

## Previsão de Cheias em Tempo Atual com Sistema Especialista Difuso

Olavo Correa Pedrollo e Antônio Eduardo Lanna

Instituto de Pesquisas Hidráulicas da Universidade Federal do Rio Grande do Sul - Caixa Postal 15029  
91501-970 Porto Alegre, RS - pedrollo@if.ufrgs.br e lanna@iph.ufrgs.br

Recebido: 27/06/01 - revisão: 07/03/02 - aceito: 05/11/02

### RESUMO

*Previsões de níveis de inundação são sempre aproximações. Intervalos de confiança resultam maiores à medida que aumentam as incertezas, podendo ser problemáticos em situações práticas, com níveis fluviais extremos. Como alternativa, propõe-se metodologias para as previsões de níveis de cheia, tanto na forma numérica tradicional quanto na forma de categorias. Utiliza-se um sistema especialista baseado em regras e inferências difusas. Procedimentos computacionais para aprendizado, simulação e consulta são idealizados, e então desenvolvidos para a pesquisa e operação. As metodologias desenvolvidas são aplicadas para previsão na bacia do rio Camaquã (15543 km<sup>2</sup>), para alcances de 1 e 2 dias. Os efeitos da variação da densidade da rede são também analisados, verificando-se que sistemas de previsão hidrológica em tempo atual são possíveis, mesmo com pequeno número de postos de aquisição de dados de chuva, para previsões sob forma de categorias difusas.*

**Palavras-chave:** Previsões de curto prazo; sistemas especialistas difusos; previsões simbólicas.

### INTRODUÇÃO

Previsões de curto prazo são baseadas no princípio de que é possível obter-se, no presente (dito “tempo atual”), informações suficientes relacionadas às causas que condicionam um evento em um futuro próximo, para, com uso de modelos que relacionam as variáveis envolvidas, prever, com precisão adequada, o estado de determinada variável que representa este evento.

Estas previsões, também chamadas de previsões em tempo atual, quando aplicadas para prever níveis fluviais, dependem do conhecimento do estado do sistema e dos volumes esperados (observados a montante do rio, ou precedentes das precipitações recentes). O estado atual é representado pelos níveis observados e pelas condições de umidade dos solos da bacia.

Na prática, porém, são possíveis apenas previsões aproximadas, uma vez que, além dos erros das observações e das inexatidões das hipóteses, formuladas com base em simplificações, existe sempre insuficiência de informações.

A forma mais comum de apresentação das previsões é a numérica ou pontual, embora se saiba que existe uma imprecisão que pode ser interpretada como erro de previsão. Na prática, este erro pode conduzir ao descrédito e conseqüente perda da utilidade da previsão.

Previsões por intervalos de confiança, por sua vez, tornam-se mais úteis à medida que a confiabilidade seja mais importante que a precisão, podendo, porém, serem problemáticas em situações extremas, uma vez que, nestes casos, tendem a extrapolar os limites físicos reais dos níveis fluviais.

O objetivo principal deste trabalho é a pesquisa do uso de sistemas especialistas para a previsão, sob forma de cate-

gorias difusas, de níveis de enchentes fluviais. Espera-se, com este sistema, superar as deficiências apontadas nos limites de confiança, experimentando uma forma inovadora de modelagem para previsão, confiável e imprecisa, que é a previsão por categorias de intensidade.

No que diz respeito às previsões quantitativas, considera-se que a obtenção de desempenho semelhante aos obtidos por modelos tradicionais já seria um indício favorável para uma técnica emergente, ainda por desenvolver plenamente, embora estas previsões não sejam propriamente o alvo principal da pesquisa.

### OS SISTEMAS ESPECIALISTAS DIFUSOS

Sistemas especialistas tradicionais resultam do mapeamento do espaço do domínio no espaço das soluções, constituídos, respectivamente, por condições ou estímulos, e ações ou respostas.

O conhecimento é formado pelas ligações estabelecidas entre os intervalos ou categorias do espaço do domínio e os correspondentes do espaço das soluções, sendo estas ligações representadas por regras associativas do tipo “Se... então...”.

O sistema obtido pode ser capaz de aproximar qualquer função implícita com qualquer grau de precisão desejado, dependendo apenas do detalhamento exigido para a categorização (Kasabov, 1996), o qual pode ser excessivo para aplicações práticas.

Utilizando-se, por sua vez, conjuntos difusos (Zadeh, 1965) parcialmente superpostos quando adjacentes, para representarem os intervalos quantizados, resulta um siste-

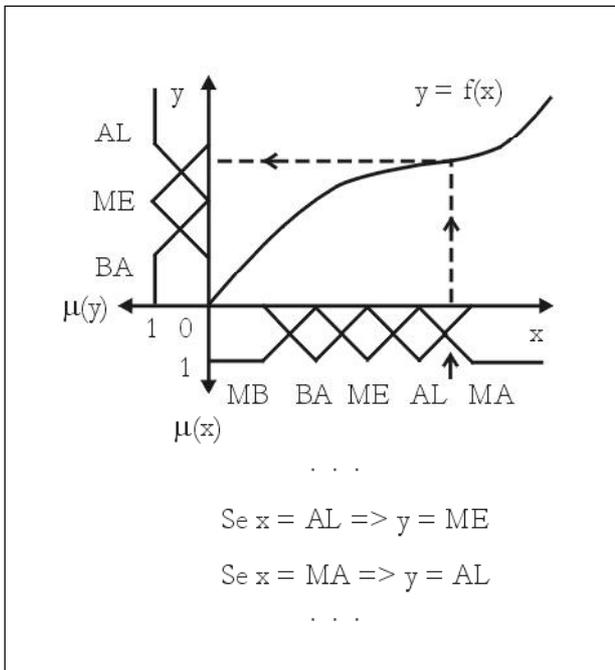


Figura 1. Aproximação de funções.

ma especialista difuso, governado por proposições lógicas representadas por regras associativas.

Kosko (1992) demonstra que um pequeno número de regras, com escolha cuidadosa dos conjuntos difusos para representação do domínio de cada variável, podem aproximar suficientemente qualquer função (Figura 1). Por esta razão, sistemas especialistas difusos são considerados aproximadores universais de funções.

Seja um universo de discurso X, compreendendo o domínio de uma variável x. Pode-se definir um conjunto difuso J como sendo um intervalo de X ao qual é associada uma função de pertinência  $\mu_J(x)$ , que representa o grau com que x pertence a J, e cujos valores podem variar de 0 a 1. No exemplo da Figura 2,  $\mu_J(x)$  representa o grau com que x pertence ao conjunto das pessoas jovens.

Ross (1995) define o núcleo, o suporte e os limites das funções de pertinência, para conjuntos normais. O núcleo de um conjunto difuso caracteriza-se pela ocorrência de pertinência completa ou grau 1, constituindo-se no protótipo, quando restrito a um único ponto. Os elementos descritos são ilustrados na Figura 3, juntamente com as formas mais comuns de funções de pertinência.

Quanto maior o grau de atendimento dos antecedentes de uma regra (ou seja, os graus de pertinência aos conjuntos envolvidos), maior a influência da declaração consequente desta regra sobre a resposta final.

Caso haja mais de uma variável explicativa cujos conjuntos são relacionados por operações conjuntivas ( $\cap$ ) ou disjuntivas ( $\cup$ ), o grau de acionamento é consequência da operação correspondente ( $\cap$ : min;  $\cup$ : max), porém entre diferentes domínios (Figura 4).

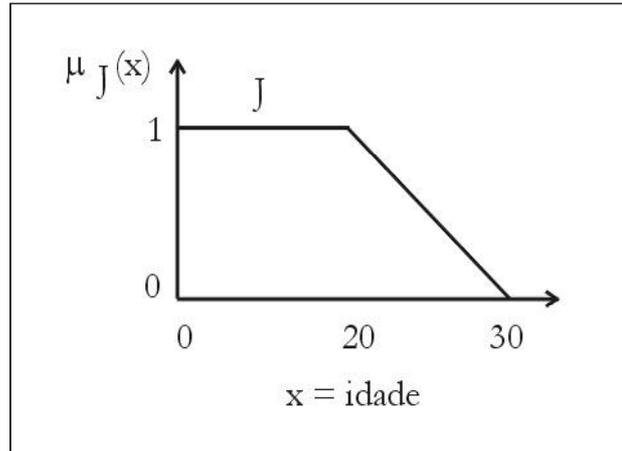


Figura 2. Exemplo de conjunto difuso: Pessoas jovens.

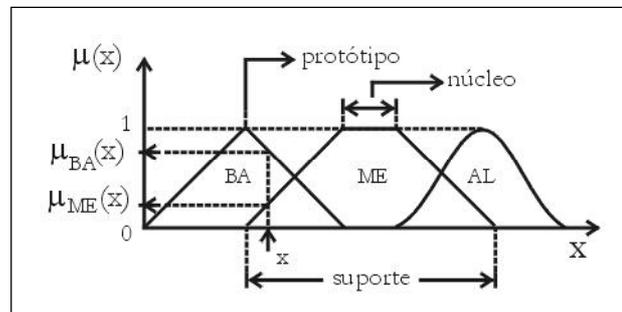


Figura 3. Características dos conjuntos difusos.

É bastante comum que diversos conjuntos consequentes sejam referenciados, uma vez que o sistema funciona em paralelo, com diversas regras sendo acionadas simultaneamente, sendo a função de pertinência resultado de uma composição.

Embora a regra de composição mais utilizada seja a adoção da envoltória superior das funções de pertinência resultantes, Kosko (1992), propõe o uso da composição aditiva, a qual incorpora a frequência de ocorrência de cada categoria (Figura 5), evitando o achatamento, comum na alternativa tradicional.

Alguns métodos para a desquantização, para apresentação de respostas numéricas do sistema, são avaliados por Driankov et al. (1996), destacando-se o método do centro de área e o método das alturas, o qual apresenta como vantagem a menor complexidade computacional (Figura 6).

Com o objetivo de estabelecer-se ponderação diferenciada para as regras, diversos autores, entre eles Kosko (1992) e Kasabov (1996), recomendam a adoção de fatores de confiança para estas, sob a forma de pesos ( $\omega_i$ ), com valores entre 0 e 1. Na prática, é possível que estes pesos alterem os resultados, na medida em que afetam as razões de interpolação entre os protótipos dos conjuntos adjacentes.

