

Avaliação de Redes de Monitoramento Fluviométrico Utilizando o Conceito de Entropia

Wilde Cardoso Gontijo Junior, Sérgio Koide

Departamento de Engenharia Civil e Ambiental da Universidade de Brasília – DF
cardosowilde@hotmail.com; skoide@unb.br

Recebido: 10/01/11 - revisado: 16/05/11 - aceito: 23/09/11

RESUMO

Este trabalho propõe aplicar o conceito de entropia na otimização de redes fluviométricas em bacias com dados coletados há vários anos, reduzindo suas redundâncias e redirecionando a locação de estações. Os resultados apresentados em dois casos de estudo mostram sua praticidade de aplicação tendo em vista necessitar unicamente da locação das estações e dos dados brutos coletados. O conceito de entropia de séries de dados utilizado neste artigo foi proposto por Shannon e Weaver (1962), em 1948, através da publicação da Teoria Matemática da Comunicação. Trata-se de incorporação, de forma análoga, do conceito concebido para a Termodinâmica na análise da quantidade de informações transmitida por uma série de dados.

Palavras-chave: Redes fluviométricas, entropia.

INTRODUÇÃO

A locação de estações para coleta de dados fluviométricos tem obedecido a demandas isoladas referentes às diferentes finalidades das redes instaladas o que promove a dispersão e duplicação de esforços, materiais e financeiros, na busca do melhor conhecimento do comportamento hidrológico dos cursos d'água. Outro fator que agrava o processo de coleta de dados é a sua descontinuidade, frequentemente afetada pelas dificuldades operacionais para manutenção e recuperação dos postos e pela indisponibilidade financeira necessária para o seu sustento.

A análise da rede com a verificação tanto da qualidade dos dados coletados como do seu comportamento temporal deve ser sistemática e expedita para que estudos subsequentes possam ter reduzida a sua imprecisão. Há de se ressaltar que as incertezas dos estudos hidrológicos não poderão ser eliminadas com coletas mais confiáveis, porém, certamente uma boa contribuição à redução das imprecisões dar-se-ia com uma melhor qualidade geral dos dados.

A integração dos operadores e das informações coletadas, esforço que vem sendo empreendido pela tentativa de instalação no Brasil do Sistema Nacional de Informações de Recursos Hídricos (SNIRH), é uma forma de otimizar e ampliar a ca-

pacidade do país na gestão de suas águas. No entanto, além disso, faz-se necessária a construção de ferramentas estatísticas que permitam avaliar qualitativamente as redes existentes e orientar o processo de escolha das estações e pontos mais adequados para o atendimento às diversas finalidades das redes como o atendimento a estudos hidrológicos básicos ou o monitoramento de metas quantitativas visando a garantir os usos múltiplos nas bacias hidrográficas.

Este artigo apresenta um método que utiliza o conceito de entropia, introduzido pela teoria matemática da comunicação, para analisar qualitativamente os dados coletados, baseando-se, exclusivamente, nesses dados. Tal proposta foi objeto de Tese de doutoramento de Husain (1979) além de artigo de Caselton e Husain (1980), ambos nos Estados Unidos, importantes referências ao estudo das ferramentas matemáticas aplicadas neste trabalho e tem inspirado artigos mais recentes buscando um melhor arranjo para as redes fluviométricas, visando à sua otimização, conforme pode ser visto em Mishra e Coulibaly (2010).

No Brasil, esse método foi utilizado na análise de dados de uma rede de monitoramento de qualidade de água (Soares, 2001) e mostrou-se adequado. Os resultados apresentados neste artigo podem contribuir com o aprimoramento e otimização da locação das estações e, também, para a avaliação de áreas de transição de regimes hidrológicos em uma bacia hidrográfica.

MÉTODOS PARA OTIMIZAÇÃO DE REDES

Neste artigo são apresentados alguns métodos utilizados para a proposição da localização das estações que permitem avaliar a qualidade da rede e dos seus dados. Diversos outros métodos são apresentados e discutidos por Gontijo Jr. (2007).

O método da OMM

Segundo a Organização Meteorológica Mundial - OMM (WMO, 1994): “Um adequado sistema de monitoramento de dados hidrometeorológicos é planejado de forma que, dentro de certos níveis de precisão pré-estabelecidos, os parâmetros de projetos utilizados no planejamento, controle e gerenciamento dos recursos hídricos possam ser perfeitamente caracterizados e definidos, de forma direta, em qualquer ponto da bacia hidrográfica. Nessas condições, seria possível afirmar que existe uma rede ótima de estações de monitoramento.”

O planejamento e o projeto de uma rede de monitoramento confundem-se com a análise da rede existente. Dificilmente se depara com regiões onde não há qualquer informação hidrológica oriunda de estações já instaladas. Nesse sentido, a OMM propõe o seguinte procedimento para análise e rearranjo de redes: a) definir a finalidade da rede sob o ponto de vista dos usos e usuários dos dados e informações geradas; b) descrever os seus objetivos específicos e as redes especialistas; c) definir as prioridades entre as redes analisadas; d) avaliar as redes existentes por intermédio de análises quantitativas e qualitativas tendo em vista os objetivos propostos na implantação da rede, além da avaliação comparativa com redes instaladas em outras bacias; e) elaborar projeto de rearranjo da rede visando a corrigir as distorções apresentadas na avaliação precedente; e f) avaliar se os custos financeiros para implementação do novo projeto são compatíveis com os recursos disponibilizados.

A avaliação financeira deve considerar os aspectos relativos à implantação ou modernização de equipamentos, a frequência de visitação às estações, a coleta dos dados e a estrutura operacional necessária para todas as ações pertinentes.

Vê-se que tal proposta é abrangente e considera fatores que explicitam a complexidade da decisão a ser tomada. Um aspecto basilar em todo o processo, no entanto, são os dados e a avaliação das informações coletadas pela rede. O procedimento proposto pela OMM sugere uma avaliação preliminar (apresentada na figura a seguir) na qual são

definidas densidades mínimas para a locação das estações fluviométricas, considerando diferentes condições morfológicas do terreno. Foi acrescentada à Figura 1 uma linha de corte visando separar as correntes de água em principais e secundárias (Llamas, 1996). As correntes principais são aquelas cujas áreas de bacia são maiores que a área “A”. As secundárias são as demais. Segundo Llamas “para uma amostragem adequada deve haver pelo menos tantas estações nas correntes secundárias quanto nas correntes principais”.

