Centinela: an early warning system for the water quality of the Cauca River

Carlos Vélez, Leonardo Alfonso, Arlex Sánchez, Alberto Galvis and
Gilberto Sepúlveda

ABSTRACT

The Cauca River is the drinking water source for 1.3 million inhabitants of the city of Cali, Colombia. Although the river discharge is sufficient to handle the water demand of the city all year long, significant water pollution events cause frequent disruption to the Puerto Mallarino Treatment Plant (PMTP) and the water supply service, with substantial social and economic impacts on the city. The sources of pollution include wastewater discharges upstream of the PMTP and important sediment transport from the upstream sub-catchments during heavy rainfall events. Both situations can lead to a closure of the PMTP when the presence of a pollution plume at its intake is evident. This paper presents the design and prototype of a water quality early warning system to anticipate the peaks of pollution in the river, in order to assist the operators in taking timely informed decisions about the operation of the treatment plant. As the published experiences of early warning systems for similar water pollution problems are very limited, the approach to solve the problem using hydroinformatics technologies is worth documenting for utility companies with a similar problem.

Key words | anticipatory water management, early warning systems, pollution control, water quality, water supply, water treatment

Carlos Vélez (corresponding author)
FlowNet, Voorstraat 225A, 3311 PE,
Dordrecht,
The Netherlands
E-mail: *carlos_velez@yahoo.com*;
carlos.velez@anteagroup.com

Leonardo Alfonso
Arlex Sánchez
Hydroinformatics Chair,
Department of Integrated Water Systems and
Governance, UNESCO-IHE,
Westvest 7 2611 AX, Delft,
The Netherlands

Alberto Galvis
Instituto Cinara – Universidad del Valle,
Calle 13 # 100-00, Cali,
Colombia

Gilberto Sepúlveda
Empresas Municipales de Cali (EMCALI),
Cali,
Colombia

INTRODUCTION

Drinking water sources must maintain certain quality standards in order for treatment systems to guarantee the delivery of safe drinking water to users. Treatment processes are designed to be operated with source waters that may fluctuate within certain water quality limits. However, if the characteristics of the source suddenly exceed these limits due to occasional contamination events upstream of the intake, the risk of supplying water that deviates from the standards increases. In this situation, operators have two options: either modify the operation of the treatment (e.g. add more chemicals) or close the intake for a period of time and resume operation once the pollution plume overtakes the treatment plant. In both operational scenarios the treatment process is disrupted. In the former, the risk of producing poor water quality is increased and, as a consequence, the risk of outbreak of water-related diseases is

increased. In the latter, the lack of water supply may have negative effects in the distribution system, causing complaints from users and loss of revenue for the utility company, and other issues related to the supply intermittency such as contamination intrusion, back siphonage, etc.

Early warning systems (EWS) have been recognized as important tools to reduce the negative effects of contamination events of water sources for water supply systems (Grayman *et al.* 2001). The International Strategy for Disaster Reduction of the United Nations defines early warning as 'the provision of timely and effective information, through identified institutions, that allows individuals exposed to a hazard to take action to avoid or reduce their risk and prepare for effective response' (UN-ISDR 2004).

EWS may play an important role by identifying critical contamination events in real time and generating

information for operators and the community, so that informed decisions can be made (USEPA 2005). The analysis made by Grayman *et al.* (2001) on data from 153 water supply systems from the USA, Canada and the UK shows that 90% of these utility companies consider EWS as an important tool. However, the survey also shows that less than 50% of the plants have any type of EWS in place. In terms of the warning time that operators have to respond to contamination events, only 16% reported that they always have sufficient time, 56% reported insufficient warning time in half of the critical events and 23% do not have a warning signal at all. This situation has been changing since 2001, especially in the USA where the threat of intentional contamination of water supply systems has forced authorities and utility companies to increase the water security by developing vulnerability assessments and emergency response plans (Roberson & Morley 2005).

In recent years, significant attention has been given to EWS in the context of water-related disasters such as tsunamis (Hadihardaja *et al.* 2010), flooding (Abebe & Price 2005; Krzhizhanovskaya *et al.* 2011; Miyamoto *et al.* 2012; Pyayt *et al.* 2012) and debris flows (e.g, Romang *et al.* 2011), because of the human casualties and the economic losses caused by these events. In contrast, less attention has been given to EWS in the context of pollution of drinking water sources.

In EWS for water supply systems, considerable research has been concentrated on the development of monitoring indicators. Some examples are: Balk *et al.* (1994) on biological indicators; Tryland *et al.* (2002) on faecal contamination; and Izydorczyk *et al.* (2005) on cyanobacteria in reservoir waters. Other studies have concentrated on the use of hydrodynamic models to identify possible locations for early warning monitoring stations (Kawka & Cousson 2012). Latest research has focussed on the detection of intentional contamination in the distribution systems rather than transient pollution in the water source. This is perhaps because distribution systems are considered more vulnerable than water sources (USEPA 2007). Authors of comprehensive reviews on EWS have concluded that more research is needed on strategies for design and implementation of EWS. For instance, research is needed on sensor selection and location, definition of triggers that may lead to a generation of warnings and how to translate data into information

useful for decision makers (USEPA 2005; Quansah *et al.* 2010; UNEP 2012).