Os mecanismos de aprendizado dos sistemas a partir de dados amostrais são classificados (Driankov et al., 1996) em Organizativos (estabelecimento das regras associativas)

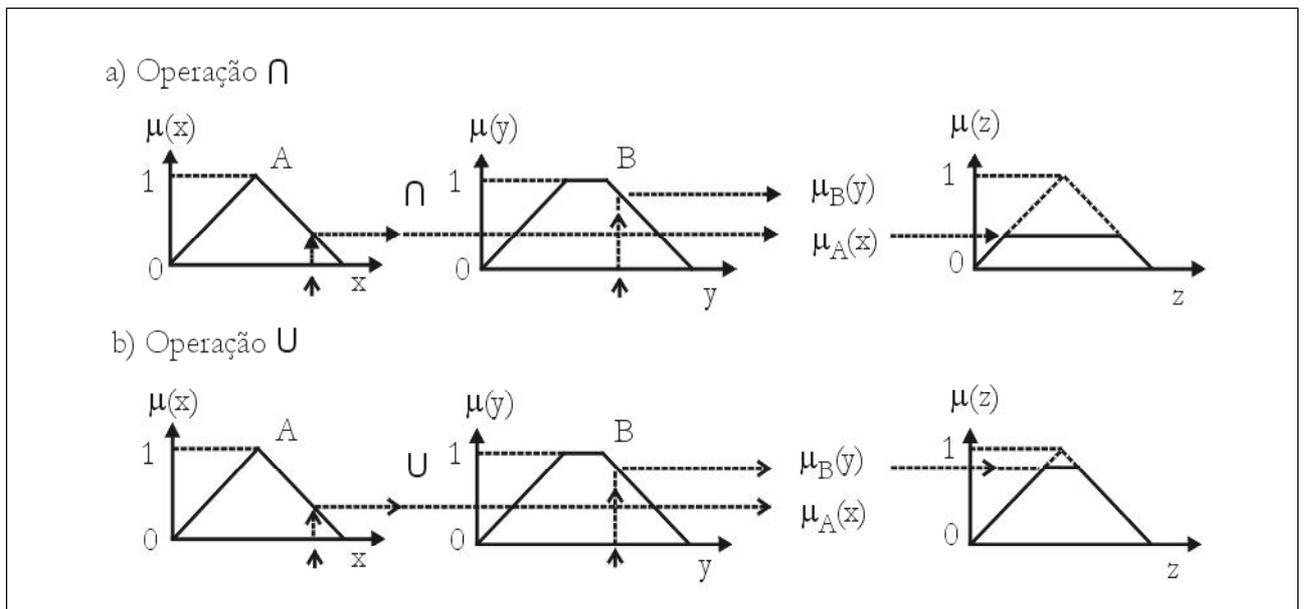


Figura 4. Operações conjuntiva e disjuntiva.

e Adaptativos (ajuste dos pesos atribuídos às regras, ou das funções de pertinência aos conjuntos difusos).

O aprendizado Organizativo é utilizado quando se dispõe de amostragens, simbólicas, numéricas ou mistas, das variáveis que se sabe estarem relacionadas, mas desconhece-se estas relações, seja parcial ou totalmente. Então, técnicas de agrupamento, após a categorização difusa do universo de discurso de cada variável, são utilizadas para produzir a descrição das relações existentes, sob forma de regras associativas.

O aprendizado Adaptativo é necessário quando o comportamento relativo entre as variáveis relacionadas, o qual pode ser expresso sob forma de regras associativas, é conhecido, porém apenas nos seus aspectos qualitativos.

A aplicação prática deste conhecimento depende de encontrar-se o número e a posição dos conjuntos difusos, representados por ordens de intensidade no universo de discurso de cada variável, bem como a forma respectiva de cada função de pertinência.

Kosko (1992) representa o conjunto de regras associativas sob forma de uma matriz multidimensional, a FAM (Fuzzy Associative Memory), onde cada dimensão corresponde a uma variável explicativa, devidamente quantizada, e os elementos são os conjuntos conseqüentes na variável explicada. No exemplo da Figura 7, as variáveis são quantizadas em categorias de intensidade, segundo os símbolos MB (muito baixo), BA (baixo), ME (médio), AL (alto) e MA (muito alto).

## METODOLOGIA

Foram constatados dois fenômenos extremos, particularmente importantes, que ocorrem durante o aprendiza-

do, os quais condicionaram alguns aspectos da aplicação e exigiram aperfeiçoamentos nas técnicas utilizadas, sendo chamados, respectivamente, de “falta de conhecimento” e “falso conhecimento”.

A “falta de conhecimento” manifesta-se quando, tendo havido insuficiência de aprendizado, ocorrem situações para as quais não existe conhecimento sob forma de regras.

A solução idealizada neste trabalho, para prevenir “falta de conhecimento”, é o recurso das bibliotecas combinadas de conhecimento, com conjuntos de regras totais (todas as variáveis contribuem com informações), e conjuntos de regras parciais, onde parte das componentes não participa da regra associativa.

Os conjuntos de regras são estabelecidos de forma que exista sempre, na biblioteca de conhecimentos, uma regra suficientemente simples para dar uma solução, mesmo que esta seja menos precisa.

Este recurso é equivalente à utilização de sistemas com diversos modelos, para prevenção da ocorrência de degradação progressiva de acesso a informações, o que costuma acontecer justamente em situações de emergência.

O “falso conhecimento”, por sua vez, ocorre quando o aprendizado conduz a conjuntos de regras que descrevem perfeitamente as situações submetidas, porém não representam os fenômenos ou as relações subjacentes, o que prejudica a generalização do uso.

Algumas regras geradas, nestes casos, descrevem particularidades de causa desconhecida, que podem ser de natureza aleatória ou mesmo devido a erros dos dados utilizados, produzindo resultados errados quando o modelo é aplicado para previsão ou simulação.

Quanto maior a complexidade do modelo, maior a frequência de regras aprendidas a partir de poucas observa-

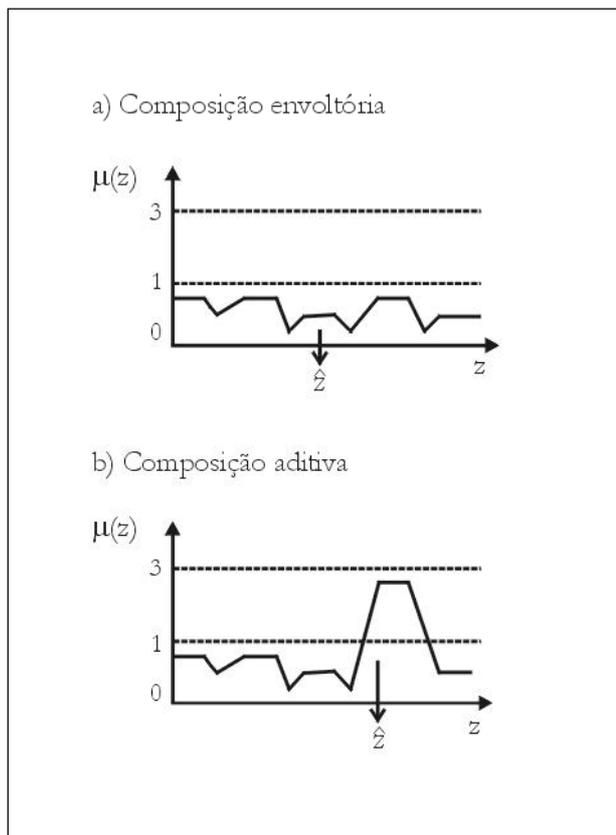


Figura 5. Técnicas de composição.

ções, ou mesmo observação única, devido a limitações amostrais. Nestes casos, pode acontecer que, embora o desempenho de um modelo mais complexo seja melhor do que o obtido com um modelo mais simples quando aplicados às amostras de aprendizado, o mesmo não ocorre quando ambos são aplicados a outras situações.

Uma vez que a ocorrência de “falso conhecimento”, que afeta a significância das regras, é de difícil prevenção, as comparações entre os resultados dos modelos investigados, ou mesmo a seleção das componentes com as quais o modelo apresenta melhor desempenho, são executadas, nesta pesquisa, pela análise do comportamento destes quando utilizados para simulação com séries de dados diferentes das que foram empregadas no aprendizado.

Os modelos empíricos (Pedrollo e Lanna, 1991) são funções contínuas, capazes de fornecer respostas em todo o domínio das variáveis envolvidas, sendo favorecidos pelos pequenos alcances, nos quais as aproximações realizadas (como a linearidade) podem ser aceitas. Para efeito de comparação em igualdade de condições, estes modelos são ajustados para as mesmas informações utilizadas no modelo correspondente resultante da modelagem com sistema especialista.

Os conjuntos difusos utilizados são tipos previamente definidos, caracterizados pela forma da função de pertinência, por proporções fixas, por um ponto de referên-

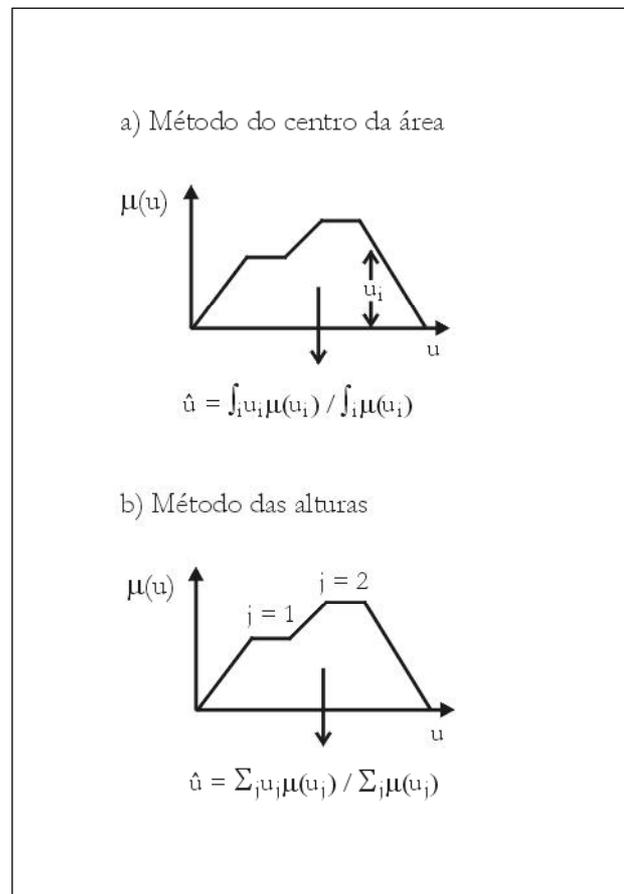


Figura 6. Principais métodos de desquantização.

cia (o protótipo), e por conceitos associados, constituídos por: EB (Extremamente Baixo), MB (Muito Baixo), BA (Baixo), BM (Baixo Médio), ME (Médio), AM (Alto Médio), AL (Alto), MA (Muito Alto) e EA (Extremamente Alto).