Condições morfológicas do terreno	Densidade mínima de estações (km ² / estação)	Área das correntes principais (km ²)
Zona costeira	2.750	-
Relevo montanhoso	1.000	100
Planícies interiores	1.875	3.000 a 5.000
Ondulações leves	1.875	3.000 a 5.000
Pequenas ilhas	300	-
Área polar ou árida	20.000	10.000

Figura 1 – Densidade mínima de estações (OMM)

A aplicação desse procedimento requer conhecimento aprofundado do relevo e da hidrografia local, dados nem sempre prontamente disponíveis.

O método do Federal Institute of Hydrology

A Alemanha possui, desde a metade do século XIX, completa rede de monitoramento dos seus rios que vem gerando informações praticamente ininterruptas até os dias de hoje (Belz e Engel, 2003). Visando à otimização de custos de manutenção dessa rede, foi elaborado um método que busca identificar a significância das estações para a geração das informações hidrológicas e para a gestão dos recursos hídricos naquele país.

A abordagem desse procedimento vem ao encontro da legislação brasileira quanto ao fundamento da garantia dos múltiplos usos da água: a sua premissa principal é que uma rede de monitoramento deve ser compartilhada pelos diversos usuários assim como as informações oriundas dessa rede. A proposta metodológica é descrita sucintamente a seguir: a) mapeamento dos usos das informações produzidas pela rede de monitoramento, dividido em dois tipos básicos de dados: níveis da água nos

corpos d'água e vazões coletadas; b) elaboração de critérios para avaliar as estações de monitoramento quanto aos seus usos separando-as em dois grupos: grupo A, estação com destacada importância; e B, estação com importância específica; e c) julgamento quanto à avaliação da estação.

Esse método utiliza o conceito de redundância das informações no processo decisório final para otimização da rede quando afirma que “uma estação é dispensável sob o ponto de vista hidrológico quando os dados de uma estação vizinha coincidem ou há uma forte correlação entre os mesmos”. Alertam os autores, no entanto, que a tomada de decisão não deve ser feita somente sob o ponto de vista hidrológico e que devem ser considerados os usos específicos da estação em questão, inclusive quanto ao período a ser contemplado pelo planejamento dos recursos hídricos. A aplicação desse método implica a aceitação de critérios subjetivos que devem nortear o processo decisório.

O método Karasiev

Karasiev, apud Llamas (1996) e Mollinedo (2000), elaborou um método para otimização dos pontos de observação de uma rede baseando-se nos trabalhos desenvolvidos por Drozdov, Shepelevski, Gandin e Kogan, apud Mollinedo (2000). Seu objetivo é estabelecer a área ótima para a locação de estações fluviométricas. Devem ser definidas, inicialmente, regiões que tenham características geográficas e hidrológicas relativamente homogêneas. Três critérios, baseados nas áreas de drenagem da bacia homogênea sob análise, são utilizados para o cálculo da densidade ótima das estações:

- a) área mínima - Amin: superfície considerada pouco sensível às variações locais de uso e ocupação do solo, sendo representativa das condições climáticas e hidrológicas em um longo período;
- b) área relativa à variabilidade hidrológica - Agrad: superfície que garante que em 95% das vezes a diferença entre as vazões específicas anuais de duas bacias subsequentes é maior que o desvio-padrão das vazões;
- c) área relativa à correlação hidrológica mínima - Acor: superfície máxima que assegura a correlação entre as estações e, conseqüentemente, um erro máximo admitido para valores medidos em cada uma das estações.

Karasiev define, então, a área de drenagem ótima para a localização de uma estação como aquela que esteja confinada entre Acor e Agrad.

O método de Karasiev sugere, ainda, que uma rede de monitoramento ótima deveria possuir entre 15 e 30% das estações instaladas em bacias de drenagem com ordem inferior àquela da área ótima. Esse método, como pode ser observado, requer conhecimento aprofundado sobre as características hidrológicas da bacia, o que é, muitas vezes, o próprio objetivo da rede a ser instalada.

UTILIZANDO O CONCEITO DE ENTROPIA

Os métodos apresentados anteriormente requerem uma série de informações que ou são objeto da própria rede ou são de disponibilidade muito improvável na maioria das bacias hidrográficas. O método utilizando o conceito de entropia visa justamente possibilitar alcançar o mesmo objetivo com uma menor quantidade de informações sobre a bacia. Para sua aplicação são necessários apenas os dados brutos das estações e a definição de sua rede de drenagem. Ele se baseia no conceito de entropia desenvolvido por Shannon e Weaver (1962) para a teoria matemática da comunicação. Eles propuseram um tratamento análogo entre os conceitos de informação transmitida e de entropia aplicável às séries de dados. A entropia é classicamente definida pela 2ª lei da termodinâmica e pode ser entendida como o aumento da desordem interna de um sistema enquanto a reserva interna de energia livre diminui (Epstein, 1988). Shannon e Weaver definem que a informação intrínseca de um evento é representada por $I = -\log_2 p_k$, denominada autoinformação. Considerando que p_k é a probabilidade de ocorrência de um evento, com valor normalizado entre 0 e 1, o valor de I será tanto maior quanto mais rara sua ocorrência. A informação contida em uma série de dados é o resultado da equação 1 a seguir:

$$I_t = - \sum_{k=1}^N \log_2 p_k \times m_k \quad (1)$$

Na qual:

- p_k = probabilidade de ocorrência de um evento
- m_k = número de vezes que o evento ocorreu
- N = número de eventos únicos da série

A informação média associada a uma série hidrológica, ou entropia própria, é dada, então, pela equação 2 a seguir:

$$I_{med} = -\frac{\sum_{k=1}^N mk \times \log_2 pk}{N} = -\left(\sum_{k=1}^N pk \times \log_2 pk\right) \quad (2)$$

A informação média será máxima (I_{max}) quando a probabilidade de ocorrência de todos os eventos for a mesma, ou seja, quando há a máxima incerteza ou máxima entropia da série ($I_{max} = \log_2 N$). Assim, quanto mais desiguais forem as probabilidades de ocorrência dos dados numa série, menor será a quantidade média de informações transportada pela série.

