Basin water quality network design: optimum sampling sites located by information theory

JANAÍNA BEZERRA MESQUITA[1] & SERGIO KOIDE[2]

1 *STN Ed.Toscana, Bloco B, apto 205. Asa Norte, Brasília – DF. CEP 70.770-100, Brazil*
jmesquita2008@gmail.com

2 *Universidade de Brasília – UnB. Campus Universitário Darcy Ribeiro, Prédio SG12, Programa de Pós-Graduação em Tecnologia Ambiental e Recursos Hídricos (PTARH), Brasília – DF. CEP 79.910-900, Brazil*

INTRODUCTION

Water quality monitoring is motivated by the information needed to maintain or restore water quality for ecological and human benefits. Despite its importance, the design of networks for monitoring water quality is still a challenging activity and there is no universally-accepted methodology for this purpose.

The design of basin water quality monitoring networks consists of: defining the location of water quality monitoring stations, water quality variables selection, and sampling frequency. As basin water quality monitoring involves high costs, the search for establishment of minimal representative networks has great importance. The literature review shows that first considerations for sampling sites definition were based on the basin topology (Sharp, 1971; Sanders *et al*., 1994; Harmancioglu *et al.*, 1999). Concepts of land use and drainage area as major factors to define contamination detection probabilities, in addition to the network monitoring ability to gather information, are introduced by this study. This proposal aims to use the entropy concept (Shannon & Weaver, 1949) to maximize the information gathered from the basin monitoring network.

INFORMATION THEORY

The mathematical theory of communication, also known as Information Theory (IT) was first studied by the mathematician Warren Weaver and then structured by the engineer Claude Elwood Shannon (Shannon & Weaver, 1949). Its first goal was the solution of some optimization problems about the transmission cost in the communication system. According to this theory the communication system can be represented as five main structures: the information source, the transmitter, the receiver, the noise and the destination. The information source selects a desired message, which is converted into a signal by the transmitter. The signal is sent over the communication channel. The receiver changes the transmitted signal back into a message and handles this message on to the destination. Unfortunately, in this process there are some unwanted distortions, called noises. The expression for the information developed in the IT is axiomatic and based on the principles of logarithmic and probability functions, represented as:

$$I = -\sum_{k=1}^{n} p_k \times \log p_k$$

where p_K is the probability of choice of a message; I is the quantity of information with the message; and n is the total number of events. As the information decreases just as soon as one message becomes more probable than another, the minus signal of the equation is necessary.

ADAPTATION OF INFORMATION THEORY TO DEFINING SAMPLING SITES

In order to apply this theory for spatial positioning of water quality monitoring stations, some considerations were assumed:

(a) the pollution detection probability is proportional to the drainage area contributing to the sample site;
(b) additional efforts must be made to find the precise location of the pollution source once it is detected at some site;

The expression to quantify the information gathered from the water quality monitoring network is:

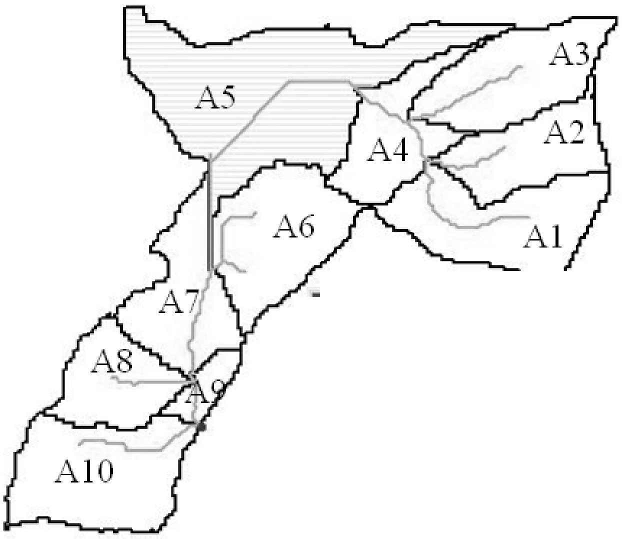
$$I = -\sum_{k=1}^{n} w_k \frac{A_k}{A} \times \log \frac{A_k}{A}$$

where A_k is the area of the drainage basin k; A is the total basin area and w_k is a factor related to the land use.

APPLICATION OF INFORMATION THEORY TO THE STUDY CASE OF THE DESCOBERTO RIVER BASIN

This methodology was applied to the Descoberto River basin, at the entrance of the Descoberto Lake, located in Distrito Federal (DF), the central region of Brazil. In the first scenario, considering two sampling points for this area, the results of this application were:

Sample points	Information
TOTAL and A1	0.508633917
TOTAL and A2	0.382799389
TOTAL and A3	0.466632884
TOTAL and A4	0.3797126
TOTAL and A5	0.788151853
TOTAL and A6	0.494469562
TOTAL and A7	0.422900349
TOTAL and A8	0.356601696
TOTAL and A9	0.139649647
TOTAL and A10	0.535747756



Scenarios for three sampling points where also simulated and this methodology seems to have a good potential to be applied, particularly for ungauged basins.

Acknowledgements This project is financially supported by the CAPES-Brazil and CNPq-Brazil.

REFERENCES

Harmancioglu, N. B., Fistikoglu, O. Ozkul, S. D. Singh, V. P. & Alpaslan, M. N. (1999) *Water Quality Monitoring Network Design.* Kluwer Academic Publishers. Water Science and Technology Library, vol. 33.

Sanders, T. G., Ward, R. C., Loftis, J. C., Steele, T. D., Adrian, D. D. & Yevjevich, V. (1994) *Design of Networks for Monitoring Water Quality*, 3ª Edição. Water Resources Publications, Littleton, Colorado, USA.

Shannon, C. E. & Weaver, W. (1949) *The Mathematical Theory of Communication.* Illinois, USA.

Sharp, W. E. (1971) A topologically optimum water sampling plan for rivers and streams. *Water Resour. Res.* **7**(6), 1641–1646.