Tendo-se em conta as considerações apresentadas na revisão, optou-se pela composição aditiva e pela desquantização pelo método das alturas. Os pesos resultam das frequências ponderadas das ocorrências, sendo explorado, também, o ajuste heurístico destes.

Uma vez que não tem sentido, em situações práticas, a execução de previsões numéricas e categorizadas não coincidentes, e que as previsões numéricas são, por enquanto, muito importantes, optou-se pelo critério de adoção, para a categoria prevista, do conjunto difuso ao qual pertence, com maior grau de pertinência, o valor numérico resultante da desquantização.

Esta categoria constitui o conjunto principal (CP), sendo que o conjunto com o segundo maior grau de pertinência é o conjunto secundário (CS). Este critério é consistente, uma vez que, se o valor observado realiza-se nas proximidades do valor previsto, é razoável supor-se que ambos ocorreram na mesma categoria.

Os modelos tradicionais, os quais apresentam, via de regra, resultados numéricos, podem ter estes resultados con-

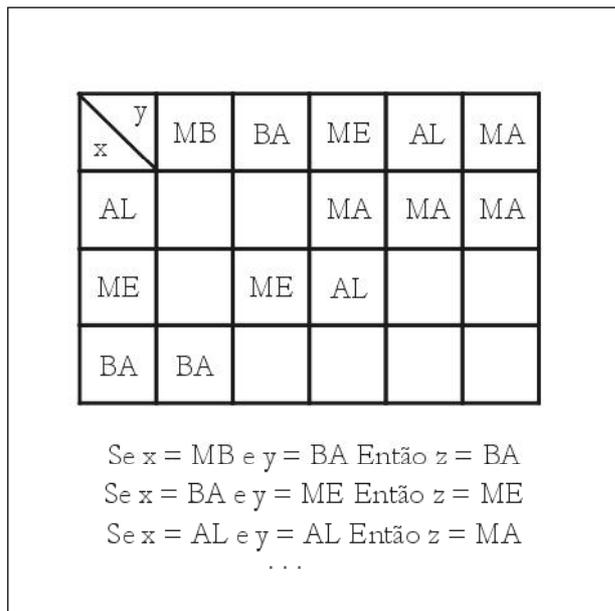


Figura 7. Representação da FAM.

vertidos à forma de categorias com graus de pertinência associados, o que possibilita comparações de resultados também sob forma de categorias.

As estatísticas utilizadas para avaliação e comparação dos desempenhos das previsões sob forma numérica são o Erro Padrão de previsão (Ep), o Coeficiente de Eficiência (Ce) e o erro máximo absoluto (Ex):

$$Ep = [1/k \sum_t (\hat{u}_t - u_t)^2]^{1/2} \quad (1)$$

$$Ce = 1 - \sum_t (\hat{u}_t - u_t)^2 / \sum_t (u_{\mu} - u_t)^2 \quad (2)$$

$$Ex = \text{MAX} (|\hat{u}_t - u_t|) \quad (3)$$

onde  $u_t$  é a observação, no intervalo  $t$ , da variável  $u$ ;  $u_{\mu}$  a média dos valores amostrais de  $u$ ;  $\hat{u}_t$  o valor previsto de  $u$  para o intervalo de tempo  $t$ ;  $k$  o número de observações.

O primeiro parâmetro quantifica a dispersão das previsões em torno das observações, enquanto que o segundo, introduzido por Nash e Sutcliffe (1970) apud Lanna e Indursky (1983), representa a proporção com que o modelo explica a variância da variável observada.

A comparação da qualidade dos resultados de modelos de previsão competitivos é feita com o Coeficiente de Eficiência Relativa (Lanna e Indursky, 1983):

$$Er = 1 - \sum_t (\hat{u}_t - u_t)^2 / \sum_t (\hat{u}_a - u_t)^2 \quad (4)$$

o qual calcula a fração adicional da variância das observações que é explicada pelo modelo testado ( $u_t$ ), além da variância explicada por um modelo alternativo ( $\hat{u}_a$ ).

Neste trabalho utiliza-se, também, uma forma ligeiramente modificada deste coeficiente, a qual é denominada

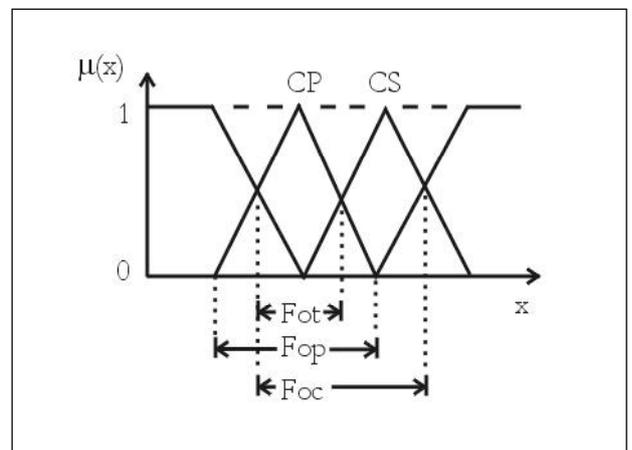


Figura 8. Domínio dos coeficientes de agrupamento.

Coeficiente de Perda de Eficiência (Pe):

$$Pe = \sum_t (\hat{u}_t - u_t) / \sum_t (\hat{u}_a - u_t) - 1 \quad (5)$$

para quantificar o quanto os resultados ficam prejudicados com a degradação do sistema, ou com simplificações posteriores dos modelos.

No caso de previsões por categorias difusas, porém, necessita-se de uma nova classe de estatísticas de desempenho, as quais possam ser utilizadas como estimativas para a quantificação da capacidade do modelo, quando utilizado para esta modalidade de previsões.

Os coeficientes de agrupamento, concebidos para atender esta necessidade, são calculados como freqüências, expressas em forma de porcentagens das ocorrências das observações nas respectivas faixas de domínio de cada coeficiente, em relação aos conjuntos principais (CP) e secundários (CS), ilustradas na Figura 8.

A freqüência de ocorrências totais (Fot) e a freqüência de ocorrências parciais (Fop) representam o quanto o modelo acertaria se fosse utilizado para previsão da ocorrência do valor observado no conjunto principal (CP), ou seja, o conjunto resposta que apresenta maior grau de pertinência.

No primeiro caso, o critério para considerar-se correta a previsão é de que a observação apresente pertinência, ao conjunto previsto (CP), maior do que a qualquer outro conjunto da variável predita. No segundo caso, o critério é menos rigoroso, sendo exigido apenas que o grau de pertinência a este conjunto seja maior do que zero.

A freqüência de ocorrências totais combinadas (Foc), por sua vez, representa a freqüência com que as observações ocorrem, com maior pertinência, ou ao conjunto principal, (CP) ou ao conjunto secundário (CS).

A idéia básica para a implementação do método organizativo é a análise de agrupamentos deduzidos da amostra, com adoção, para cada combinação de antecedentes, do conjunto conseqüente cuja ocorrência é mais freqüente.

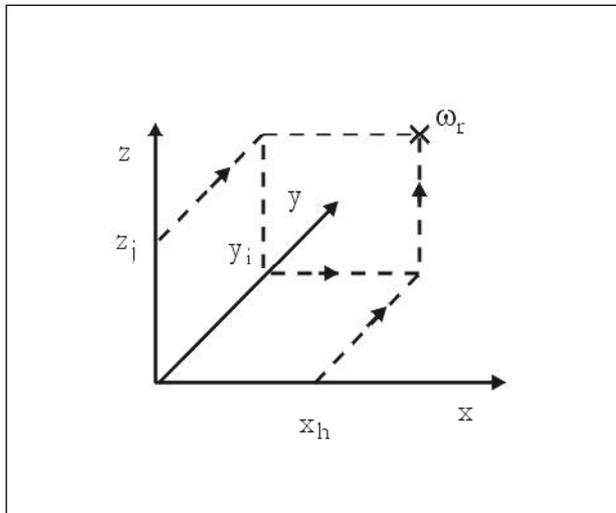


Figura 9. Elemento da matriz cognitiva.

Para tanto, é incorporada uma dimensão adicional à representação matricial das FAM's (Kosko, 1992), para representar os conjuntos da variável explicada, liberando desta finalidade os elementos da matriz, que passa a ser chamada de matriz cognitiva, devido à função exercida, de agregação de conhecimentos.

Cada elemento (Figura 9) representa, no decorrer do processo, a frequência relativa com que cada conjunto da variável explicada é acionado, uma vez que resulta de somatórios de ocorrências, as quais, por sua vez, são previamente ponderadas pelos graus de pertinência aos conjuntos difusos envolvidos, uma vez que cada valor numérico pode pertencer a mais de uma categoria.

Esta forma de representação do conhecimento é mais completa do que a FAM, pois incorpora a informação da importância, atribuída ou calculada, da conexão. Com a aquisição de novas informações, os elementos são alterados pela reconsideração destas frequências, resultando modificação ou incorporação de conhecimentos.

Após o processamento de toda a amostra de aprendizado, um procedimento específico seleciona, para cada combinação ocorrida das variáveis explicativas, o conjunto mais frequente na variável explicada, compondo uma regra associativa. O fator de ponderação é também uma medida da confiabilidade da regra pesquisada.

Os processos de aprendizado Adaptativos implicam em pesquisas orientadas, ou na forma e posição (f-adaptativo) dos conjuntos difusos, ou nos pesos atribuídos a cada regra associativa (p-adaptativo), gerada ou existente, sendo aplicados os dois procedimentos, sucessivamente.

As investigações são orientadas no sentido de obtenção do mínimo para a função objetivo, para a qual adotou-se o somatório dos desvios quadráticos entre previsões e observações.