Quando se analisa uma função distribuição de frequência de uma série de dados a percepção visual da disposição dos dados pode indicar a forma da curva para a qual tenderia sua distribuição. Nesse caso, os dados não são tão desordenados e há a tendência de que seu comportamento seja previsível. Chamamos esse “bom” comportamento de redundância ou sobreposição de informações contidas na série. Shannon propõe medir essa redundância como a diferença entre a máxima informação que poderia ser transportada e a informação média verificada na série. Essa redundância própria de uma série é calculada pela equação 3 a seguir:

$$R = \frac{I_{max} - I_{med}}{I_{max}} = \frac{(\log_2 N) + \left(\sum_{k=1}^N pk \times \log_2 pk\right)}{\log_2 N} \quad (3)$$

Por sua vez, analisadas duas séries de dados, essas podem conter informações independentes ou trazer redundâncias nas informações que transportam. Nesses casos, analisada a quantidade de informação transportada pelas duas séries, ela será menor do que se quantificadas separadamente. Por exemplo, ao compararmos duas séries fluviométricas instaladas em sequência em um mesmo rio ou bacia, provavelmente, haverá uma quantidade de informação que foi “transportada” da série a montante para a de jusante. A essa “redundância” denomina-se entropia condicional.

Pode-se, então, afirmar que quando há estações com séries de dados “bem comportados” há um comportamento “padrão” para uma determinada região hidrográfica e esta poderia ser qualificada como região homogênea. Outra proposição a ser feita nesse caso é que novas informações para essas

séries não são tão importantes quanto naquelas estações cujos valores coletados não se submetem ao mesmo bom comportamento, com uma variância maior. Ou seja, se for considerado que as estações fluviométricas apresentarão um comportamento padrão com a evolução temporal da coleta, um alto valor de entropia de uma série indicação objetivamente que é necessário conhecer mais sobre o regime fluviométrico naquela estação. Pode-se resumir a teoria revisada com as seguintes definições:

- a) entropia própria: característica intrínseca de uma série de dados resultante da equação 4 a seguir:

$$H(x) = -\int_{-\infty}^{\infty} p(x) \log p(x) dx \quad (4)$$

- b) entropia associada: característica relativa à redundância de informações contida em várias séries de dados, considerando a análise conjunta dessas séries, resultante da equação 5 a seguir:

$$H(x,y) = -\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} p(x,y) \log p(x,y) dx dy \quad (5)$$

- c) entropia condicional: característica relativa à redundância sequencial entre duas séries de dados, resultante das seguintes equações 6 e 7 a seguir:

$$H_x(y) = -\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} p(x,y) \log \left(\frac{p(x,y)}{p(x)}\right) dx dy \quad (6)$$

$$H_y(x) = -\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} p(x,y) \log \left(\frac{p(x,y)}{p(y)}\right) dx dy \quad (7)$$

Utilizando-se dos estudos de Shannon, Soares (2001) desenvolveu matematicamente as equações a serem utilizadas para aferir as entropias de séries de dados fluviométricos cujos comportamentos esperados sejam funções-distribuição multivariadas Normal ou Log-Normal. Esses são comportamentos frequentes para situações de redes de monitoramento consideradas básicas. A entropia associada é dada pela equação 8 a seguir:

$$H(X_m) = \left(\frac{M}{2}\right) \ln(2\pi) + \left(\frac{1}{2}\right) \ln|C| + \frac{M}{2} - M \ln(\Delta X) \quad (8)$$

Na qual:

H = entropia associada (ou própria com M=1)

M = número de amostras (ou estações)

X_m = vetor de M séries de dados

ICI = determinante da matriz das covariâncias C

ΔX = intervalo de amostragem dos dados

Utilizando-se essa formulação de Soares (2001) para a análise de séries de dados fluviométricos, verificou-se que a parcela referente ao intervalo de amostragem, considerada igual para todas as séries, poderia ser desprezada. Assim, adaptando-se a proposta de Soares, tem-se para a entropia condicional a equação 9 a seguir:

$$H(X_1, X_2, \dots, X_j - 1 | X_j) = H(X_1, X_2, \dots, X_j - 1, X_j) - H(X_j) \quad (9)$$

Na qual:

$$H(X_1, \dots, X_j) = \left(\frac{j}{2}\right) \ln(2\pi) + \left(\frac{1}{2}\right) \ln|C| + \frac{j}{2}$$

Nas quais:

H(X/Y) = entropia condicional de X, dado Y

j = número de amostras (ou estações)

O Método aplicado a redes fluviométricas

O método utilizado neste artigo foi adaptado de Soares (2001) e consiste no seguinte procedimento:

1ª etapa – verificação de premissas

A metodologia pode ser utilizada nas seguintes situações: a) estações fluviométricas com séries com comportamento hidrológico ergódico (nas quais os eventos são estatisticamente estacionários) e com área de drenagem superior a 1.000km²; b) amostra com um mesmo período de tempo (?t) para todas as séries de dados das estações, privilegiando aqueles períodos suscetíveis a menores influências com relação às dependências espacial e temporal dos dados coletados; e c) bacias cuja função-distribuição esperada seja Normal ou Log-Normal.

2ª etapa - entropia própria de cada série

- a) calcular a entropia própria da série X pela equação 10 a seguir:

$$H(x) = \left(\frac{1}{2}\right) \ln(2) + \left(\frac{1}{2}\right) \ln\{\sigma^2\} + \frac{1}{2} \quad (10)$$

Na qual:

H(x) = entropia própria

σ² = variância da série

- b) classificar as M estações em ordem decrescente (de 1 a M) quanto ao valor da entropia própria.

3ª etapa - cálculo do transporte de informações

O cálculo do transporte de informações permitirá a hierarquização das estações de suas formas diferentes: considerando o conjunto das estações na bacia ou considerando a situação espacial dessas estações. Para o primeiro caso, utilizar o seguinte procedimento:

- a) calcular o menor transporte de informações (TR) entre a estação de maior entropia (n^o 1) e, subsequentemente, cada uma das demais estações ordenadas na 2ª etapa, por intermédio da equação 11 a seguir:

$$\min [TR(X_1, X_j)] = \min [H(X_1) - H(X_1 | X_j)] \quad (11)$$

A entropia condicional será calculada, substituindo-se j pelo número da estação subsequente (de 2 a M), conforme equação 12 a seguir:

$$\min \{TR(X_1, X_j)\} = \min \left\{ H(X_1) - \left[\ln(2\pi) + \frac{1}{2} \ln|C| + 1 - H(X_j) \right] \right\} \quad (12)$$

- b) renumerar as estações, a partir da estação n^o 2 da 2ª etapa, classificando como a nova n^o 2 aquela estação cujo par com a n^o 1 apresenta o menor de transporte de informações;
- c) repetir os passos “a” e “b” para todas as estações renumeradas até a estação M; o resultado desse procedimento será a hierarquização das estações segundo a equação 13 a seguir:

$$\min \{TR(X_k, X_j)\} = \min \left\{ H(X_k) - \left[\ln(2\pi) + \frac{1}{2} \ln|C| + 1 - H(X_j) \right] \right\} \quad (13)$$

- d) utilizar os resultados dessa hierarquização na 4ª etapa.

