

UTILIZAÇÃO DE PARÂMETROS FÍSICO-QUÍMICOS PARA ESTIMAR O APORTE DE NUTRIENTES NO ESTUÁRIO DO RIO PARAÍBA DO SUL

Torres, R.¹; Andrade, L. E. V.¹; Gatts, C. E. N.¹; Ovalle, A. R. C.²; Rezende, C. E.²

¹ Laboratório de Ciências Físicas, Universidade Estadual do Norte Fluminense Darcy Ribeiro, Av. Alberto Lamego 2000, 28013-602 Campos dos Goytacazes, RJ

² Laboratório de Ciências Ambientais, Universidade Estadual do Norte Fluminense Darcy Ribeiro, Av. Alberto Lamego 2000, 28013-602 Campos dos Goytacazes, RJ
gatts@uenf.br

RESUMO

Neste trabalho foram utilizadas técnicas de mineração de dados a uma série temporal de indicadores da qualidade da água na foz do Rio Paraíba do Sul, medida em Campos dos Goytacazes, para verificar uma eventual associação entre parâmetros físico-químicos e o aporte instantâneo de nutrientes no estuário. Os resultados demonstram que é possível prever o aporte de nitrato e fosfato com razoável acurácia, com o coeficiente de correlação entre o conjunto de previsões e o de valores observados na faixa de 60%.

Palavras chave: Séries temporais, mineração de dados, modelagem empírica.

INTRODUÇÃO

Estudos ambientais usualmente envolvem um número grande de variáveis, o que se mostra como um fator a mais de complicação na interpretação de resultados, particularmente quando se trata de estudos de longo prazo, posto que estes, invariavelmente, resultam em grandes conjuntos de dados.

Por outro lado, o acelerado desenvolvimento de recursos computacionais vem resultando em sistemas de análise de dados cada vez mais sofisticados e robustos. Dentre as ferramentas analíticas para este fim, merece destaque a KDD – Knowledge Discovery in Databases.

A proposta essencial da KDD consiste, grosso modo, na utilização de um verdadeiro arsenal de metodologias analíticas (análise de aglomerados, redes neurais, algoritmos evolucionários, inferência fuzzy etc) para encontrar padrões não-evidentes em grandes conjuntos de dados, que correspondam a informações relevantes a respeito do sistema em estudo (MAIMAN et al., 2008).

Neste trabalho foram utilizadas técnicas baseadas em redes neurais e análise de aglomerados por lógica fuzzy (fuzzy clustering) para analisar uma série temporal de amostragens realizadas entre janeiro de 1997 e dezembro de 2007 com periodicidade quinzenal, em Campos dos Goytacazes, na saída da bacia do Rio Paraíba do Sul, com o objetivo de investigar uma eventual associação entre parâmetros físico-químicos e o aporte instantâneo de nutrientes no estuário.

MATERIAIS E MÉTODOS

Os dados utilizados como entrada para o modelo foram as séries temporais da vazão, pH, temperatura, condutividade, oxigênio dissolvido e material particulado em suspensão e, para saída, nitrato e fosfato.

O tratamento dos dados foi realizado no ambiente estatístico R (CRAN, 2010), utilizando pacotes (nnet, cluster) desenvolvidos pela comunidade de usuários.

Do conjunto de dados 75% foram aleatoriamente escolhidos para treinamento e os demais 25% reservados para teste do modelo.

RESULTADOS E DISCUSSÕES

Os resultados encontrados com a abordagem fuzzy e por redes neurais foram de certa forma equivalentes, com o coeficiente de correlação entre os resultados dos modelos e os valores observados em torno de 60%. Como a resposta do modelo foi semelhante para os dois nutrientes, optamos por apresentar os dados de nitrato.

Na Fig. 1 são apresentados os resultados obtidos para o nitrato através da aplicação de uma rede neural tipo perceptron com 3 neurônios na camada escondida. No topo observa-se o gráfico comparativo entre os valores observados e os resultados do modelo para o conjunto de teste, enquanto na parte de baixo da figura é apresentado o comparativo para a série completa.