A melhoria da distribuição, no universo de discurso, dos conjuntos difusos das variáveis a partir de uma configu-

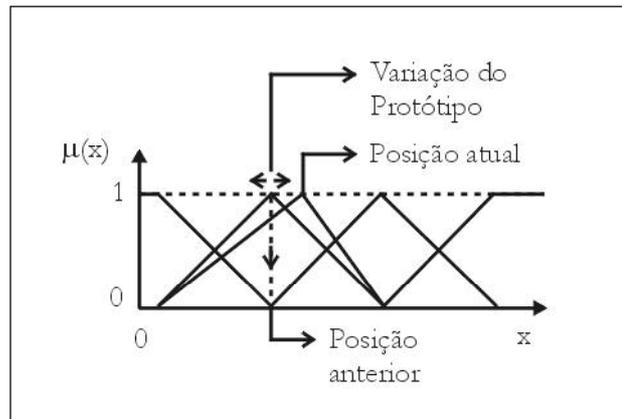


Figura 10. Variação dos protótipos.

ração inicial, é um processo muito lento, uma vez que depende de pesquisa, por tentativas, da posição de cada protótipo, de cada variável (Figura 10), em iterações sucessivas, escolhendo-se, a cada pesquisa, o menor valor verificado para a função objetivo.

A investigação das funções de pertinência mais convenientes, por sua vez, não é um procedimento de otimização por aproximação, mas uma pesquisa exaustiva, por meio de experimentação, da melhor configuração.

Finalmente, no processo de adaptação dos pesos, determina-se o valor de cada parâmetro, associado a cada regra da biblioteca de conhecimentos. O procedimento consiste na busca de uma ponderação mais eficiente para cada regra, por simulação com o cálculo da função objetivo.

Neste caso, os pesos funcionam como parâmetros os quais podem alterar a qualidade dos resultados, perdendo, portanto, o caráter de medida da confiabilidade ou importância da regra. Evidentemente, o processo não verifica todas as combinações possíveis, mas consegue-se melhoras, em iterações sucessivas.

Os procedimentos adaptativos produzem alterações nos conjuntos difusos, de forma que cada modelo, representado pela respectiva biblioteca de conhecimentos, passa a possuir uma definição própria para estes.

Como conseqüência, uma previsão em determinada categoria possui significados diferentes em função do alcance a que se destina, podendo gerar confusões ao ser divulgada.

Uma solução adequada para este inconveniente é a definição, independentemente das quantizações particulares de cada modelo, de uma escala fixa de categorias da variável explicada. O valor previsto é submetido a esta escala, a qual transforma a previsão, da forma numérica para a forma de conjuntos difusos, principal e, se houver, secundário, com os respectivos graus de pertinência (Figura 11).

As categorias utilizadas podem estar associadas à noção existente de níveis ALTOS, MÉDIOS, BAIXOS, etc.. Porém, uma alternativa interessante é supor-se que existe uma frequência de acertos que deve ser atingida, para a

qual pesquisa-se a categorização correspondente, por simulação iterativa.

As freqüências de acerto, sob forma dos coeficientes de agrupamento, são calculadas ao final da simulação. Adota-se, então, a quantização difusa mais detalhada possível, que atende às exigências formuladas.

Com este mecanismo, são possíveis previsões mais confiáveis, por categorias, sem que sejam alteradas as precisões das estimativas numéricas originais.

Outro ponto relevante é que, uma vez que previsões numéricas de qualquer origem podem ser convertidas para a forma categorizada, pode-se adotar uma categorização específica de referência.

## APLICAÇÃO E RESULTADOS

### Considerações gerais

A bacia hidrográfica do rio Camaquã (Figura 12) tem forma alongada, estando compreendida entre as latitudes  $30^{\circ}30'$  e  $31^{\circ}15'S$  e longitude de  $51^{\circ}45'$  e  $54^{\circ}30'WGR$ , nas regiões geomorfológicas do Escudo Cristalino e da Planície Costeira.

Entre os postos hidrológicos existentes, destaca-se o de Passo do Mendonça, com registrador para precipitações e medição de descargas (Pedrollo, 1997), que controla uma área de  $15543 \text{ km}^2$ , ou seja, a maior parte da bacia. Os limites de variação histórica situam-se entre 0 e  $800 \text{ cm}$ , e entre 0 e  $5000 \text{ m}^3 \text{ s}^{-1}$ , para níveis e descargas, respectivamente.

Dois períodos, um para ajuste amostral e outro para verificação, são estabelecidos. O primeiro compreende o intervalo de 01 de janeiro de 1991 a 01 de janeiro de 1997, sendo propositadamente maior, devido às dificuldades existentes para aquisição suficiente de conhecimentos a partir de amostras disponíveis. O segundo período compreende o intervalo de 24 de abril de 1997 a 31 de dezembro de 1999.

São utilizadas onze estações pluviométricas, com base nos critérios de existência e disponibilidade dos dados em todo o período pesquisado, e na abrangência da distribuição espacial. As chuvas representativas utilizadas são as médias aritméticas dos totais acumulados diariamente nos postos.

Uma vez que as observações de totais pluviométricos acumulados diariamente são feitas às 07:00 horas, e que o máximo proveito das previsões é obtido quando o tempo atual coincide com a obtenção da última informação necessária, adota-se esta hora para a execução das previsões, para todos os alcances considerados.

Três alcances de previsão são apresentados, ou seja, 1 dia,  $1 \frac{1}{2}$  dias e 2 dias. Alcances maiores não são, por enquanto, cogitados, devido ao decaimento progressivo da qualidade dos resultados, e à utilidade duvidosa.

As previsões são efetuadas, portanto, para as 07:00 horas ou 17:00 horas do dia seguinte, ou ainda para as 07:00

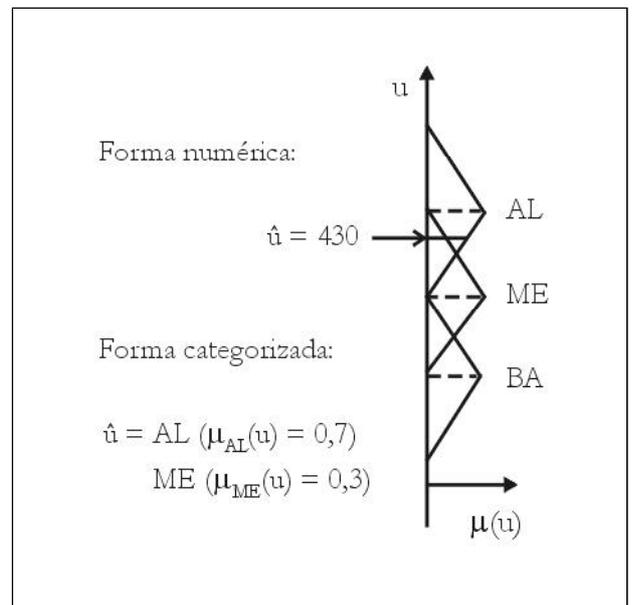


Figura 11. Transformação para a forma categorizada.

horas do próximo dia, conforme o alcance seja, respectivamente, para 1,  $1 \frac{1}{2}$  ou 2 dias. Portanto,  $1 \frac{1}{2}$  dia, significa, na verdade, 34 horas de alcance.

### A identificação dos modelos

A identificação dos modelos consiste na investigação dos grupos de variáveis, bem como do número de conjuntos difusos adotados para estas, que proporcionam respostas mais eficientes à previsão.

Para efeito destas aplicações, emprega-se o termo variável de forma particularizada. No caso de um modelo autorregressivo com duas componentes, por exemplo, cada componente é interpretada como sendo uma variável, podendo ter atribuída quantização difusa diferente.

A ordem de investigação é importante, uma vez que cada modelo pode complementar suas regras com as do modelo anterior, mais simples, se necessário, para suprir falta de conhecimento.

À medida que aumenta a complexidade do modelo e, portanto, o número possível de combinações de antecedentes, aumenta também o número de regras geradas. Estas podem ser de duas naturezas: As que descrevem as relações lógicas existentes entre as variáveis (grupo 1), e as que descrevem as combinações que são ocorrências específicas das séries utilizadas para o aprendizado (grupo 2).

Quando uma componente adicional não é importante para a construção de regras do grupo 1, esta contribui para a construção de regras grupo 2, gerando “falso conhecimento”, o qual contribui para erros quando o modelo é aplicado para outras séries. Por isso, o excesso de componentes não



Figura 12. Localização da bacia do rio Camaquã.

apenas deixa de trazer benefícios, mas também pode prejudicar os resultados obtidos na verificação.

Portanto, as pesquisas são interpretadas sempre com estatísticas resultantes da verificação, com séries diferentes das utilizadas no aprendizado.

A função objetivo utilizada para as comparações, com a finalidade de identificação dos modelos mais eficientes, é o Erro Padrão (Ep), expresso em cm, o qual transmite a noção da ordem de grandeza dos erros.

Os conjuntos difusos são, inicialmente, arbitrados, tendo os protótipos extremos sido escolhidos após inspeção visual das séries de dados amostrais.

Os níveis são representados por nove conjuntos difusos, com valores dos protótipos representados pelo conjunto (0, 100, 200, 300, 400, 500, 600, 700, 800), e funções de pertinência triangulares.

As chuvas, cujos valores correspondem às médias dos totais acumulados diariamente em cada posto, são representadas por três conjuntos difusos com protótipos correspondentes de 0, 30 e 60 mm, e funções de pertinência sigmoidais.

São testadas diversas alternativas para o número de componentes auto-regressivas, porém, para todos os alcances, observa-se que os melhores resultados correspondem ao uso de apenas uma componente auto-regressiva, o que significa que a captação de tendências de crescimento ou recessão induz mais a erros do que a acertos.

Apresenta-se resumo, na Tabela 1, dos erros padrões correspondentes aos resultados mais favoráveis, para modelos auto-regressivos e com componentes de chuvas. Nesta tabela, “p” representa o número de componentes auto-

regressivas, e “q” o número de componentes de chuvas, sendo a primeira componente o total diário acumulado no tempo atual, e as outras, os totais diários acumulados nos dias anteriores, seqüencialmente em direção ao passado, e sempre com valores totalizados às 07:00 horas.

A eficiência relativa,  $Er_{i/j-1}$  (%), dos modelos com três componentes de chuvas, em relação aos modelos auto-regressivos, ou seja, o quanto de variância estes explicam a mais do que os anteriores, resulta em 41,58%, 45,41% e 47,98%, para os alcances respectivos de 1, 1 ½ e 2 dias.

A variação do número de categorias de chuvas é efetuada com estes modelos, para a investigação da quantização mais adequada, apresentando-se, na Tabela 2, as comparações, com uso do erro padrão.

São experimentados dois, três e cinco conjuntos difusos, com protótipos adotados de 0 e 60, para o primeiro caso, 0, 30 e 60 para o segundo, e de 0, 15, 30, 45 e 60 para o terceiro.