Repete-se o procedimento anterior para o segundo caso (considerando a localização geográfica das estações), ou seja, calculando o transporte de informações considerando, duas a duas, estações a montante e jusante em uma mesma bacia hidrográfica. Nesse caso, os resultados devem ser apresentados em diagrama unifilar da bacia e permitirão avaliar o significado físico das entropias própria e condicional além de servir como referência para a avaliação final da rede.

4ª etapa – otimização das estações pelo cálculo da entropia associada

- a) calcular todos os transportes de informações (TR) na cadeia hierarquizada de estações definida na 3ª etapa, com a associação sucessiva de estações, conforme equação 14 a seguir:

$$TR[(X_1, X_2, \dots, X_{j-1}), X_j] = H(X_1, X_2, \dots, X_{j-1}, X_j) - [H(X_1, X_2, \dots, X_j) - H(X_j)] \quad (14)$$

Na qual “j” variará de 2 a M e a entropia associada calculada pela equação 15 a seguir:

$$H(X_1, \dots, X_j) = \left(\frac{j}{2}\right) \ln(2\pi) + \left(\frac{1}{2}\right) \ln|C| + \frac{j}{2} \quad (15)$$

- b) construir o diagrama de fluxo do transporte das informações, a cada passo, desde a estação nº 1 até a estação nº M, calculando-se a informação percentual transferida por intermédio da equação 16 a seguir:

$$TR \% (X_1, X_j) = \frac{TR[(X_1, X_2, \dots, X_{j-1}), X_j]}{H(X_1, X_2, \dots, X_{j-1})} \quad (16)$$

- c) alternativamente, pode-se calcular a razão da entropia associada (RH) a “j” estações e a entropia associada a todas as estações da rede (M), conforme equação 17 a seguir:

$$RH(X_1, X_j) = \frac{H(X_1, X_2, \dots, X_j)}{H(X_1, X_2, \dots, X_M)} \quad (17)$$

5ª etapa – avaliação da rede quanto à qualidade dos dados coletados nas estações

Os resultados apresentados na 1ª etapa indicam aquelas que necessitam de um melhoramento na coleta dos dados. Deve-se proceder, nesse caso, a uma investigação específica sobre as causas da alta entropia (pouco tempo de medição, equipamento, treinamento do leiturista etc.) visando a esse encaminhamento.

6ª etapa – avaliação da rede quanto à redundância de informações

Os resultados do transporte de informações entre estações subsequentes (3ª etapa – objetivo 2º) e aqueles oriundos dos resultados apresentados pelas razões da 4ª etapa precisam ser cotejados e, assim, se proceder a uma avaliação geral. Essa avaliação permitirá a proposição da supressão de estações (alto transporte de informações), a inclusão de estações (muito baixo transporte de informações), ou a otimização geral da rede retirando aquelas que são, no conjunto, redundantes.

CASOS DE ESTUDO

A bacia hidrográfica do rio São Francisco foi escolhida por contar com diferentes avaliações da rede instalada em suas sub-bacias, além de possuir todas as condições requeridas pelo procedimento proposto. Verificar-se-á a aplicação do método utilizando o Conceito de Entropia nas sub-bacias do rio das Velhas e do Alto Grande.

Bacia do rio das Velhas

A bacia hidrográfica do rio das Velhas tem área de drenagem de 27.815 km² e densidade de drenagem de 0,29 km/km². O curso principal tem 688 km de extensão e vazão média anual de 320 m³/s. Seus cursos d’água não apresentam barramentos significativos para a regularização das vazões. A rede de monitoramento fluviométrico é constituída por 105 estações. Segundo esse quantitativo, pelos padrões OMM, essa bacia é considerada superdensa relativamente à quantidade de estações, com 265 km²/estação. Observa-se, no entanto, que apenas 29 (27%) estações possuem medição de descarga líquida e, dessas, 12 estações (11%) apresentam área de drenagem superior a 1.000 km². Há uma predominância de estações com medição de parâ-

metros de qualidade (67 estações) o que sinaliza uma demanda de gestão mais acentuada sob o aspecto da qualidade da água.

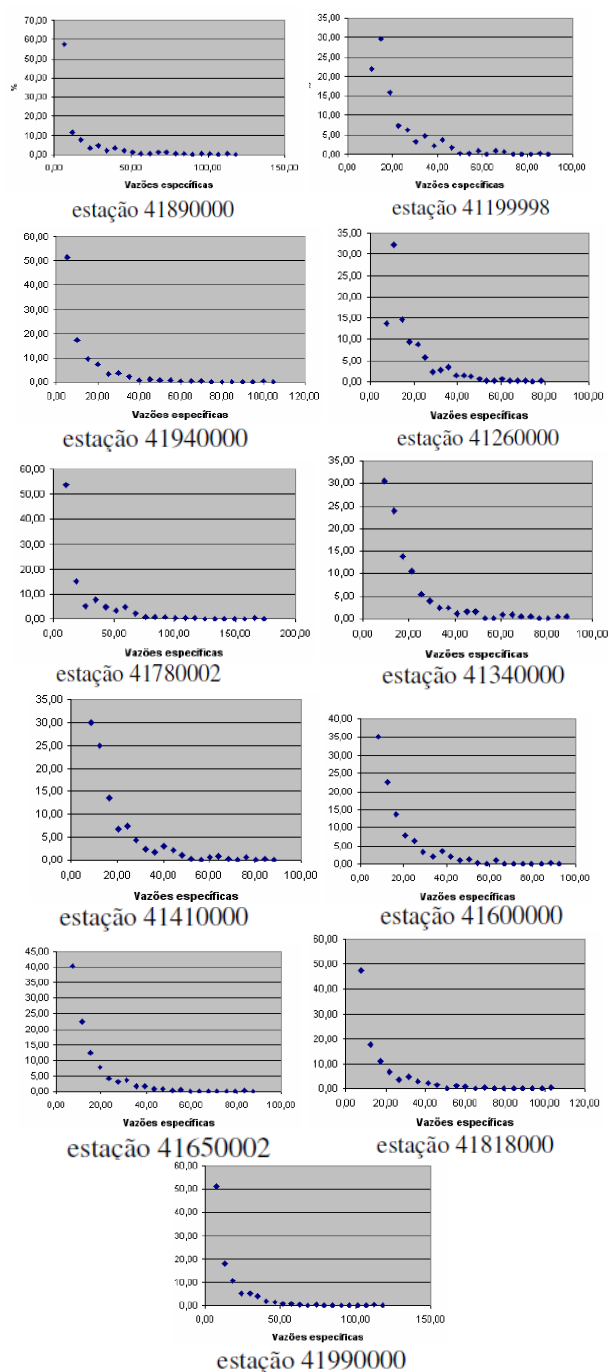


Figura 2 – Função distribuição estações BH Velhas

Para efeito de estudo foram selecionadas as estações com área de drenagem superior a 1.000

km2 e com tempo de coleta superior a 36 meses. No caso, todas possuem tempo de coleta superior a 28 anos, porém, com períodos diferentes de medição.

