (1988) The weather attractor over very short time scales. *Nature*, 333, p. 545-547.
- WOLF, A., SWIFT, J. B., SWINNEY, H. L., VASTANO, J. A. (1985) Determining Lyapunov

exponents from a time series. *Physica*, 16D, p. 285-317.

## **Chaos and Prediction of Hydrological Time Series**

### **ABSTRACT**

*This paper reviews briefly the main concepts associated with the analytical identification of chaotic behaviour in non-linear dynamic systems. The approach is directed to dynamic systems, which do not have a mathematical formulation, but only the time series as representative of the system evolution, which is the case of natural phenomena of interest in water resources engineering. In the context of the chaotic behaviour identification, we consider determination of Hausdorff dimension, estimation of the Lyapunov exponents and estimation of the Kolmogorov entropy. The application of these procedures on a synthetic time series of precipitation indexes, estimated for a basin in the semi-arid region N-E Brazil, is presented and commented in relation to its efficiency in investigating the presence of a low-dimension strange attractor.*

*The relationship is explained between the presence of a strange attractor and the possibility of forecasting future states of the system for which only the time series is available. The suitability of a forecast model based on local approximation using the nearest neighbours in the phase space is analysed for a synthetic time series of precipitation indexes. Aspects related to the error function of the forecast and the structure of the correlation between predicted and real precipitation for different lead times are presented. The predictions are made for 1 and 2 month steps. The possibility is also discussed of using similar models for identifying a chaotic dynamic in hydrological time series.*