O número de regras associativas geradas não apresenta variação em função do alcance. Verificou-se, porém, dependência em relação ao número de categorias, com crescimento do número de regras difusas em proporção maior do que o refinamento da quantização, resultando 41, 96 e 205 regras para, respectivamente, 2, 3 e 5 categorias difusas.

O número ideal de categorias não é necessariamente o máximo possível, conforme observado. Isto se deve, em parte, a faltas de conhecimento ocorridas, com uso de cinco categorias (1, 3 e 2 falhas para, respectivamente, 1, 1 ½ e 2 dias de alcance), as quais são, automaticamente, supridas com conhecimento proveniente de modelos

Tabela 1. Identificação dos modelos: EP (cm).

(p, q)	Alcance (dias)		
	1	1 ½	2
(1, 0)	37,89	48,44	85,22
(1, 1)	31,08	39,11	53,74
(1, 2)	29,53	37,79	59,70
(1, 3)	<b>28,96</b>	<b>35,79</b>	<b>45,98</b>
(1, 4)	30,43	36,43	46,09

Tabela 2. Categorias de chuva: Ep (cm).

Número de categorias	Alcance (dias)		
	1	1 ½	2
2	30,73	36,96	50,46
3	<b>28,96</b>	<b>35,79</b>	<b>45,98</b>
5	31,74	39,45	50,22

mais simples (com menor número de variáveis), cujas regras são previamente geradas e incorporadas às regras associativas do modelo principal.

Os protótipos adotados para a análise do número de categorias de níveis (Tabela 3), foram os conjuntos (0, 100, 200, 350, 500, 650, 800), e (0, 200, 400, 600, 800) para, respectivamente, sete e cinco categorias difusas.

Observa-se redução na qualidade dos resultados, e também redução no número de regras associativas geradas, sendo de 96, 78 e 66 regras para, respectivamente, 9, 7 e 5 categorias difusas de níveis, utilizando-se o número de três categorias para as chuvas.

### As técnicas adaptativas de aprendizado

Até então, a geração de regras associativas é feita pelos procedimentos organizativos. Verifica-se, a seguir, o quanto os procedimentos adaptativos podem contribuir, seja pelas modificações das formas dos conjuntos difusos (f-adaptativos), seja pelas alterações posteriores dos pesos atribuídos a cada regra associativa (p-adaptativos).

Na Tabela 4, são apresentados os erros padrões resultantes da simulação, com as amostras de ajuste e verificação, após os sucessivos procedimentos de aprendizado.

Os recursos de aprendizado são aplicados seqüencialmente, aproveitando-se, em cada etapa, os progressos obtidos na etapa anterior.

As principais alterações obtidas nos conjuntos difusos, ao final dos procedimentos f-adaptativos, ocorrem nos valores dos protótipos extremos das precipitações, que passam a ser de 65, 80 e 85 mm, para, respectivamente, alcances de 1, 1 ½ e 2 dias. As chuvas  $P_{t-1}$  têm suas funções de pertinência modificadas, do tipo sigmoidal para o tipo triangular.

Tabela 3. Categorias de níveis: Ep (cm).

Número de categorias	Alcance (dias)		
	1	1 ½	2
9	<b>28,96</b>	<b>35,79</b>	<b>45,98</b>
7	34,36	38,57	50,50
5	56,79	59,97	79,36

Tabela 4. Desempenho do aprendizado: Ep (cm).

Procedimento	Aplicação	Alcance (dias)		
		1	1 ½	2
Organizativo	Ajuste	23,4	29,0	38,9
	Verificação	29,0	35,8	46,0
F-Adaptativo	Ajuste	26,1	31,4	42,4
	Verificação	23,2	29,3	39,4
P-Adaptativo	Ajuste	23,6	29,2	40,3
	Verificação	22,0	28,0	39,1

Os progressos mais significativos, porém, ocorrem após a aplicação do procedimento p-adaptativo.

Na Tabela 5, o coeficiente de eficiência relativo,  $Er_{i/i-1}$  (%), é utilizado para apresentar a evolução da variança explicada, ao longo das pesquisas.

Compara-se a eficiência do modelo final, após o aprendizado p-adaptativo, com a eficiência resultante da aplicação do método imediatamente anterior (f-adaptativo), e a deste com a resultante do aprendizado organizativo, observando-se que menores incrementos de qualidade ocorrem com maiores alcances.

### As previsões categorizadas

A pesquisa da quantização que atende critérios estabelecidos de confiabilidade, com possível aumento da amplitude das categorias, pode ser feita sem que isto implique em alteração da precisão das previsões pontuais.

O uso do recurso de conversão de categorias, para operação com menor número de conjuntos, não resulta em degradação da eficiência das previsões pontuais, enquanto que a modelagem direta, com menor número de conjuntos, implica necessariamente em redução desta eficiência.

No caso das previsões de níveis com os modelos pesquisados, tem-se, na Tabela 6, os valores dos coeficientes de agrupamento, para cada alcance e para cada número de categorias possíveis com os recursos disponíveis.

As funções de pertinência utilizadas são triangulares, os protótipos adotados para os níveis são os mesmos usados inicialmente, para nove categorias, e os conjuntos (0, 150, 300, 400, 500, 650, 800), (0, 150, 300, 500, 800), (0, 500, 800) e (0, 800), para sete, cinco, três e duas categorias, respectivamente.

**Tabela 5. Eficiência relativa dos procedimentos.**

Recurso	Alcance (dias)		
	1 Er <sub>i/j-1</sub> (%)	1 ½ Er <sub>i/j-1</sub> (%)	2 Er <sub>i/j-1</sub> (%)
Organizativo (j=1)	-	-	-
F-Adaptativo (j=2)	17,5	22,3	14,1
P-Adaptativo (j=3)	18,3	13,5	9,53

O critério empregado para a investigação da categorização mais desejável pode ser simples ou múltiplo, conforme dependa de um ou mais coeficientes de agrupamento, sendo a frequência de ocorrência total (Fot) o coeficiente mais restritivo.

No presente estudo têm-se, para a quantização em nove conjuntos difusos, coeficientes de ocorrência total combinada favoráveis até mesmo para alcances de dois dias.

Para as previsões com alcance de um dia, os coeficientes obtidos na simulação indicam previsões corretas nas categorias difusas (combinadas), com 98,8% de acerto, com ocorrência de 86,7% na categoria principal prevista.

Pode-se obter, por exemplo, coeficientes de ocorrência total (Fot) de pelo menos 95%, para qualquer um dos alcances pesquisados, com uso de apenas três categorias (níveis *alto*, *médio* ou *baixo*). Neste caso, ainda pode-se contar com acertos de 100% na categoria secundária (Foc = 100%), quando as observações ocorrerem fora da categoria principal.

É importante, neste ponto, a ilustração do significado das previsões em categorias combinadas. Por exemplo (Figura 13), as previsões de níveis AL e MA, com graus de pertinência, respectivamente, de 0,8 e 0,2. Significa que se espera, para o alcance considerado, nível entre alto (AL: em torno de 600 cm), e muito alto (MA: em torno de 700 cm), com tendência maior para alto.

### Comparações com modelos empíricos

Existe uma forte analogia entre os modelos empíricos lineares e os modelos que são representados pelos sistemas especialistas difusos, no que diz respeito à escolha e identificação das variáveis e à capacidade de adaptação às novas situações.

Este tipo de modelo freqüentemente aparece, sob denominações variadas, na literatura especializada em previsões em tempo atual, apresentando, geralmente, melhor desempenho que outras abordagens, especialmente se utilizadas técnicas recursivas de calibragem.

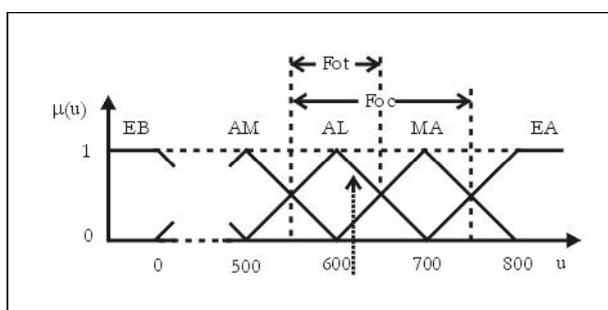
Os modelos empíricos lineares ajustados são representados pela equação:

$$\hat{u}_{t+\tau} = \alpha u_t + \sum_{i=1,3} \beta_i P_{t+i} \quad (6)$$

onde  $\hat{u}$  e  $u$  representam, respectivamente, os níveis previstos e observados, e  $P_t$ ,  $P_{t+1}$  e  $P_{t+2}$ , as precipitações, res-

**Tabela 6. Coeficientes de agrupamento obtidos na verificação das previsões.**

Alcance (dias)	Coeficiente (%)	Número de conjuntos				
		2	3	5	7	9
1	Fot	98,1	97,6	93,2	90	86,7
	Fop	100	100	99,9	99,0	98,7
	Foc	100	100	100	99,0	98,8
1 ½	Fot	96,3	96,5	92,1	87,3	83,0
	Fop	100	100	99,9	97,8	97,3
	Foc	100	100	100	98,2	97,4
2	Fot	94,5	95,3	87,5	82,8	78,3
	Fop	100	100	99,7	96,1	95,6
	Foc	100	100	99,8	96,7	96,0



**Figura 13. Previsões em categorias combinadas.**

pectivamente, no tempo atual (t), e nos dias sucessivamente anteriores.

Na Tabela 7, são apresentados os parâmetros ( $\alpha$  e  $\beta_i, i = 1, 3$ ) obtidos para os modelos empíricos lineares, ao final do ajuste, para os três alcances considerados.

Os resultados dos modelos empíricos lineares são convertidos para previsões em categorias difusas, uma vez que este é um recurso proposto para uso geral, com qualquer tipo de modelo matemático. Desta forma, torna-se possível a comparação do desempenho dos modelos, tanto nos termos tradicionais (previsões pontuais, ou seja, numéricas), quanto nos termos adequados aos sistemas categorizados, em função dos coeficientes de agrupamento.

Para permitir comparações de eficiência das previsões difusas de níveis, é adotada uma escala comum de conjuntos difusos, utilizando-se a quantização mais eficiente obtida na análise anterior, de previsões categorizadas. Os níveis são, portanto, quantizados em nove conjuntos difusos, com funções de pertinência triangulares e protótipos de 0, 100, 200, 300, 400, 500, 600, 700 e 800 cm.

As previsões obtidas com os modelos são analisadas com uso dos coeficientes de eficiência e dos coeficientes de agrupamento, comparando-se, para os diversos alcances, o desempenho dos modelos.