O período a ser considerado será de 1978 a 2006, intervalo no qual, à exceção da estação 41780002, todas as demais possuem número relevante de amostragens. A definição desse período deve-se, ainda, ao número de estações com leituras concentradas no período e cujo comportamento pode ser considerado adequado tendo em vista as mudanças ocorridas na ocupação socioeconômica nessa região hidrográfica. Destaque-se que a bacia é impactada pela aglomeração urbana da cidade de Belo Horizonte (MG) e uma série abrangendo períodos muito longos poderia ser bastante suscetível às mudanças espaciais promovidas por ela.

Considerou-se que a função-distribuição Log-Normal é a que melhor se ajusta aos padrões hidrológicos da bacia. A Figura 2 apresenta os gráficos dessas funções para as diferentes estações.

2ª etapa - entropia própria

A Tabela 1 a seguir apresenta os resultados dessa etapa. Em função dos valores das entropias próprias apresentarem-se muito próximos resolveu-se manter todas as estações como prioritárias para o monitoramento. No caso do descarte de algumas, optar-se-ia pelas de menor valor da entropia própria associada, ou seja, as estações 41260000, 41199998 e 41410000.

Tabela 1 – Entropia estações BH Velhas

Hierarquia inicial	Estação	Intervalo amostral relativo	Entropia própria 1978 a 2006
1	41890000	1,00	3,96
2	41780002	1,00	3,91
3	41990000	1,00	3,89
4	41940000	1,00	3,89
5	41818000	1,00	3,82
6	41650002	1,00	3,72
7	41600000	1,00	3,63
8	41340000	1,00	3,56
9	41410000	1,00	3,55
10	41199998	1,00	3,52
11	41260000	1,00	3,48

3ª etapa - cálculo do transporte de informações

Objetivo 1º - hierarquização das estações utilizando o conceito da entropia condicional.

Tabela 2 – Transporte informações estações BH Velhas

Hierarquia	Estação	Transporte
1	41890000	
2	41260000	0,10
3	41780002	0,14
4	41199998	0,29
5	41940000	0,35
6	41410000	0,60
7	41990000	0,62
8	41340000	0,62
9	41818000	0,65
10	41600000	0,76
11	41650002	0,88

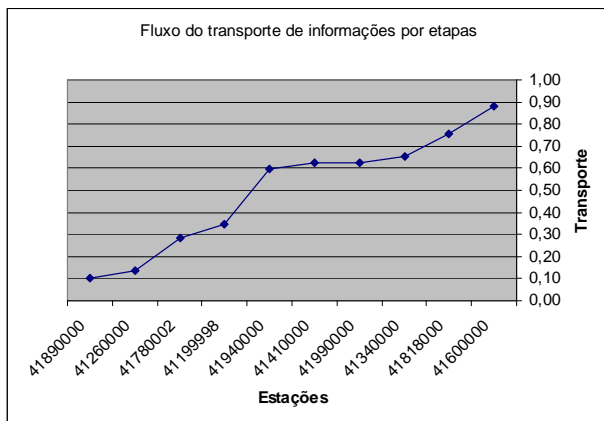


Figura 3 – Fluxo do transporte de informações BH Velhas

Objetivo 2º - cálculo do transporte de informações considerando a localização geográfica das estações.

No diagrama a seguir, os valores junto às setas correspondem ao valor do transporte e aqueles ao lado das estações à respectiva área de drenagem.

Observe-se na hierarquização do 1º Objetivo que o transporte de informações torna-se mais significativo entre as estações 41818000, 41650002 e 41600000. No diagrama da bacia, pode-se observar

que essas estações são sequenciais e encontram-se no centro da bacia.

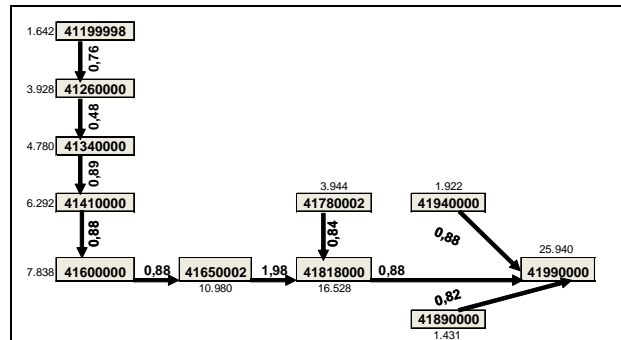


Figura 4 – Transporte informações BH Velhas (diagrama)

No cálculo para o 2º objetivo (diagrama da bacia) observa-se que há um menor transporte de informações entre as estações 41260000 e 41340000. A pequena diferença entre as áreas de drenagem (852 km²) dessas estações sugere que se trata de uma zona de transição quanto às características hidrológicas da bacia. Verifica-se, também, que a transferência de informações entre as estações 41650002 e 41818000 é grande, em que pese a grande diferença entre as áreas de drenagem (5.548 km²). Esse fato sugere que se trata de uma região homogênea e que a estação 41650002 deveria ser considerada candidata caso haja necessidade de uma redução do número de estações da rede.

Os valores altos para as transferências a partir das estações 41780002, 41890000 e 41199998 que, porém, abrangem regiões de cabeceiras, mostram que a avaliação conjunta dos objetivos (1º e 2º) não pode ser desconsiderada.

As estações sujeitas a um provável descarte nesta etapa não coincidem com aquelas com menor entropia apresentadas na 2ª etapa (41260000, 41199998 e 41410000). Esse fato indica que cada uma das etapas proporciona oportunidades distintas de julgamento e que se deve aplicar o método a toda a rede para depois se avaliar globalmente as possibilidades de otimização.