This paper contributes to the research in the area of designing EWS for water sources used for water supply systems. It is focussed on detecting transient pollution and generating information useful for the operation of water treatment plants (WTP). A real case study is used to present methods and approaches to design an early warning system for the water quality of the Cauca River that is the main source of water for the city of Cali (Colombia). The design includes the evaluation of sensors and monitoring stations, data processing, modelling, and communication to support decision-making. Existing and customized hydroinformatics tools were used to process data and develop a web-based prototype named Centinela. This paper corresponds to an extended version of the paper presented at the 10th International Conference of Hydroinformatics (Vélez *et al.* 2012).

EWS FOR WATER SOURCES

According to the Inter-Agency Secretariat of the International Strategy for Disaster Reduction (UN 2006), a complete and effective early warning system comprises four inter-related elements: risk knowledge, monitoring and warning service, dissemination and communication, and response capability. Even though these elements are associated with disaster reduction, we used them here to schematize the main components of an EWS for water sources of treatment plants.

Risk knowledge

Risks arise from the combination of hazards and vulnerabilities at a particular location. Assessments of risk requires systematic collection and analysis of data and should consider the dynamic nature of hazards and vulnerabilities that arise from processes such as urbanization, rural land-use change, environmental degradation and climate change (UN 2006). An interesting initiative is the US Source Water Protection Plans, which focused on diminishing pollution coming from agriculture on source water by implementing improved voluntary practices at a local level. In the case of WTP, the hazards on the catchment of

the water source should be assessed. The vulnerability is associated with the constraints of the WTP to deal with the peaks of pollution and the flexibility to adjust the processes to those fluctuations in the water quality.

Monitoring and warning service

The monitoring equipment is the basis of this component. The sensor must have the capability to detect the presence and intensity of hazardous contaminants in the water sources. In addition, methods should be developed to assess the data of detected contaminants, determining the nature of the contamination event, the concentration of the pollutants and the time that they may affect the water at critical points, like the intake of a treatment plant (Grayman *et al.* 2000). Mathematical models may be developed to predict the movement of the contaminant in the water body.

The warning services are the core of the EWS. There must be a sound scientific tool for the generation of reliable forecasts and warnings. Continuous monitoring of hazardous contaminants is essential to generate accurate warnings in a timely fashion (UN 2006). In WTP, warning services for different contaminants should be available.

In general, the monitoring and warning service should have the following functions (USEPA 2005): provide a rapid response; screen for a number of contaminants while maintaining sufficient sensitivity; operate as an automated system that allows remote monitoring and operation; reduce the number of incorrect warnings to the minimum (false-positives/false-negatives); operate continuously.

Dissemination and communication

This component requires the use of communication infrastructure to disseminate the information to the target groups (e.g. the utility's operational and managerial staff). A careful selection of target groups and level of information transmitted must be carried out. Clear messages containing simple, useful information are critical to enable proper responses that will help decision making. The use of multiple communication channels is necessary to ensure that all the people who must be warned are reached, to avoid failure of any channel, and to reinforce the warning message (Grayman *et al.* 2000).

Response capability

Decision makers must understand warning signals and know how to react to them and therefore education and preparedness programmes play a key role (UN 2006). Options for operating treatment plants during a critical pollution event can include closure of intakes, use of alternative intakes and modification of treatment processes. The presence of varying amounts of raw water storage, finished water storage, and alternative water sources influences the possible actions of the water utility during a contaminant event (Grayman *et al.* 2000).

In general, an EWS for the water source should be conceived as an integral part of the water supply system. To become a widely used, effective, and reliable part of the system it should demonstrate the characteristics and components presented above. In the next section the design and prototype of a Water Quality Early Warning System (WQ-EWS) to anticipate the peaks of pollution of the Cauca River in Colombia is presented, following the components described above, but concentrating mainly on the monitoring and warning service element.