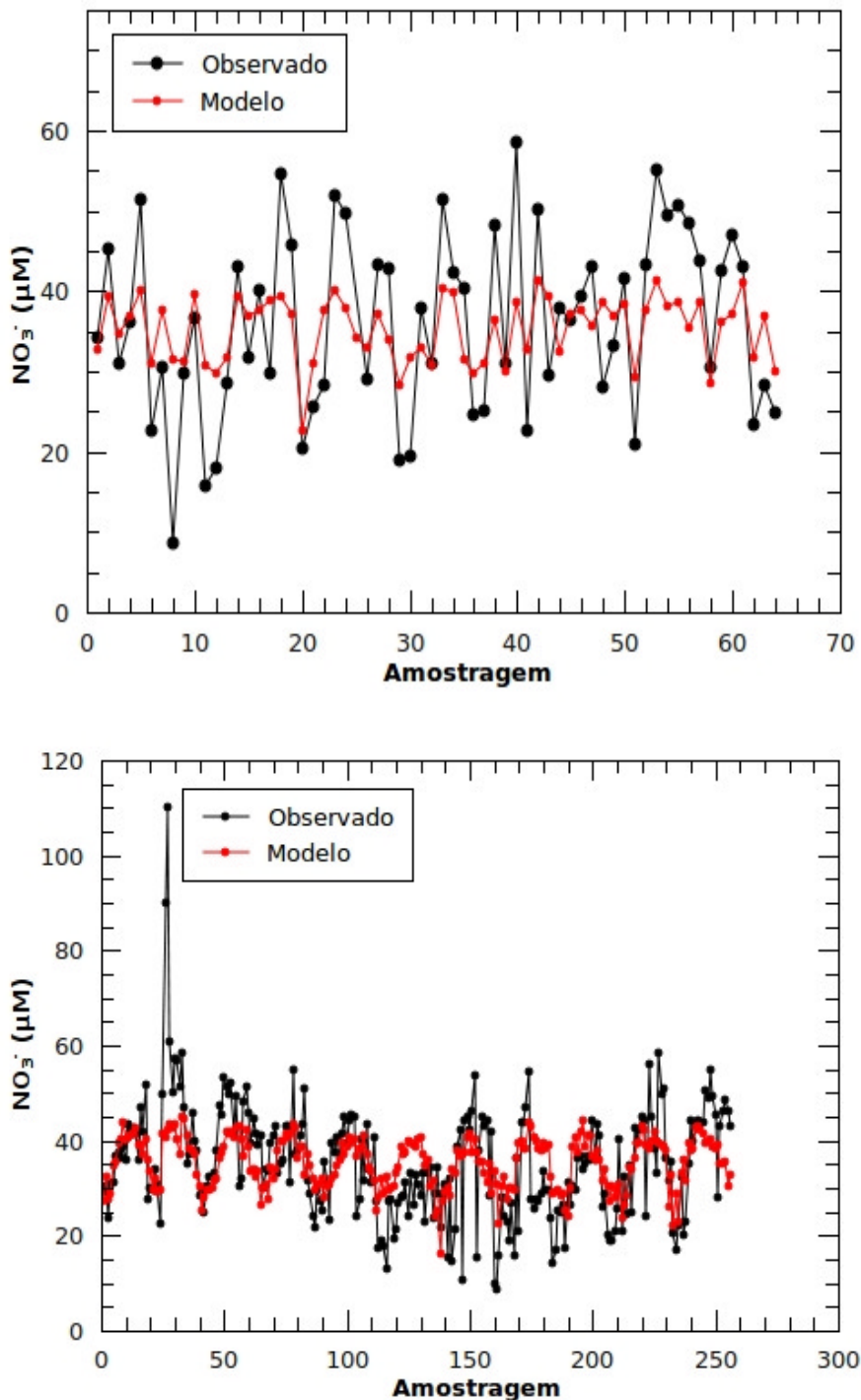


Figura 1: Resultados obtidos para a série temporal do nitrato com a utilização de uma rede neural tipo perceptron em 3 camadas. No topo é apresentado o resultado para o conjunto de teste e abaixo para a série como um todo.

A resposta do modelo representa de forma geral a tendência de longo prazo da concentração de nitrato nas águas do Rio Paraíba do Sul (p.ex. a exemplo da redução dos valores mínimos observada entre as amostragens 100 e 200). Por outro lado, a resposta do

modelo a eventos episódicos e transientes de aumento ou diminuição de concentração não é bem ajustada (p. ex. picos de concentração nas amostragens 30, 150 e 220). Essa característica é interessante, posto que representa uma certa prevalência da tendência geral da série temporal em relação a eventos episódicos.

AGRADECIMENTOS

Este trabalho contou com apoio financeiro do projeto Instituto do Milênio Estuários - (CNPq Proc. 420050/2005-1)

CONCLUSÕES

A utilização de técnicas de mineração de dados a séries temporais de indicadores da qualidade da água do Rio Paraíba do Sul demonstrou a viabilidade de se modelar o aporte instantâneo de nutrientes, cuja determinação demanda análise em laboratório, a partir de parâmetros físico-químicos, de determinação imediata. Deve-se ressaltar, no entanto, que, a qualidade dos resultados, mais do que à metodologia analítica empregada, está vinculada à consistência da série temporal considerada.

REFERÊNCIAS

CRAN 2010. <http://cran.r-project.org>, acessado em 29/01/2010.

MAIMON, O. e ROKACH, L. 2008. *Soft Computing for Knowledge Discovery and Data Mining*. Springer, Heidelberg. 437pp.