4ª etapa – otimização das estações pelo cálculo da entropia associada

A hierarquização da 3ª etapa é utilizada para o processo de otimização por intermédio do cálculo do transporte transferido e da razão da entropia geral ao se agregar a estação seguinte. Os resultados são apresentados na Tabela 3 a seguir.

Tabela 3 – Informações transferidas BH Velhas

HIERARQUIA TR (%)	ESTAÇÃO	TR(%)
1	41890000	
2	41260000	2,51%
3	41780002	12,11%
4	41199998	9,11%
5	41940000	8,02%
6	41410000	7,12%
7	41990000	6,03%
8	41340000	5,28%
9	41818000	4,80%
10	41600000	4,48%
11	41650002	4,11%

Considerando que um valor máximo do transporte de informações não deva ultrapassar 10% (TR% < 10% - proposição a ser pesquisada e confirmada com o acúmulo de experiência na aplicação desse procedimento), pela Tabela anterior, somente a estação 41780002 seria descartada.

Se, por outro lado, for considerado que o valor máximo da razão das entropias na rede (RH) deve ser igual a 90%, segundo a Tabela 4 a seguir, a estação a ser descartada seria a 41650002.

Tabela 4 – Razão das entropias BH Velhas

HIERARQUIA TR (%)	ESTAÇÃO	TR(%)
1	41890000	
2	41260000	12,72%
3	41780002	23,56%
4	41199998	33,26%
5	41940000	41,53%
6	41410000	50,68%
7	41990000	58,47%
8	41340000	67,43%
9	41818000	75,29%
10	41600000	83,94%
11	41650002	91,83%

Além dessas análises é necessário que os aspectos relativos à gestão dos recursos hídricos, para o atendimento às necessidades de usos específicos, tais como estudos de transporte de poluentes, e aquelas relativas ao custo-benefício de sua operação

sejam considerados antes da decisão final de descartar das estações.

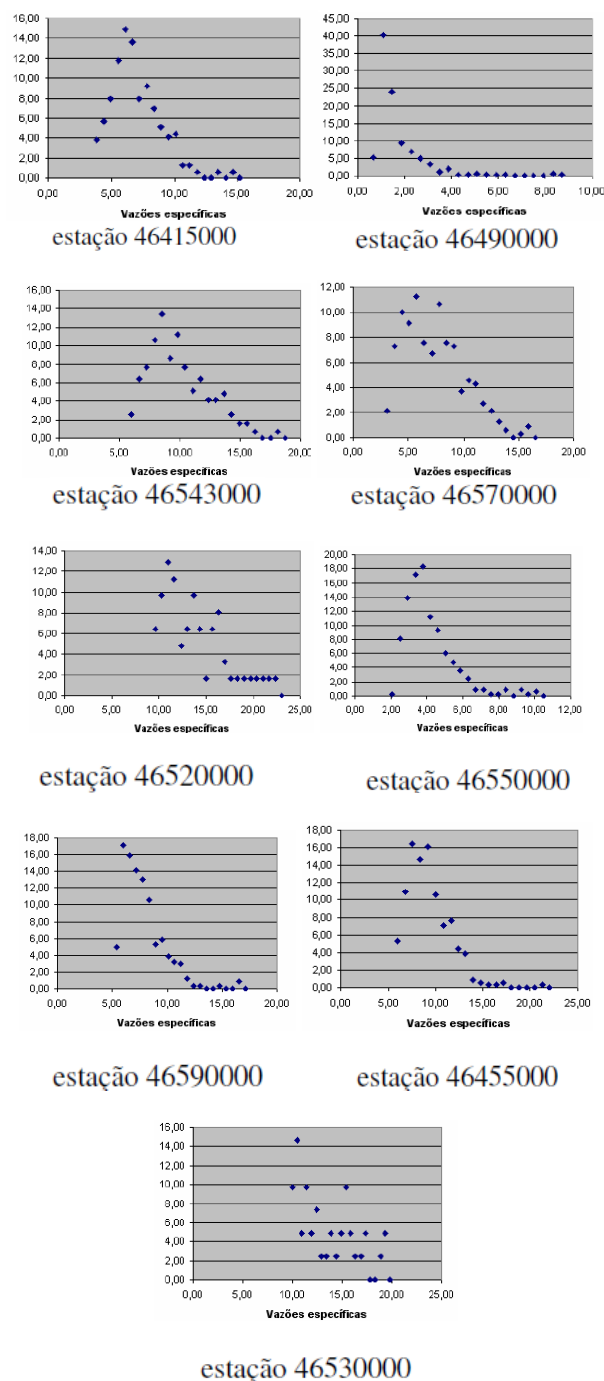


Figura 5 – Função distribuição estações BH Alto Grande

Bacia do Alto rio Grande

A bacia hidrográfica do Alto rio Grande tem área de drenagem de 33.544 km² e densidade de

drenagem de 0,13 km/km². O curso principal tem 204 km de extensão e vazão média anual de 155 m³/s. Seus cursos d'água não apresentam barramentos significativos para a regularização das vazões.

A rede de monitoramento fluviométrico é constituída por 25 estações. Segundo esse quantitativo, pelos padrões OMM, essa sub-bacia é considerada superdensa relativamente à quantidade de estações. Observa-se que 22 (88% das estações) possuem medição de descarga líquida e, dessas, 12 (48%) apresentam área de drenagem superior a 1.000 km². Há uma predominância de estações com medição de descarga o que sinaliza uma demanda de gestão mais acentuada sob o aspecto do controle quantitativo.

Para efeito deste estudo foram selecionadas 9 estações sendo que 7 estações apresentam tempo de operação superior a 28 anos e 2 estações possuem em dados coletados entre 3 e 6 anos.

Será considerada uma amostra das séries para o período de 1978 a 2006. Essa definição deve-se ao histórico do número das estações, com leituras concentradas nesse período e que, tendo em vista as mudanças ocorridas na ocupação socioeconômica na região, que abrange área muito impactada por projetos de irrigação (oeste baiano), é prudente não utilizar os dados da estação 46550000 para o período entre 1934 e 1978. Uma série abrangendo períodos muito longos poderia ser bastante suscetível à ocorrência de mudanças significativas de uso do solo.

Da mesma forma que para a sub-bacia do rio das Velhas, considerou-se que função-distribuição Log-Normal é a que melhor se ajusta aos padrões hidrológicos da bacia. A Figura 5 apresenta os gráficos dessas funções.