DESCRIPTION AND METHODOLOGY FOR THE CASE STUDY OF CALI

Cali is the third biggest city in Colombia, with a population of 2,075,380 inhabitants in 2005 (DANE 2012). The main source of water for the city is the Cauca River, which supplies 77% of the demand of the distribution system. The Cauca River is the second largest river in Colombia and is used for different purposes: as a source for drinking water, as a waterway for transportation of goods, for recreation, irrigation, a water source for industry, generation of electricity, extraction of sand and gravel and as a final receiving system of both treated and untreated wastewater. Figure 1 shows the Cauca River crossing the city of Cali from south to north. The map also shows seven tributary rivers that cross the city. These tributaries also serve as drainage of part of the city, which includes some combined sewer overflows.
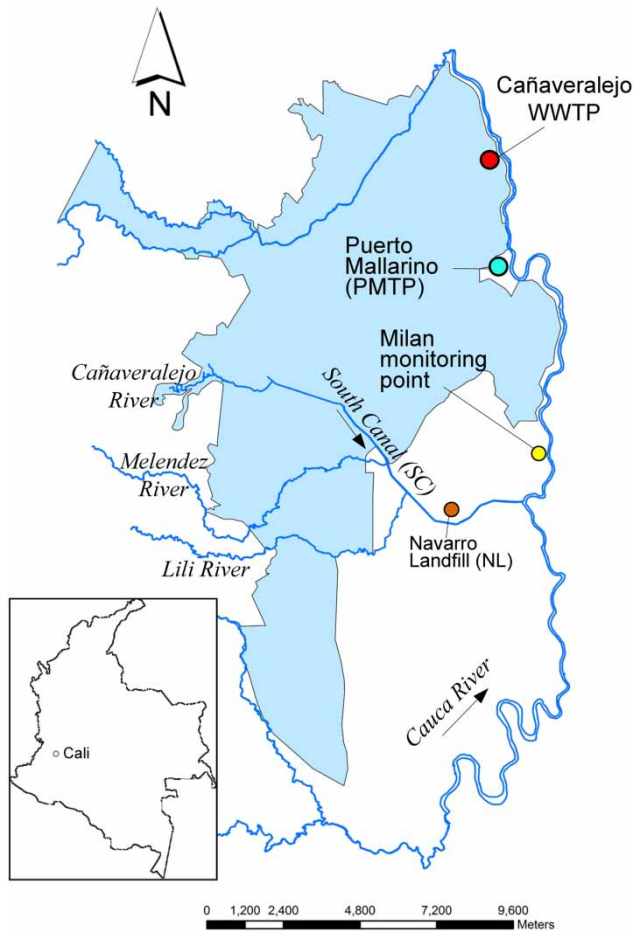
Figure 1 │ Map of the main rivers and drainage system of Cali, Colombia. Shaded area indicates the urban boundary.

Water quality-related risk in Cali, Colombia

As discussed above, risk arises from the combination of hazards and vulnerabilities. For the particular case of Cali, the hazard is associated with pollution events that deteriorate the water quality of the Cauca River, while the vulnerability is associated with the impact of the WTP operation on the population. Two possible operational alternatives are distinguished under a pollution event: to close the intake of the treatment plant or to let the highly polluted water pass through the system. The first alternative implies that the supply of water to the city is interrupted with the associated negative impacts on economy, and the associated unpressurized network problems. The second alternative could lead to an increment in the risk to deliver inadequately treated water, with the associated impacts on public health.

Hazard

The main sources of transient pollution in the region are related to wet weather events. They are a consequence of the deterioration of catchments and the discharge of untreated wastewater and combined sewer overflows in the natural drainage systems (Galvis *et al.* 2001). Acute pollution may also be caused by industrial spills but, if they exist, they are not well identified.

One of the most critical sources of transient pollution is the South Canal (SC). This manmade channel is used to drain three rivers: the Cañaveralejo, Melendez and Lili (Figure 1). The upper catchment of these rivers is affected by coal mining and deforestation. In addition, they together with the SC are the final receiving systems of urban drainage in the southern part of Cali. The impacts of the wash-off of sediments from the canal on the water quality of the Cauca River are associated with reduction in the amount of dissolved oxygen (DO) in the river and with peaks in turbidity. The utility company Empresas Municipales de Cali EICE ESP (EMCALI) has records of transient events caused by this tributary. A well-documented example of reduction of concentration of DO from 5 to 1 mg L^{-1} in the Cauca River as a consequence of wash-off from the SC is presented in Vélez *et al.* (2006). The impacts of the SC discharge into the Cauca River are important because of its proximity (11 km) to the intake of the water treatment plant (PMTP) named Puerto Mallarino (Figure 1), although other tributaries upstream of SC may also have significant effects on transient pollution.

The critical pollution caused by transient events has a negative impact on the water quality characteristics of the Cauca River. The level of pollution can be so high that it surpasses the limits of treatability of the existing treatment processes of the PMTP. One of the concerns of the operators is related to the deterioration of the treatment processes (sludge blanket and sand filtration) due to the high content of total suspended sediments (Montoya *et al.* 2011). The other concern is related to the generation of by-products after chlorination due to the peaks in organic matter present during acute pollution events.

According to information collected by the EMCALI, the operation of the PMTP was stopped 28 times in 2010. On nearly half of these occasions the intake was closed due to concentrations of DO of about 2 mg L^{-1}, as an indicator of high organic matter pollution in the river. The others were caused by the high concentration of turbidity (above 1,000 NTU) used as an indicator for the content of suspended sediments in the water.

Vulnerability

Figure 2 shows the number of times that the PMTP has been shut down in recent years. The figure also shows the reason for closing the intake: namely high turbidity or high organic pollution load measured as low DO concentration (EMCALI 2009).

On average, the shutdowns of the PMTP last for 2 hr. During this time no water is pumped into the treatment system, so the distribution relies entirely on the volume of water stored in the city's tanks. Depending on the time of the day in which the event occurs (i.e. during peak demands) the situation may be worsened and the hydraulic capacity of the distribution system may take more time to recover. The population in general, as well as the industrial and commercial clients, are affected for several hours before normal operation is reached (Univalle 2008). The disruption in the water supply system creates

discomfort for the users and loss of revenue for EMCALI. In addition, the lack of water in the system could increase the chances of re-contamination in the distribution network and affect the health of the population (Madera *et al.* 2007).