Nas Figuras 14, 15 e 16, são apresentadas as séries de níveis observados no período de verificação, bem como os erros produzidos pelo sistema especialista difuso (SED) e

**Tabela 7. Parâmetros dos modelos empíricos lineares.**

Parâmetros	Alcance (dias)		
	1	1 ½	2
$\alpha$	0,927	0,902	0,876
$\beta_1$	2,371	3,309	4,132
$\beta_2$	0,955	0,977	0,947
$\beta_3$	0,306	0,399	0,453

pelo modelo empírico linear (MEL), para os alcances de 1, 1 ½ e 2 dias.

É notável o incremento da magnitude dos erros com o aumento do alcance, bem como sua redução durante os períodos contínuos de recessão. Também, observa-se que os erros dos diferentes modelos não são coincidentes, sendo os dos sistemas especialistas um pouco menores.

Na Tabela 8, apresenta-se o desempenho comparativo dos modelos construídos com os sistemas especialistas difusos e dos modelos empíricos lineares, através das estatísticas propostas, obtidas pela simulação de previsões para o período de verificação.

Os modelos com base no sistema especialista difuso apresentam, em geral, desempenhos um pouco superiores aos dos modelos empíricos lineares, para previsões pontuais, conforme análise do erro padrão ( $E_p$ ), erro máximo ( $E_x$ ) e coeficiente de eficiência ( $C_e$ ), para todos os alcances.

Porém, um fato que chama a atenção é justamente que os erros máximos são muito grandes, para ambas as técnicas, e ocorrem, provavelmente, nos momentos em que previsões confiáveis são mais necessárias.

A confiabilidade das previsões pontuais pode ficar sob suspeita, pois, ainda que com coeficientes de eficiência que atestam varianças explicadas da ordem de 95%, os erros máximos chegam a valores maiores do que um metro.

As frequências de ocorrência total ( $F_{ot}$ ), nas categorias principais, são, em geral, baixas, para as duas classes de modelos, enquanto que as frequências de ocorrência parcial ( $F_{op}$ ) e frequências de ocorrência combinada ( $F_{oc}$ ) são, em geral, altas.

As frequências de ocorrência parcial e combinada, por sua vez, apresentam valores muito semelhantes entre si. As previsões combinadas em categorias difusas, são muito semelhantes, para alcances menores, quando executadas com o modelo formado unicamente por bibliotecas de conhecimento e pelo modelo empírico linear. Para o alcance de dois dias, porém, os modelos empíricos lineares apresentam desempenho claramente desvantajoso, em relação aos primeiros.

É notável como, para 2 dias de alcance, mesmo com as limitações observadas nas previsões numéricas, o modelo gerado pelo sistema especialista resulta em 96% de acerto nas categorias combinadas, contra 93,3% do modelo empírico linear.

Para exemplificar, os coeficientes de agrupamento indicam que as previsões acertam 78,3% se forem formuladas para categorias simples (categoria principal, com maior grau de pertinência), e 96% dos casos, se forem formuladas difusamente, com duas categorias, principal (CP) e secundária (CS), como: “CP ou CS, com tendência maior para CP”.

Se a previsão com dois dias de antecedência, com a precisão proporcionada pelas categorias adotadas, puder ser útil, provavelmente terá sido viabilizada para este alcance graças ao sistema especialista difuso e às técnicas propostas para uso e interpretação das previsões categorizadas, tendo-se em conta os resultados até então obtidos para as previsões numéricas com modelos empíricos lineares.

### As dimensões da rede de aquisição

A principal alteração nas variáveis de entrada, como consequência da utilização de menor número de postos para sua determinação, é o aumento da amplitude, uma vez que menor número de postos resultam em médias máximas maiores.

Por exemplo, na série de verificação, os máximos totais diários médios das precipitações resultaram em 71, 89 e 92 mm para, respectivamente, médias obtidas de 11, 7 e 4 postos, e máximo de 124 mm para totais diários simples, do posto de Passo do Mendonça.

Em um modelo baseado em equações, a adaptação deste para operação com as novas variáveis de entrada implicaria em modificações dos coeficientes de ajuste, enquanto que, no sistema especialista, as alterações devem ser feitas no espaço de domínio das variáveis que representam as precipitações.

As alterações necessárias, por sua vez, podem ser efetuadas por meio dos recursos desenvolvidos para o aprendizado Adaptativo. Este procedimento é utilizado, portanto, de modo a não interferir com as regras aprendidas, resultando em modificações sensíveis nos conjuntos difusos das variáveis que representam os totais diários precipitados.

Os protótipos dos conjuntos extremos variam, após a execução dos ajustes propostos, de 65, 80 e 85 mm para os alcances de, respectivamente, 1, 1 ½ e 2 dias, para os valores apresentados na Tabela 9.

Observam-se, nitidamente, deslocamentos no sentido da expansão do domínio, tanto na passagem das condições originais (médias de 11 postos) para as reduzidas (médias de 4 e de 7 postos), quanto na passagem destas para a condição de uso de apenas um posto.

Na Tabela 10, são apresentados os coeficientes de perda de eficiência, em relação ao uso das médias de totais pluviométricos diários de 11 postos, dos modelos que empregam médias correspondentes de sete e quatro, bem como dados de um único posto pluviométrico.

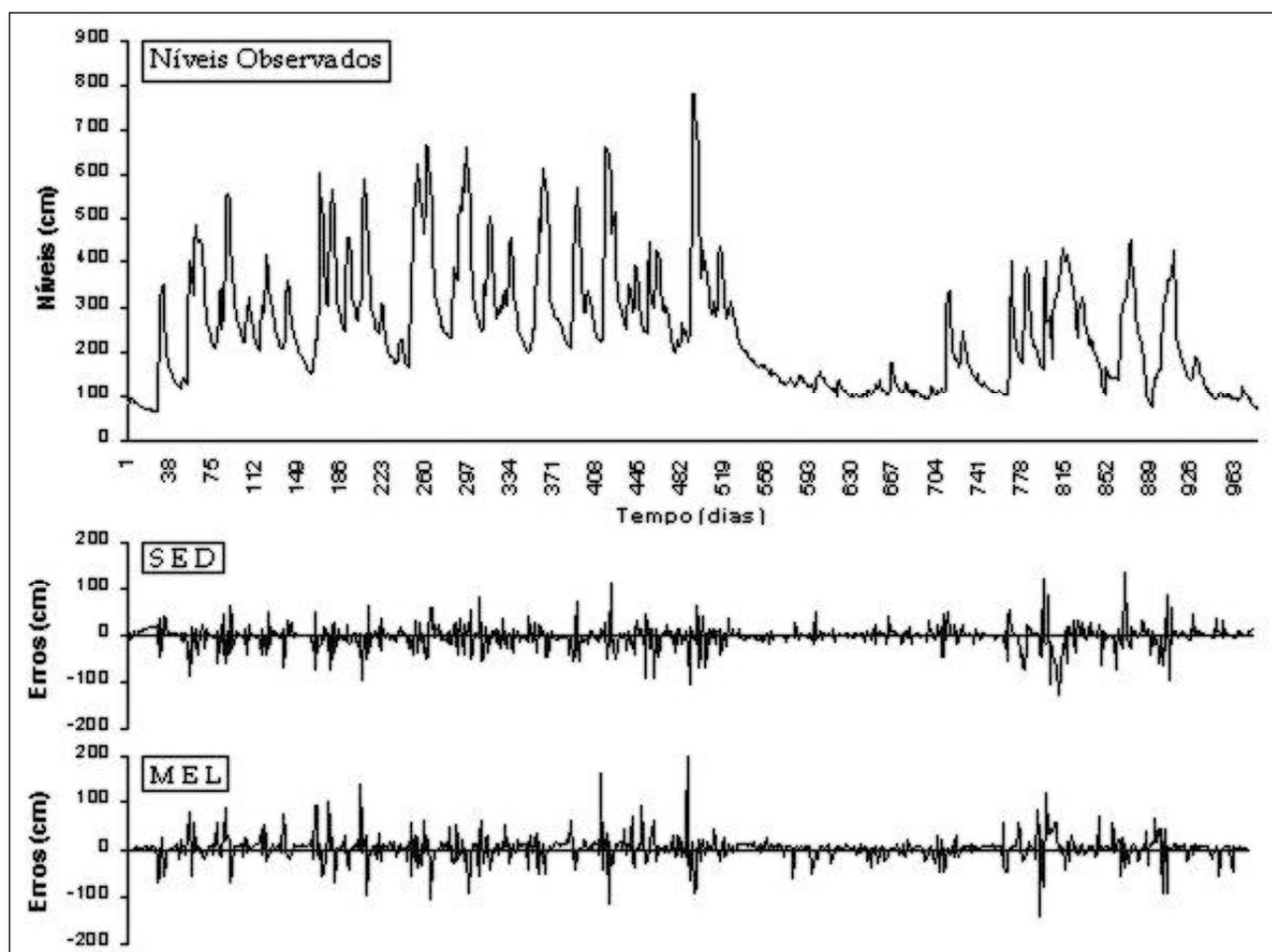


Figura 14. Previsões com Sistema Especialista Difuso (SED) e Modelo Empírico Linear (MEL): Alcance de 1 dia.

São examinadas duas condições de uso para cada alternativa, ou seja:

- a condição original (os protótipos dos conjuntos difusos são rigorosamente os mesmos que são usados com as médias de 11 postos);
- a condição modificada (os protótipos dos conjuntos que representam as categorias de chuvas são ajustados, pelo procedimento Adaptativo).

A visualização do significado das perdas de eficiência observadas é apresentada na Figura 17, onde se têm os erros resultantes da aplicação do sistema especialista difuso com uso de número variável de estações pluviométricas.

São contempladas previsões com alcance de um dia, utilizando-se médias de totais diários precipitados de onze, sete e quatro estações, bem como leituras simples, de um único posto. Portanto, com perdas de eficiência em torno de 21%, 35% e 76%, respectivamente, para o segundo, terceiro e quarto caso, em relação ao primeiro.

Observa-se que as modificações dos conjuntos difusos das variáveis que representam as precipitações resultam em

reduções significativas da perda de eficiência, em todos os casos, em relação ao uso dos conjuntos difusos originais. Porém, permanece uma perda de eficiência residual, a qual pode ser atribuída principalmente à redução da capacidade de representação que ocorre com a diminuição do número de postos de aquisição de dados de chuva.

Na Tabela 11, são apresentadas algumas estatísticas, para a comparação do desempenho dos modelos, para os diversos alcances, e com uso de diferentes configurações, em termos do número de postos de coleta de dados de chuva.