2ª etapa - entropia própria

A Tabela 5 a seguir apresenta os resultados desta etapa. Observa-se uma grande variação dos valores das entropias próprias. Optou-se, então, pela retirada somente da estação 46570000 de ambos os estudos e para as simulações das etapas posteriores. Essa opção levou em consideração a posição geográfica da estação, conforme ver-se-á à frente.

3ª etapa - cálculo do transporte de informações

Objetivo 1º - hierarquização das estações utilizando o conceito da entropia condicional

Vê-se que o transporte de informações é pouco significativo entre as quatro primeiras estações (46490000, 46415000, 46530000 e 46550000). A partir da estação 46550000, com a associação da estação 46520000, há uma transferência importante de informações ou uma redundância maior na coleta dos dados.

Tabela 5 – Entropia própria estações BH Alto Grande

Hierarquia inicial	Estação	Intervalo amostral relativo	Entropia própria 1978 a 2006
1	46490000	1,00	3,72
2	46550000	1,00	3,21
3	46590000	1,00	3,19
4	46455000	1,00	3,19
5	46415000	1,00	2,97
6	46543000	1,00	2,79
7	46520000	1,00	2,75
8	46570000	1,00	2,73
9	46530000	1,00	2,72

Tabela 6 – Transporte informações estações BH Alto Grande

Hierarquia	Estações	Transporte
1	46490000	
2	46415000	0,05
3	46530000	0,08
4	46550000	0,06
5	46520000	0,32
6	46543000	0,32
7	46590000	0,37
8	46455000	0,81

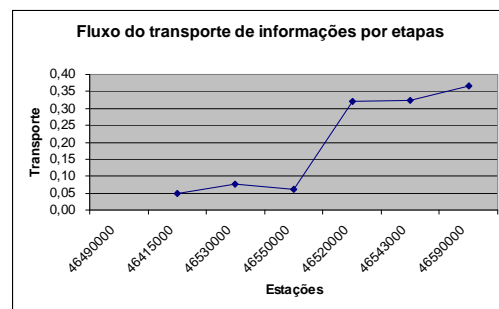


Figura 6 – Fluxo do transporte de informações BH Alto Grande

Objetivo 2º - calcular o transporte de informações considerando a localização geográfica das estações.

No diagrama a seguir, os valores junto às setas correspondem ao valor do transporte e aqueles ao lado das estações à respectiva área de drenagem.

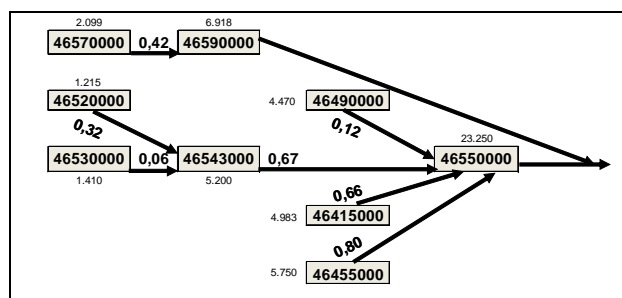


Figura 7- Transporte de informações BH Alto Grande (diagrama)

Vê-se pelo diagrama unifilar da bacia que a estação 46550000 é a que possui maior redundância nas informações relativas às estações 46543000, 46415000 e 46455000. No entanto, possui pouca informação da estação 46490000. Observa-se que as áreas de drenagem das três estações a montante da estação 46550000 são relativamente próximas o que não parece ter ajudado para que houvesse um comportamento hidrológico próximo entre as séries. Analisando-se a região drenada pela seção da estação 46490000, observa-se que se trata de uma região bastante diferente das demais sob aspectos hidrológicos, inclusive quanto ao regime pluviométrico. Por sua vez, a estação 46590000 possui pequena redundância com a estação 46570000 e nenhuma relação hidrológica com as demais. Assim, considerando as condições físicas da rede de drenagem, não devem ser descartadas as estações colocadas nos últimos lugares da hierarquização tais como as estações 46590000 e 46455000.

Por outro lado, observando-se o diagrama, verifica-se a necessidade de que seja instalada uma estação próxima ao exutório da bacia, o que, provavelmente, proporcionaria o descarte de alguma estação a montante.

4ª etapa – otimização das estações pelo cálculo da entropia associada

A hierarquização da 3ª etapa é utilizada para o processo de otimização por intermédio do cálculo do transporte transferido e da razão da entropia geral ao se agregar a estação seguinte. Os resultados são apresentados a seguir.

Tabela 7 – Informações transferidas BH Alto Grande

Hierarquia TR %	Estação	TR %
1	46490000	
2	46415000	1,31%
3	46530000	4,10%
4	46550000	8,60%
5	46520000	4,01%
6	46543000	6,79%
7	46590000	7,34%
8	46455000	6,25%

Tabela 8 – Razão das entropias BH Alto Grande

Hierarquia TR %	Estação	RH
1	46490000	
2	46415000	18,83%
3	46530000	33,59%
4	46550000	45,96%
5	46520000	58,22%
6	46543000	69,77%
7	46590000	79,17%
8	46455000	89,48%

Utilizando os mesmos critérios definidos para a sub-bacia do rio das Velhas, considerando que um valor máximo do transporte de informações não deva ultrapassar 10% e que o valor máximo da razão das entropias na rede (RH) deveria ser igual a 90%, pela análise das tabelas acima não há qualquer estação a ser descartada.

AVALIAÇÃO GERAL DOS RESULTADOS

Aplicado a duas regiões hidrologicamente muito distintas, o método mostrou-se capaz de apontar as diferenças intrínsecas relativas ao comportamento hidrológico das bacias.

A região hidrográfica do rio das Velhas possui estações mais antigas, séries mais consolidadas e muito homogêneas quanto ao tempo de coleta. No entanto, as variâncias das séries de dados são muito altas o que proporcionou valores maiores de entropia.

pia própria (acima de 3) para todas as estações. Por sua vez, as entropias próprias apresentaram valores muito próximos entre as estações o que poderá representar uma especificidade das vazões dessa bacia.

Pela análise da transferência de informações entre as séries da bacia do rio das Velhas pode-se supor que há uma maior redundância nas informações coletadas já que seus valores são sensivelmente maiores que aqueles apresentados na bacia do Alto rio Grande, o que pode indicar a necessidade de otimização das estações ali instaladas.