In the last 10 years, EMCALI has implemented several actions to reduce the effects of transient pollution. A master plan to collect all untreated wastewater that affects the SC has been implemented, including sediment traps in the tributaries and in the SC. A reservoir of 80,000 m^3 to store pre-treated water for emergency situations was built in 2010. The reservoir can be used as an alternative source for about 6 hr. In addition a set of sensors to collect on-line information from the Cauca River has been implemented. The monitoring system includes sensors in the Cauca River at the intake of the PMTP and at a place 8 km upstream named Milan (Figure 1). The system allows for real-time data collection of temperature, conductivity, turbidity and DO. Data are stored in a database and analysed after high pollution events. Basic statistics are estimated and information is available for the operator upon request.

However, despite these measurements the data collected by the monitoring system missed the analysis required to provide information to take decisions at the operational level. For these reasons, EMCALI needed a decision support system (DSS) to process the data and generate another level of information at the right time for
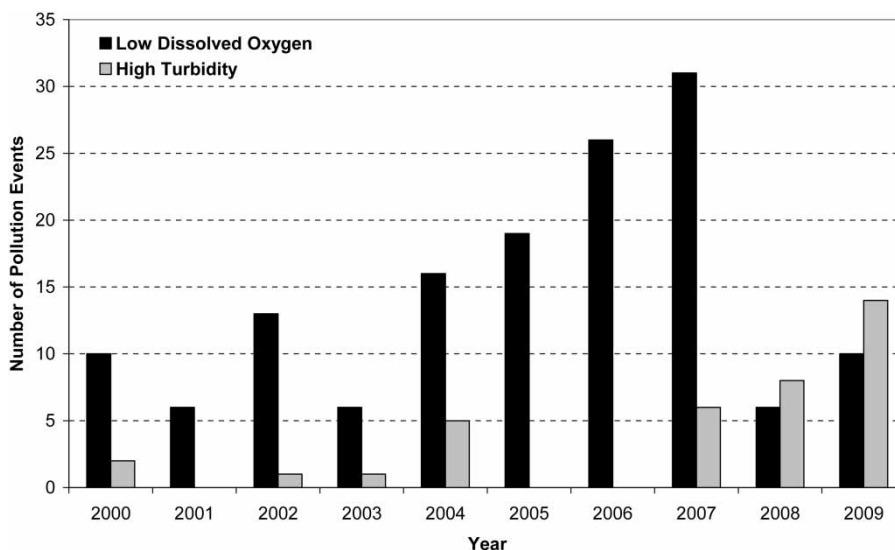

Figure 2 │ Number of events and type of pollution that have triggered the closure of PMTP.

decision makers involved in the operation of the PMTP. The need for better information to take decisions in real-time motivated EMCALI to design an Early Warning System for the water quality of the Cauca River. The remainder of this paper describes the methodology used for the design of the EWS and the results obtained for each of the components.

Evaluation of the existing monitoring system

Monitoring systems are very important because they provide data that can be processed to generate information that support decisions. Although there are many ways to evaluate monitoring networks with multiple sensors (e.g., Mishra & Coulibaly 2009; Alfonso *et al.* 2013), for the case of Cali we concentrate only on the location of one water sensor. Before 2005, EMCALI only monitored the water quality variables for operational purposes and inside the PMTP facility. Because of the frequent pollution events in the Cauca River, in 2006 EMCALI deployed two monitoring stations in the river: one at the intake and the other at Milan (8 km upstream of the intake). It is important to note that there are no discharges coming from industries, cities or tributaries between these two sensors.

The location of the sensor in the water body is important. Rivers may not be completely mixed either laterally or vertically, and discharges can hug one bank or the other for large distances (Grayman *et al.* 2001). Therefore, the evaluation focuses on the validity of the measurements at Milan. It was of special interest to establish if the measurements at Milan were representative of the cross section and if the discharge from the SC was completely mixed at this point. Two tests were carried out: a monitoring campaign at Milan to analyse the distribution of the concentrations in the cross section and a tracer study to identify the complete mixture of a plume coming from the SC. For the tracer experiment, a solution of rhodamine was discharged at the mouth of the SC and the concentration was monitored in the Cauca River at three cross sections approximately 1, 2 and 4 km downstream of the SC.

Using the methodology proposed by Constain *et al.* (2008), we estimate the mixing length of a discharge from the SC as approximately 1 km. Therefore Milan, which is located about 2.5 km from the mouth of the SC, properly

represented the conditions of the water quality in the river at that point.

Objectives and specifications

The main objective is to use data for the Cauca River collected upstream of the intake to generate early warnings related to critical pollution events. The system should facilitate the decision-making for operation of the PMTP when it is threatened by pollution events associated with high organic matter and/or suspended solids. The EWS is named 'Centinela' (the Spanish word for sentinel). Centinela should: perform real-time data acquisition, validation and analysis; forecast the water quality trends at the PMTP intake; generate early warnings and warn users by using sound and visual signals and text messages.