Comparando-se os coeficientes obtidos utilizando-se sete postos pluviométricos com os coeficientes apresentados na Tabela 8, observam-se diferenças de desempenho, porém estas diferenças são menos importantes quando se considera que as previsões são feitas por categorias difusas combinadas.

Por exemplo, as frequências de ocorrência combinada, para previsões com 1, 1 ½ e 2 dias de alcance variam, respectivamente, de 98,8%, 97,4% e 96% para 98,6%, 97,2% e 95,2%, quando se passa da condição de uso de 11 para sete postos.

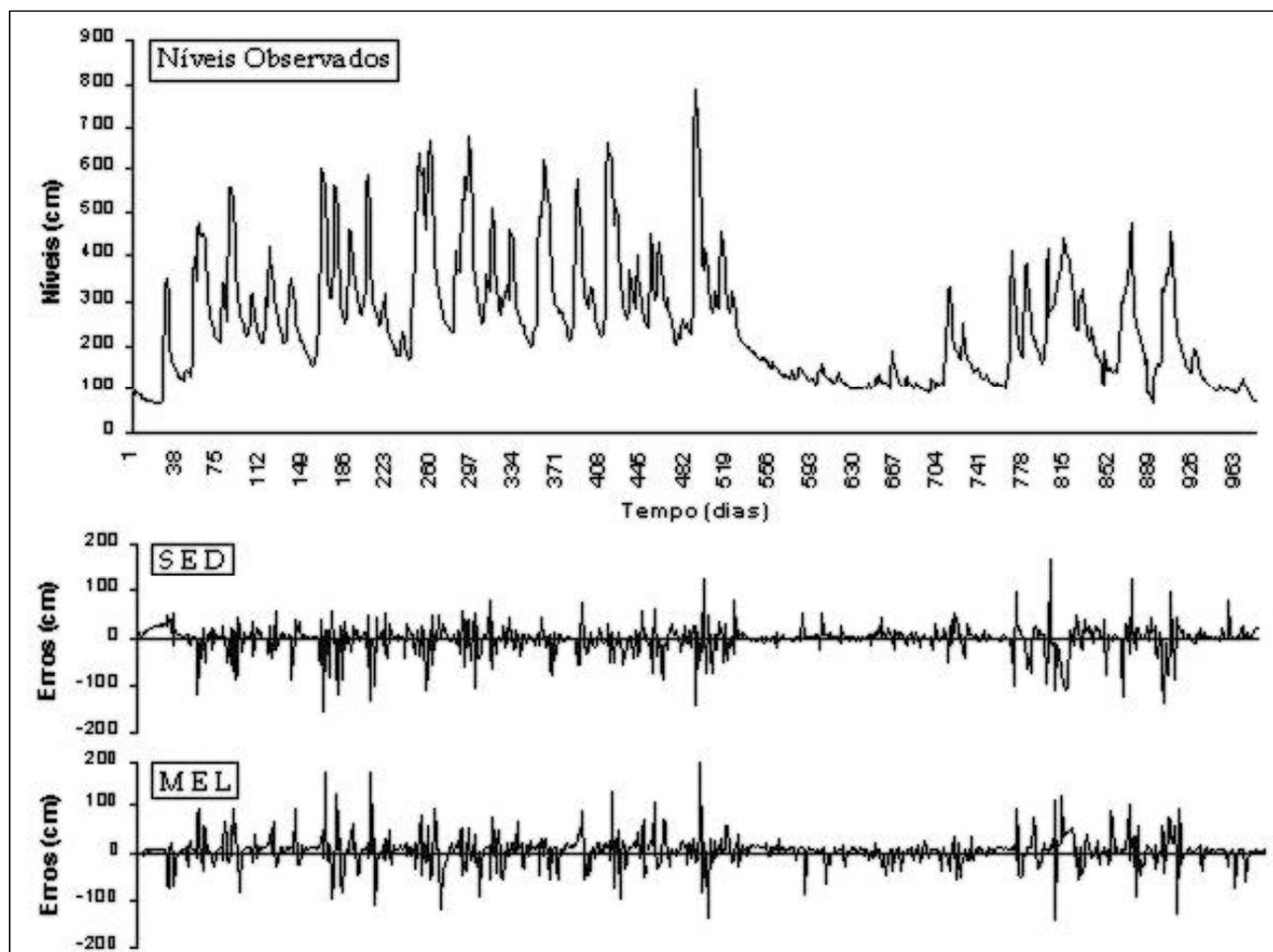


Figura 15. Previsões com Sistema Especialista Difuso (SED) e Modelo Empírico Linear (MEL): Alcance de 1½ dias.

É interessante que previsões difusas podem ser realizadas com uso de sete postos pluviométricos, com cerca de 95% de acerto nas categorias, para até dois dias de alcance.

Caso sejam necessários rigorosamente 98% de acerto nas categorias combinadas, o alcance máximo que pode ser obtido com os modelos pesquisados é de um dia. Porém, esta margem de acerto é possível, neste caso, mesmo com apenas quatro postos pluviométricos.

Finalmente, o uso de apenas um posto pluviométrico prejudica bastante os resultados, em uma taxa que aumenta com o alcance da previsão. É possível, porém, que, mesmo nestas circunstâncias, previsões sob forma de categorias difusas possam, na prática, ter utilidade.

## CONCLUSÕES

Os principais recursos matemáticos que foram concebidos para o sistema tiveram por objetivo o aprendizado, sendo baseados em uma concepção alternativa às FAM's para a manipulação de conhecimento, a qual foi denominada "matriz cognitiva".

Uma situação real foi utilizada para testes destes recursos, durante a qual foram identificadas dificuldades específicas da técnica dos sistemas especialistas para a previsão ou simulação de fenômenos naturais. Estas dificuldades são típicas de aplicações práticas, uma vez que resultam do uso de amostragens limitadas, sendo, principalmente, devidas aos dois fenômenos observados, denominados de "falta de conhecimento" e "falso conhecimento".

As observações realizadas revelaram que, enquanto a ausência de componentes no modelo pode produzir resultados insuficientes, o excesso destas pode prejudicar, pela produção de "falso conhecimento".

Uma vez que o aumento do número de combinações exige amostragens maiores para o aprendizado eficiente e que, em situações práticas, dispõe-se quase sempre de séries limitadas de observações, a "falta de conhecimento" decorrente constitui um fator restritivo à complexidade dos sistemas especialistas difusos.

Esta dificuldade, porém, foi superada pelo recurso das "bibliotecas combinadas de conhecimento", desenvolvido ao longo desta pesquisa. A idéia é que a mesma biblioteca

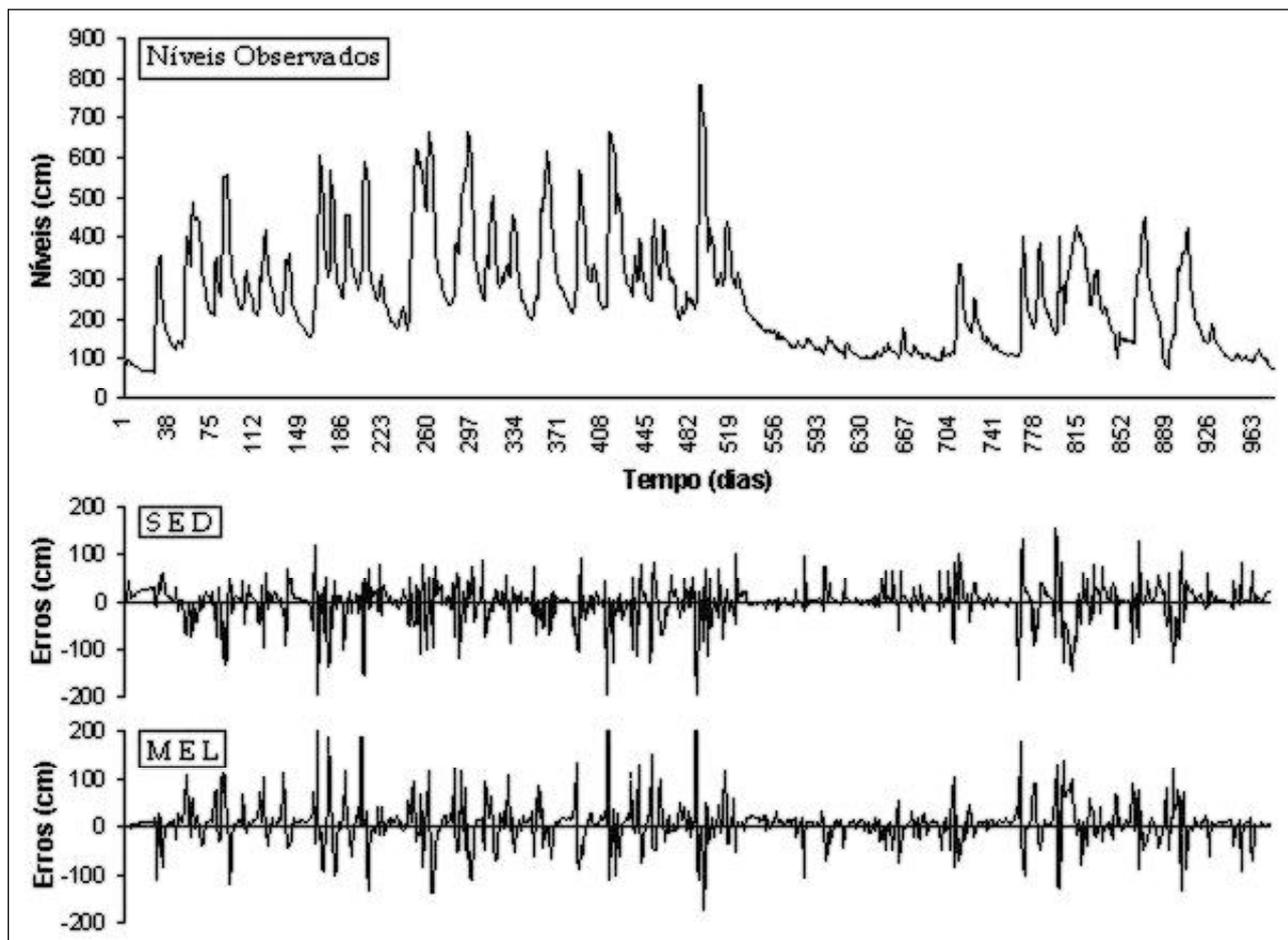


Figura 16. Previsões com Sistema Especialista Difuso (SED) e Modelo Empírico Linear (MEL): Alcance de 2 dias.