Na bacia do Alto rio Grande, a diferença entre as entropias próprias é significativa. Esse fato indica a necessidade de que nessa bacia algumas estações tenham aprimoradas suas medições, reduzindo a influência do fator tempo na avaliação dos resultados. A transferência de informações entre as estações, no entanto, indica que há uma independência grande entre as informações coletadas nos diferentes locais. Isto pode ser explicado pelas diferentes características hidrológicas da região: a área de drenagem da estação 46490000, que influencia diretamente a série da estação 46550000, é composta por cabeceiras em região semi-árida, com baixa produção hídrica, diferentemente das demais cabeceiras dos cursos monitorados. A transferência de informações pode indicar, também, que existe um número de estações menor do que seria necessário para essa região.

A construção de diagramas unifilares da rede de monitoramento, apresentando a locação geográfica na bacia das estações sob análise e permitindo a avaliação da transferência de informações à luz dessa localização, é um importante auxiliar no julgamento da interdependência das estações e para a análise da consistência dos resultados do processo de hierarquização e otimização da rede.

Observa-se no diagrama para a bacia do rio das Velhas que a dependência (ou redundância de informações) entre as estações é muito maior nessa bacia do que na bacia do Alto rio Grande. Vê-se, também, a dependência das informações geradas pelas estações hierarquicamente inferiores na rede (3ª etapa) e os valores nominais que podem sinalizar “degraus” nas condições hidrológicas da bacia. Um exemplo dessa última afirmação pode ser visto no transporte de informações entre as estações 41260000 e 41340000 na bacia do rio das Velhas ou entre as estações 46490000 e 46550000, na bacia do Alto rio Grande, com valores indicativos da transição de regiões hidrológicamente diferentes. Em ambos os casos há uma forte diferenciação dos valores nominais do transporte das informações.

Esses diagramas unifilares são importantes, ainda, para a avaliação do descarte de estações uma vez que as relações apresentadas pelo procedimento proposto para hierarquização e otimização não permitem sua visualização espacial, ferramenta importante para uma maior validação dos resultados.

CONCLUSÕES E RECOMENDAÇÕES

O método mostrou-se adequado à aplicação na avaliação de redes de monitoramento fluviométrico.

A análise dos resultados do procedimento sugere que sua aplicação pode ser muito útil para a avaliação da “qualidade” das séries de dados. Assim, as séries que têm valores relativos de entropia própria muito superiores àqueles observados nas demais devem ter suas medições priorizadas. Quanto ao descarte de estações, sugere-se que os valores de transporte de informações sejam observados com cautela visando verificar a correlação espacial entre as estações. Quando, além do procedimento proposto, elaborou-se o fluxo de informações entre as séries, de montante para jusante na bacia, os resultados indicaram com maior precisão a origem das redundâncias e aquelas que efetivamente podiam ser consideradas descartáveis. O resultado isolado da 4ª etapa (otimização geral) pode não guardar correspondência direta com a situação física real, o que poderá indicar avaliações incorretas dos resultados apresentados.

Certamente, a aplicação desse procedimento em um número maior de bacias hidrográficas deverá proporcionar uma visão mais crítica das suas vantagens e desvantagens, visando ao aprimoramento ou rejeição. Este trabalho indica, então, um possível caminho a ser seguido para que haja, de forma sistemática e menos dependente de outros fatores que não as próprias medições, uma avaliação técnica sistemática e expedita do funcionamento das redes de monitoramento fluviométrico. É prudente observar, no entanto, as premissas necessárias para a aplicação do método, principalmente quanto ao aspecto ergódico das séries, além de uma análise criteriosa de alterações da bacia hidrográfica (como do uso do solo) que possam, no intervalo temporal analisado, imputarem potencial modificação hidrológica.

REFERÊNCIAS

- BELZ, J. e ENGEL, H. Optimization of surface water observation networks. In: International workshop on hydrological networks for integrated and sustainable water resources management, 2003, Koblenz, Alemanha. Anais. p. 115-122.
- EPSTEIN, I. Teoria da Informação. 2ª. ed. São Paulo: Editora Ática, 1988. 77p.
- CASELTON, W. F. e HUSAIN, T.. Hydrologic networks: information transmission. In: Journal Water Resources Planning and Management, 1980.
- HUSAIN, Tahir. Shannon's information theory in hydrologic network design and estimation. Tese de doutoramento na Universidade British Columbia, EUA, 1979.
- LLAMAS, J. Curso intensivo de planejamento e projeto de redes meteorológicas e hidrométricas. Brasília: Secretaria de Recursos Hídricos do Ministério do Meio Ambiente, dos Recursos Hídricos e da Amazônia Legal, 1996. 66p.
- GONTIJO JR., Wilde Cardoso. Avaliação e redimensionamento de redes para o monitoramento fluviométrico utilizando o método Sharp e o conceito de Entropia. Dissertação de Mestrado na Universidade de Brasília, Brasília, 2007.
- MISHRA, A. K; e COULIBAY, P. Hydrometric network evaluation for Canadian watersheds. Journal of Hydrology (380), 2010. p. 420-437.
- MOLLINEDO, M. A. O. Optimizacion de la red hidrológica básica bajo los escenarios del cambio climático en la Republica de Bolívia. Resumo de Tese de doutoramento - Universidad Estatal Hidrometeorologica de Rusia, San Petersburgo, Rússia, 2000.
- SHANNON, C. E. e WEAVER, W. The Mathematical Theory of Communication. 9ª ed. Illinois: The University of Illinois Press: Urbana, 1962. 117p.
- SOARES, P. F. Projeto e avaliação de desempenho de redes de monitoramento de qualidade de água utilizando o conceito de entropia. Tese de doutoramento – Escola Politécnica da Universidade de São Paulo, Departamento de Engenharia Hidráulica e Sanitária, USP, São Paulo, 2001. 211p.
- WMO. Guide to hydrological practices – data acquisition and processing, analysis, forecasting e others applications. 15º ed. 1994. World Meteorological Organization n° 168. p. 259-287.

Evaluation of Fluvimetric Monitoring Using the Concept of Entropy

ABSTRACT

This work proposes to apply the concept of entropy in optimising networks within basins in which data have been collected during several years, reducing their redundancies and directing the stations to others places. The results of two case studies present its practical application because it only needs the station location and the collected data. The concept of data series entropy was proposed by Shannon and Weaver (1962), in 1948, through the publication of the Mathematical Theory of Communication. It is the analogous incorporation of the notion used in Thermodynamics in analysing the quantity of information transmitted by a series of data.

Key-words: *Fluviometric networks, entropy.*