Methodology

In general, the methodology includes acquisition and analysis of existing data, fieldwork and laboratory analysis, use of existing modelling tools for the river, development of process-based and data-driven models and prototyping a web-based application as user interface. Although these steps were considered in Centinela, we concentrate here on the data and methods used in the components of monitoring, warning and communication. To this end, the next section presents the analyses, results and discussion of the characterization of the pollution events, data acquisition and validation, definition of warning signal and lead time, the prediction of the pollution travel time, the forecast of DO concentrations at PMTP intake and the communication of warnings.

ANALYSIS AND DISCUSSION OF WQ-EWS COMPONENTS

Characterization of pollution events

The critical pollution events of the Cauca River are associated with high organic matter and/or suspended solids. As mentioned above, DO and turbidity have been used as surrogate variables to characterize those two types of pollution

events. Conventional sensors for DO and turbidity are relatively inexpensive, widely available and easily used. Therefore, there was no doubt whether they should be used to identify the presence of contaminants in the river.

Figure 3 shows an example of a pollution event associated with high contamination of the river with organic matter compounds. Measurements at Milan and at the intake of the PMTP are shown. The figure also includes the time when the PMTP operator decides to stop the treatment process and restarts after the event has passed. Analysis of these data helps to identify key characteristics of the event, for instance: the trend of the concentration of the variable, the concentration that triggers the closure of the intake (about 2 mg L^{-1} O_2), the minimum value reached at the intake (1 mg L^{-1} O_2), the duration of the critical concentration level (1.5 hr) and the travel time of the pollution from Milan to the intake (3 hr). From the 69 pollution events, it can be reported that the average event duration (from stop to start in Figure 3) is 2.13 hr with a standard deviation of 2.34 hr, with a maximum of 6.67 hr and a minimum of 0.75 hr. For the travel time, we have information of 12 events that were recorded simultaneously at Milan and PMTP with the following statistics: average = 2.91 hr, min = 2.25 hr, max = 4.0 hr and SD = 34 min. All these characteristics of the event are used to develop the warning system.

Data acquisition and validation

One of the challenges of a continuous, real-time monitoring system is management of the large amounts of data that are generated. Use of data acquisition software and a central data management centre is critical (USEPA 2005). EMCALI has a data acquisition system in place that includes remote sensors and data loggers for on-line monitoring at Milan and PMTP, telemetry to transfer the data to the central information system and an Oracle database in which to store the data. Routines for data validation are part of the data acquisition system.

When data acquisition systems are in place, the EWS may take advantage of exiting routines for data validation. Data processing includes data analysis and trending in order to assess whether an alarm level has been exceeded or whether it is faulty data. The EWS designed for the Cauca River, which uses the existing EMCALI database as the source for the input data, incorporates such procedures and some of them are described in this manuscript. This input is validated to filter out outliers and to check for missing data. Outliers are detected by comparing the input data with the physical limits of the measured water quality variables and with the detection range of the sensor used, among other statistical techniques, while missing values are estimated by linear interpolation using the past n records
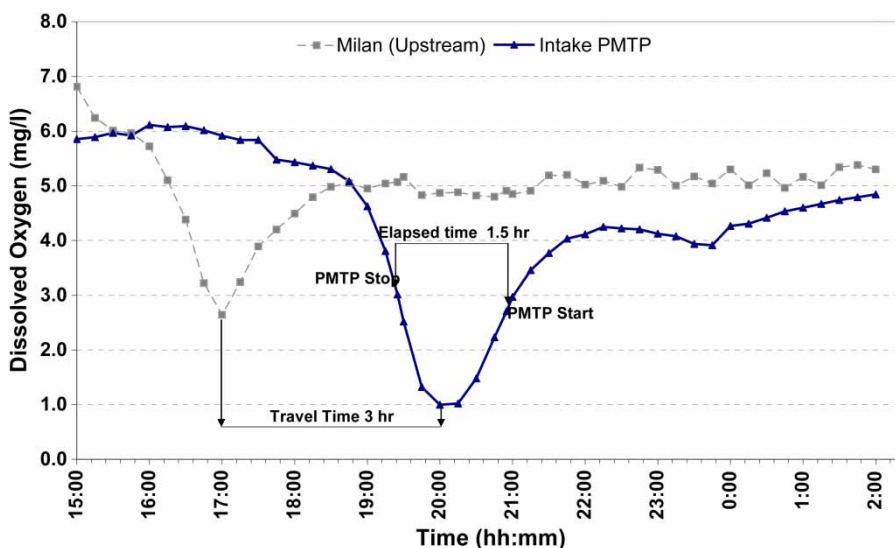

Figure 3 │ Typical DO event that triggers the closure of the PMTP.

if at a particular time t no input data are detected. The validated data are displayed and stored in the Centinela Database, to be used in the future to improve the characterizations and fine-tune the parameters.

