Fault detection in a real-time monitoring network for water quality in the lagoon of Venice (Italy)

S. Ciavatta*, R. Pastres*, Z. Lin, M.B. Beck**, C. Badetti*** and G. Ferrari*****

* Department of Physical Chemistry, University of Venice, 2137 Dorsoduro, 30123 Venice, Italy (E-mail: *ciavatta@unive.it*; *pastres@unive.it*)

** Warnell School of Forest Resources, University of Georgia, Athens, Georgia 30602, USA (E-mail: *linzhulu@uga.edu*; *mbbeck@uga.edu*)

*** Sezione Antinquinamento, Magistrato alle Acque di Venezia, 19 San Polo-Rialto, 30125 Venice, Italy (E-mail: *cristian.badetti@magisacque.it*; *ferrari@magisacque.it*)

Abstract In the context of monitoring water quality in natural ecosystems in real time, on-line data quality control is a very important issue for effective system surveillance and for optimizing maintenance of the monitoring network. This paper presents some applications of recursive state-parameter estimation algorithms to real-time detection of signal drift in high-frequency observations. Two continuous-discrete recursive estimation schemes, namely the Extended Kalman Filter and the Recursive Prediction Error algorithm, were applied to assuring the quality of the dissolved oxygen (DO) time series, as obtained from the Lagoon of Venice (Italy) during August 2002, through the real-time monitoring network of the Magistrato alle Acque (the Venice Water Authority). Results demonstrate the effectiveness of the methodology in early detection of a probable drift in the DO signal. Comparison of these results with those obtained from the application of a related recursive scheme (a Dynamic Linear Regression procedure) suggests the strong benefits of approaching the problem of on-line data quality control with several (not merely a single) independent such estimation methods.

Keywords Data quality analysis; fault detection; lagoon of Venice; recursive estimation; real-time monitoring; water quality models

Water Science and Technology **Vol 50 No 11 pp 51–58** © IWA Publishing 2004

Introduction

Real-time observing networks are becoming ever more widely used in monitoring water quality in natural environments. They provide large volumes of high-frequency data to serve a variety of purposes, such as, for example, water pollution control, detection of harmful algal blooms, and commercial aquaculture activities.

Nevertheless, the real-time data must be transformed into information for effective surveillance and control of the processes occurring in the environment. This is a difficult challenge, because of the high sampling intensity enabled by automated instruments and the multivariable character of the data sets, which requires interpretation of the interactions among many contemporaneously measured quantities (Beck and Lin, 2003). Conceptual or "mechanistic" models, when embedded in recursive estimation algorithms, can be advantageously applied to processing the data on-line, in real time (Beck, 1987; Young, 1998; Beck and Lin, 2003; Lin, 2003; Pastres *et al*., 2003). Such schemes may be referred to as "data assimilation" methods, as commonly found in the subjects of oceanography and climatology.

Data collected by automated instruments, just like any other means of observation, may be not reliable in certain circumstances, for instance, as a consequence of improper instrumental calibrations, various kinds of abrupt failure, and more gradual impairment of the sensors. This last source of error may often be due to dirt or bio-fouling, which can attach to