Linear Algebra and Minimum Relative Entropy to Investigate Contamination Events in Drinking Water Systems

Marco Propato[1]; Fanny Sarrazy[2]; and Michael Tryby[3]

Abstract: A two-step approach is proposed to assist forensic investigation of possible source locations following a contaminant detection in drinking water systems. Typically this identification problem is ill posed as it has more unknowns than observations. First, linear algebra is employed to rule out potential contaminant injections. Second, an entropic-based Bayesian inversion technique, the minimum relative entropy method, solves for the remaining variables. This formulation allows for the less committed prior distribution with respect to unknown information and can include model uncertainties and measurement errors. The solution is a space-time contaminant concentration probability density function accounting for the various possible injections that may be the cause of the observed data. Besides, a probability measure quantifying the odds of being the actual location of contamination is assigned to each potential source. Effectiveness and features of the method are studied on two example networks.

DOI: 10.1061/(ASCE)WR.1943-5452.0000059

CE Database subject headings: Water distribution systems; Water quality; Water pollution; Algorithms; Entropy methods; Drinking water.

Author keywords: Water distribution; Water quality; Water pollution; Pollution sources; Algorithms.

Introduction

The protection of consumers from contamination events has always been a concern in drinking water distribution system (DWDS) management. Routine monitoring plans, disinfectant residual maintenance and backflow prevention devices are very helpful in limiting consumer exposure risks and are probably sufficient in controlling chronic, low intensity water quality problems. Consumers are at risk of exposure, however, from accidental or malicious events that may occur suddenly and with a high intensity of contamination.

Although technological limits still exist, recent progress in the development of water quality sensor devices may soon allow a continuous water quality monitoring system as an additional barrier against the spread of a contaminant. Several approaches have been proposed to design sensor networks to provide early warning in order to prompt the appropriate response strategy to limit and mitigate consumer exposure (Propato 2006; Berry et al. 2006; Ostfeld et al. 2006; Watson et al. 2004; Ostfeld and Salomons 2004; Uber et al. 2004). Currently, there is a growing interest in investigating methodologies that process the information obtained by sensors to identify the locations from which the contamination might have started. Such capability would be very important for emergency response strategies (e.g., by isolating only part of the network rather than affecting the entire DWDS), or more important to better investigate, and possibly eliminate once and for all, the cause of contamination. A better understanding of this problem could also be helpful in the optimal design of routine monitoring sampling plans that are required to periodically characterize water quality conditions in DWDS.

The main challenges of the contaminant source identification problem are (1) inherent nonuniqueness of the solution due to limited sensor data available compared to the large number of potential contaminant source locations in a real DWDS and to hydraulic properties; (2) measurement errors and model uncertainties; and (3) computational effort that significantly increases with network size. On the contrary, one mathematical feature that may help finding the solution and ease computation is that a *linear* input/output (I/O) model for water quality can be written when contaminant is conservative or first order reactive. Recently considered contamination source identification models include linear least-squares optimization (Laird et al. 2006), online continuous parameter estimation (Guan et al. 2006) and hybrid model trees-linear programming (Preis and Ostfeld 2006). Each of these models gives a yes/no answer if a location is or is not a contamination source. Yet, they all recognize that in general these solutions are not unique.

The objectives of this manuscript are (1) to present an entropic Bayesian stochastic method for inversion, the minimum relative entropy (MRE), to assist the forensic investigation following a contaminant detection at water quality sampling locations; (2) to show that with linear algebra several potential contamination sources can be ruled out, improving computational efficiency and system invertibility; (3) to define a probability measure assigned to each potential source quantifying the odds of that node being the actual location of contamination; and (4) to study effective-

[1]Research Engineer, Networks, Water Treatment, and Water Quality Research Unit, CEMAGREF, 50 Ave. de Verdun-Gazinet, 33612 Cestas, France (corresponding author). E-mail: marco.propato@cemagref.fr

[2]Research Engineer, Networks, Water Treatment, and Water Quality Research Unit, CEMAGREF, 50 Ave. de Verdun-Gazinet, 33612 Cestas, France. E-mail: fanny.sarrazy@cea.fr

[3]Environmental Engineer, U.S. EPA Ecosystems Research Div., 960 College Station Rd., Athens, GA 30605. E-mail: tryby.michael@epa.gov

Note. This manuscript was submitted on April 22, 2009; approved on November 1, 2009; published online on November 11, 2009. Discussion period open until December 1, 2010; separate discussions must be submitted for individual papers. This paper is part of the ***Journal of Water Resources Planning and Management***, Vol. 136, No. 4, July 1, 2010. ©ASCE, ISSN 0733-9496/2010/4-483–492/$25.00.