Definition of the warning signal

A first activity towards the development of the warning signal consists of the statistical analysis of historical data collected by EMCALI. Data for 69 critical events, associated with low DO concentrations measured at the intake of the PMTP, were included in the analysis. The events occurred between August 2006 and August 2009. The objective was to define the relevant parameters that could be used to identify the starting point of a potential critical event. An equivalent analysis was carried out for turbidity but it is not included due to space restrictions.

The analysis concluded that the most suitable parameters to characterize an event are the following (Figure 4). (1) Average of the last four records of concentration, C_{ave}. This parameter indicates the state of concentrations in the river. (2) Average of the slopes within the last four records, S_{ave}. This is a measure of how fast the concentration of the water quality indicator falls or rises. The slope parameter indicates the gradient at which the water quality variable may become critical. (3) Gradient continuity, F, to detect a systematic tendency of the records. For instance, if all the negative slopes of the
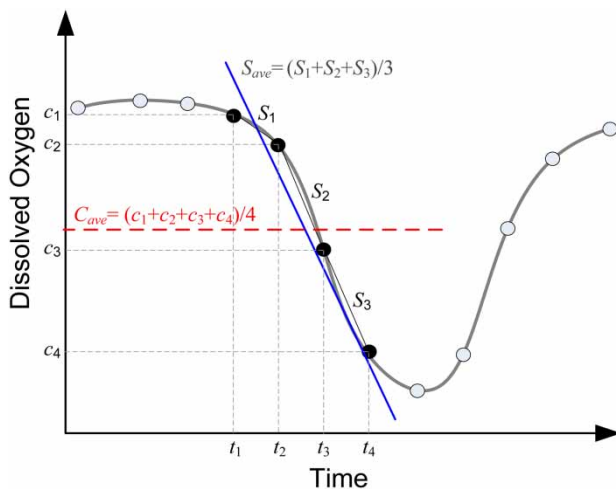
last four records point to the same direction then $F = 1$; otherwise, $F = 0$.

In order to understand how the parameters C_{ave} and S_{ave} change during a contamination event, Figure 5 was prepared. The graph is the result of a data processing of the 69 DO events recorded at PMTP. The records include, in addition to DO value at each time step, information about the time at which the closure of the intake took place. In this way, the analysis focused on knowing the characteristics of different events, encapsulated in the parameters C_{ave} and S_{ave} at the time of closure and also at other times before the closure (up to 3 h before). Data used for this figure correspond to the descending part of the event characterized by reduction of DO in the river (i.e. $F = 1$). It can be observed that from half an hour before the closure of the intake C_{ave} varies between 2.5 and 4.3 mg L^{-1}, and S_{ave} varies between -3.5 and -6.0 mg L^{-1} h^{-1}, which give an important indication of the possible values that trigger a warning. It can also be noted that there is no evident pattern to allow for detecting a potential event with one hour or more before the closure of the intake.

Definition of warning levels

A warning level, $N(t)$, is defined as the degree of significance of a pollution event at a particular time t, and it is represented by a value that ranges from 1 (critical event) to 8 (normal conditions). The warning level is assigned by the combination of the parameters C_{ave} and S_{ave} and F that
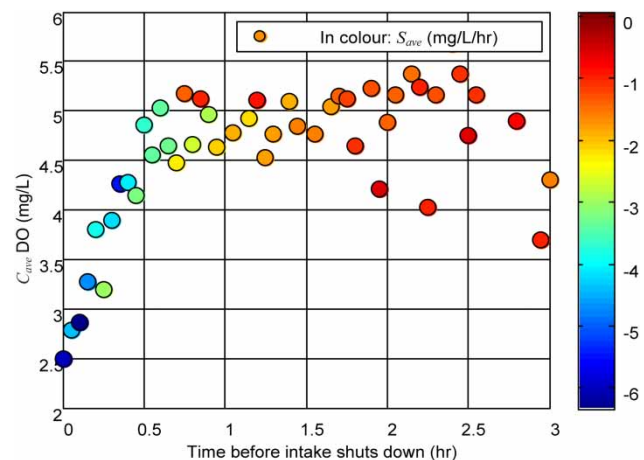

Figure 4 │ Set of parameters to characterize a critical DO event in the Cauca River. For this example $F = 1$ because the diminishing tendency is maintained.

Figure 5 │ Variation of parameters C_{ave} and S_{ave} for DO before the closure of the intake.
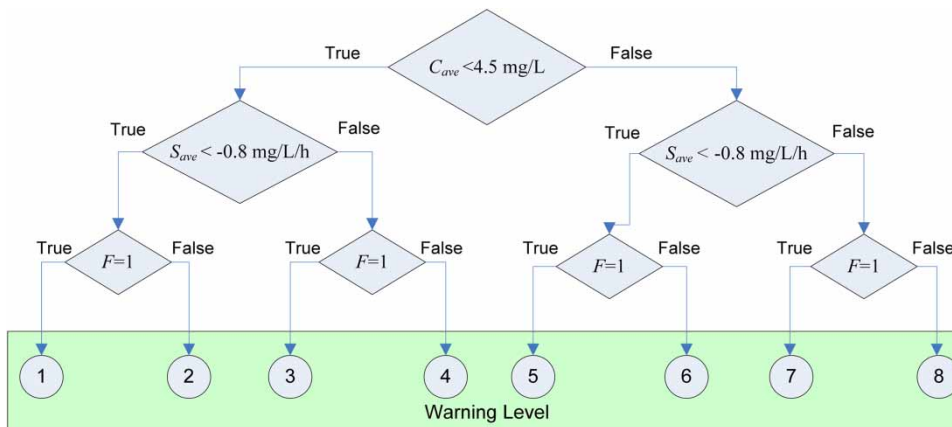
Figure 6 | Decision tree for the generation of warning levels associated with DO events. C_{ave} is the average DO concentration, S_{ave} the average slopes and F is the continuity in the slope tendency ($F = 1$ continuous, $F = 0$ otherwise), all for the last four records.