Tabela 8. Desempenho comparativo dos modelos.

Alcance (dias)	1		1 ½		2	
	SED	MEL	SED	MEL	SED	MEL
Ep (cm)	23,6	25,0	29,2	30,9	40,3	46,0
Ex (cm)	135	236	169	236	329	374
Ce (%)	96,9	96,5	95,3	94,7	91,0	88,3
Fot (%)	86,7	85,4	83,0	83,5	78,3	76,0
Fop (%)	98,7	98,6	97,3	97,2	95,6	93,1
Foc (%)	98,8	98,7	97,4	97,2	96,0	93,3

Tabela 9. Novas posições dos protótipos extremos.

Número de postos	Alcance (dias)		
	1	1 ½	2
11 postos	65	80	85
07 postos	115	120	120
04 postos	130	115	115
01 posto	160	170	180

de conhecimento pode conter regras associativas de modelos mais simples, que requerem menor número de combinações, e que são acionadas à medida que os conjuntos de regras correspondentes aos modelos mais complexos não são suficientemente completos.

A cada novo recurso experimentado para o aprendizado as melhoras foram pouco aparentes, resultando, porém, significativa ao final, o que não recomenda a supressão de nenhum destes procedimentos.

Os indicadores concebidos para avaliação das previsões categorizadas, chamados de coeficientes de agrupamen-

to, são a frequência de ocorrências totais (Fot), a frequência de ocorrências parciais (Fop) e a frequência de ocorrências combinadas (Foc). Estes coeficientes constituíram as bases para a interpretação, bem como para a avaliação da qualidade das previsões categorizadas.

Observou-se que previsões em categorias difusas, compostas de duas categorias adjacentes, com respectivos graus de pertinência, ainda são possíveis com frequências aceitáveis de acertos, até mesmo para alcances para os quais as previsões numéricas resultam inadequadas.

Uma vez que resultam valores excessivos para o erro máximo (Ex), mesmo quando o coeficiente de eficiência (Ce) é favorável, conclui-se que este coeficiente, bem como for-

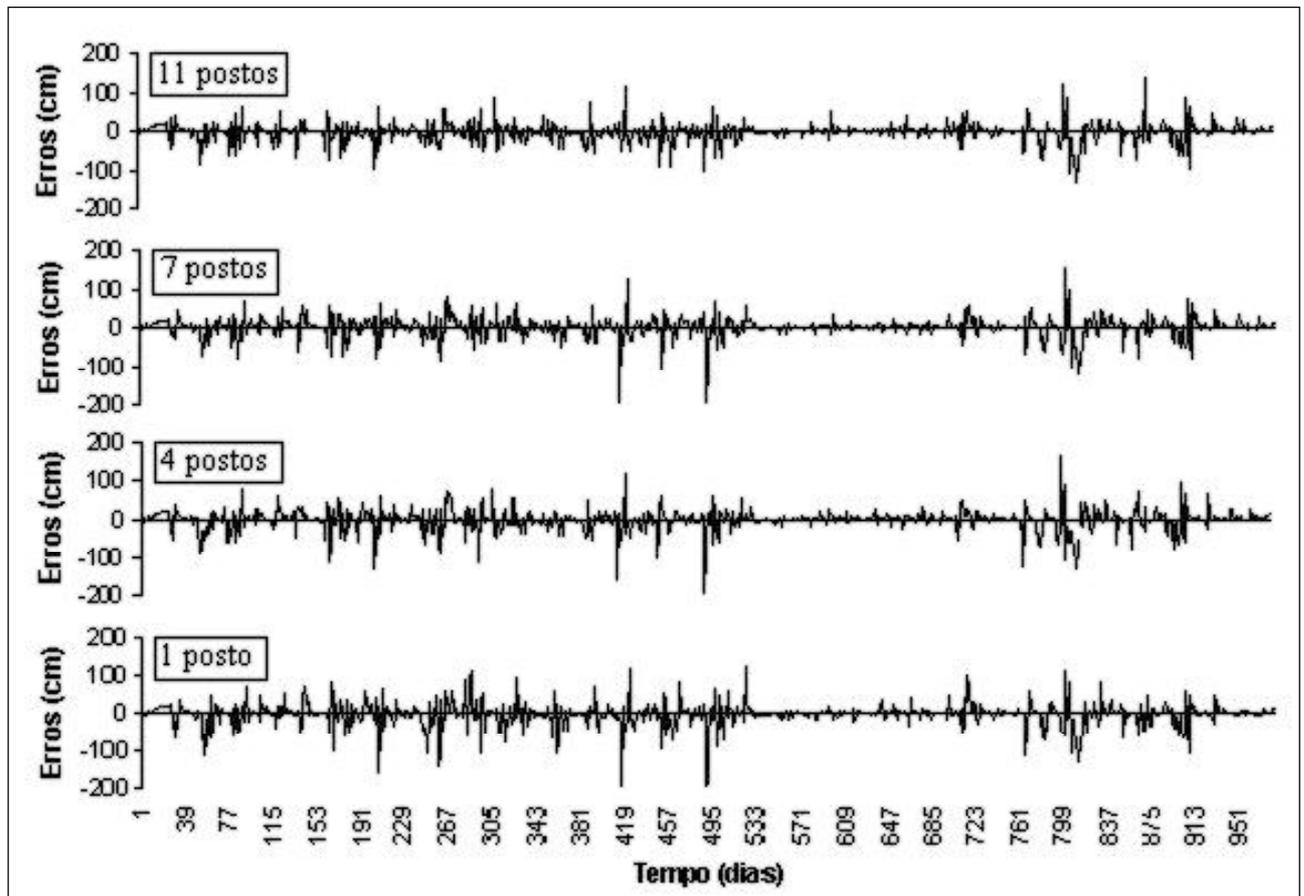


Figura 17. Previsões com alcance de 1 dia em função do número de postos.

Tabela 10. Perda de eficiência (%): Conjuntos originais (O) e modificados (M).

Alcance (dias)	1		1 ½		2	
	O	M	O	M	O	M
11	-	-	-	-	-	-
07	39,9	21,0	36,9	22,6	33,5	25,7
04	46,1	34,7	54,3	35,8	46,9	39,8
01	193	76,1	175	82,7	143	78,9

mas equivalentes baseadas no desvio quadrático, não é uma estatística confiável para previsões.

Foram identificadas as duas principais fontes de erros quando estes modelos são utilizados para previsão em situações reais, em sistemas com menor número de postos de aquisição de dados pluviométricos. Estas são a mudança da amplitude do domínio das variáveis que representam as precipitações, e a diferença na capacidade de representação das precipitações sobre a bacia.

A análise do efeito da redução do número de postos pluviométricos, realizada com uso de um coeficiente concebido para este fim (coeficiente de perda de eficiência), revela que, apesar da redução, ainda são possíveis pre-

Tabela 11. Desempenho comparativo por número de postos pluviométricos.

Alcance (dias)	1		1 ½		2				
	N. de Postos		N. de Postos		N. de Postos				
Ep (cm)	25,9	27,3	31,2	32,3	34,0	39,4	45,2	47,7	53,9
Ce (%)	96,3	95,9	94,6	94,2	93,6	91,4	88,7	87,4	83,9
Fot (%)	85,7	84,9	85,1	82,6	79,7	80,5	75,5	74,5	74,5
Fop (%)	98,5	98,3	97,7	97,2	96,6	95,8	94,6	93,6	92,1
Foc (%)	98,6	98,4	97,9	97,2	96,9	96,0	95,2	94,0	92,6

visões sob forma de categorias, mesmo para alcance de dois dias.

## AGRADECIMENTOS

Os autores agradecem às pessoas e instituições que forneceram os dados hidrológicos, essenciais para a execução desta pesquisa, especialmente a ANEEL (Agência Nacional de Energia Elétrica), e a CEEE (Companhia Estadual de Energia Elétrica).

## REFERÊNCIAS

- DRIANKOV, D.; HELLENDORRN, H. & REINFRANK, M. (1996). *An introduction to fuzzy control. 2<sup>nd</sup>*, Berlim: Springer-Verlag. 316p.
- KASABOV, N. K. (1996). *Foundations of neural networks, fuzzy systems and knowledge engineering*. Cambridge: Massachusetts institute of technology. 550p.
- KOSKO, B. (1992). *Neural networks and fuzzy systems: a dynamical systems approach to machine intelligence*. Englewood Cliffs: Prentice Hall. 449p.
- LANNA, A. E. e INDURSKY, A. B. (1983). Predição de vazões médias mensais com modelos empíricos. *Revista Brasileira de Engenharia. Caderno de Recursos Hídricos*. Rio de Janeiro: ABRH, vol.1, n°2, p.59-91.
- PEDROLLO, M. C. R. P. (1997). *Projeto análise de consistência de dados fluviométricos: bacias do Atlântico Sul, trecho sudeste, sub-bacia 87, sub-bacia do rio Camaquã*. Porto Alegre: DNAEE. 54p.
- PEDROLLO, O. C. e LANNA, A. E. (1991). Previsão de cheias em tempo atual em sistemas fluviais com contribuições laterais relevantes. *Revista Brasileira de Engenharia. Caderno de Recursos Hídricos*. Rio de Janeiro: ABRH, vol.9, n°2, p.25-51.
- ROSS, T. J. (1995). *Fuzzy logic with engineering applications*. New York: McGraw-Hill. 600p.
- ZADEH, L. A. (1965). Fuzzy Sets. *Information and Control*. New York: Academic Press. vol.8, p.338-353.

## *Real-Time Flood Forecasting Using Fuzzy Expert Systems*

### **ABSTRACT**

*Flood level forecasts are always approximations, and confidence intervals are not always suitable, particularly with low confidence probabilities, which result in excessively wide intervals. These intervals are therefore troublesome when river levels are very low and very high. Flood level forecasts are proposed as alternatives, both in the traditional numerical form, and in the form of vague categories. This is accomplished using an expert system based on fuzzy rules and fuzzy inference.*

*The methodologies are applied to real time river level forecasts in the Camaquã river basin (15543 km<sup>2</sup>), for lead times of one and two days. Practical difficulties related to the use of fuzzy systems are identified and explored.*

*The effects of raingauge network density are analyzed, and it is found that forecasting systems may be operated even where network density is sparse, given that fuzzy expert systems are available for symbolic predictions.*

*Key-words: short-term forecasts; fuzzy expert systems; symbolic forecasts.*