are checked with a series of nested conditionals, and for the case of DO is presented in Figure 6. It must be noted that the warning levels might not change consecutively. For instance, N can be five at $t = t_0$ (S_{ave} is critical but C_{ave} is not) to $N = 1$ at $t = t_0 + \Delta t$ (both S_{ave} and C_{ave} are critical). The point here is that a level 5 warning does not mean 'no threat' but it means 'check C_{ave} closely'. In addition, warning levels of $N = \{1,2,3,4\}$ are, in principle, the criteria to identify a potential alarm situation, but levels $N = \{1,2\}$ represent imminent intake closure. Each of these levels has an associated alarm message and recipient, as discussed further.

The definition of the thresholds for the parameters S_{ave} and C_{ave} is obtained by sensitivity analysis exploring several values for S_{ave} and C_{ave} and checking their impact on the number of false positives and false negatives. This is done by closely looking at the performance of the EWS through contingency tables, such as Table 1 (Doswell *et al.* 1990; Verkade & Werner 2011). In particular, false alarm rate (FAR), defined as the proportion of false alarms with respect to the total released alarms, and probability of detection (POD), defined as the proportion of hits with respect to

the total observed events (Doswell *et al.* 1990), are indexes commonly used to assess EWS.

There are two conflicting aspects that ideally should be avoided. (1) If the conditions in the decision tree become difficult to fulfil (e.g., if the thresholds for DO, C_{ave} and S_{ave} are too low), no warning would be released when there is an event; that is named a missed alarm or false negative. (2) If those conditions are easily fulfilled (e.g., if the thresholds for DO, C_{ave} and S_{ave} are too high), then warnings will frequently pop up even when there is no real event; that is named a false alarm or false positive. In both cases the credibility of the EWS will be negatively affected.

In order to tune the thresholds of the model tree, the historical data of DO and turbidity were used. A sensitivity analysis of the effect of the parameters C_{ave} and S_{ave} on the number of false and missed alarms was carried out considering the combination of warning levels 1, 2, 3 and 4, although only the analysis for levels 1 and 2 (the ones that imply imminent intake closure) are shown. Figure 7 shows the effect of changing both thresholds for DO on (a) POD and (b) FAR. Additionally, Figure 7(c)) shows the difference between plots (a) and (b), in order to make evident the combination of thresholds that provide bigger POD and simultaneously a lower FAR.

The parameters C_{ave} and S_{ave} for DO and turbidity were defined based on the combined objectives of POD and FAR, which corresponds to the darker zone in Figure 7(c) for the case of DO. The optimal values for C_{ave} and S_{ave} are

Table 1 | Contingency table for an EWS

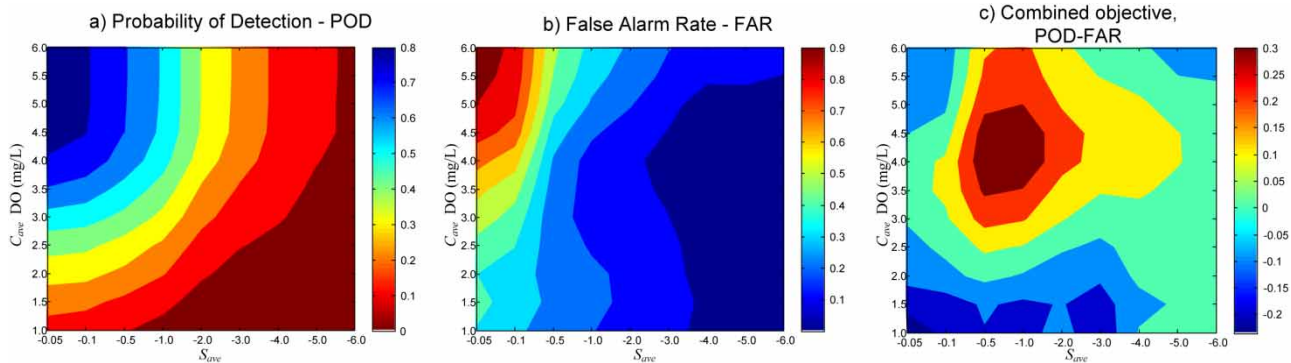|	Event observed	Event not observed
Warning issued	Hit (a)	False alarm (b)
Warning not issued	Missed alarm (c)	Correct negatives (d)

Figure 7 | Distribution of (a) POD, (b) FAR and (c) POD-FAR to determine optimal threshold values of C_{ave} and S_{ave} for DO and $N(t) = 1$ or 2.

presented in Table 2. The receiver operating characteristic (ROC) curve (Mason 1979; Mason & Graham 1999) is an accepted approach for representing the trade off between the false negative and false positive rates based on contingency tables. An acceptable EWS performance is obtained when the ROC curve rises towards the upper left hand corner of the graph (high POD and low FAR). However, if the ROC curve follows a diagonal from the origin to the upper right hand corner, it means that every improvement in POD corresponds to deterioration in FAR, and the

EWS does not perform well. The ROC curves for the optimal values for DO are shown in Figure 8.

Lead time definition

For an early warning system to be effective, the lead time will need to be sufficient for appropriate action to take place. Lead time includes the time of detection, time of decision and time of response (Verkade & Werner 2011; Parker & Fordham 1996; Carsell *et al.* 2004). For this case study, the estimated lead time is composed of two aspects, as follows. (1) The estimation of presence of pollution consisting of a detection time of 15 min, a computational time of a few seconds and a travel time of the pollutants that may vary between 2 and 5 hr depending on the current discharge in the river as explained in the next section. (2) The decision to take appropriate actions

Table 2 | Estimated parameters C_{ave} and S_{ave} for DO and turbidity

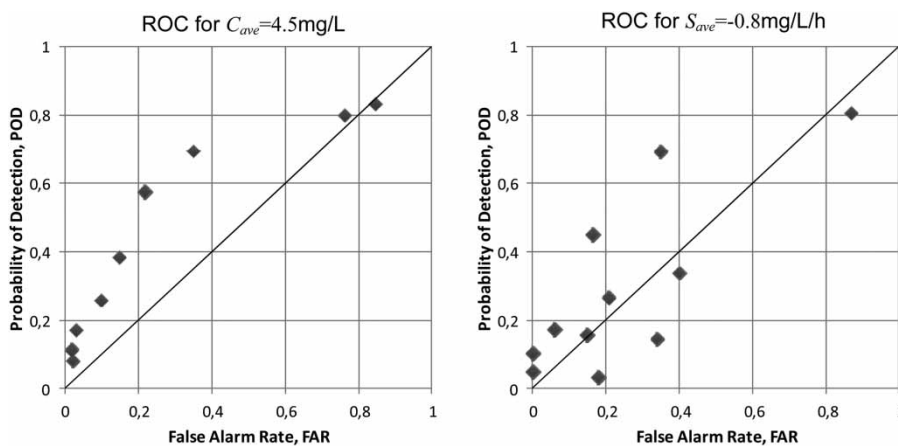| Parameter	DO	turbidity
C_{ave}	4.5 mg/l	800 NTU
S_{ave}	− 0.8 mg/l/h	400 NTU/h

Figure 8 | ROC curves for the selected threshold values of $C_{ave} = 4.5$ mg/L and $S_{ave} = −0.8$ mg/L/h for $N(t) = 1$ or 2.

(operational time), consisting of turning off the intake pumps (10 min) and turning on backup storage pumps (alternative source) (10 min). Critical lead time is therefore approximately 2 hr and 35 min, and is mainly driven by the travel time of the pollution.

Prediction of pollution travel time

The travel time is the time that a pollutant particle takes from the monitoring point at Milan up to the PMTP intake. This information is used in the EWS to inform the operator of the estimated arrival time of the pollution.

The transport of a substance in a river is determined by advection and dispersion processes. To estimate the effects of these processes we used a calibrated model of the Cauca River developed in Mike 11 software (Holguin *et al.* 2005). The model was used to simulate tracer experiments. In the upstream boundary, a digital particle with a certain mass was released as a point source and the model calculated the concentration at the downstream boundary. The travel time was estimated as the time difference between the centroids of each of the concentration curves. The mathematical experiment was repeated for different flow conditions in the river. The results of the model were compared with travel times estimated using data measured by EMCALI during pollution events. Calibration parameters

were fine tuned to fit the observed data. After that, the model was used to extrapolate the travel time for different flows. The results of travel times as a function of the river flow are presented in Figure 9. A power function was fitted to the results. This was done in order to simplify the implementation of the travel time model in the Centinela prototype. In consequence, the EWS not only requires the monitoring of water quality variables but also water quantity (i.e. flow). The river has an on-line monitoring station for water level that is used to estimate the flows at PMTP with the help of a rating curve. As pointed out before, the Cauca River does not receive any tributary or lateral discharge in the reach between Milan station and the PMTP and therefore it can be assumed that the flow is constant along this reach and that the mass of pollutants is only affected by transport and transformation processes.

Forecast of DO concentrations at the intake

To forecast the DO concentration at the intake of the PMTP a data-driven model was developed. The reason for choosing a model based on data is that the data available (i.e., travel time, DO concentrations at Milan and PMTP), are insufficient to run a process-based model, which needs, among other data, the concentration of organic matter discharged to the river at every time step.
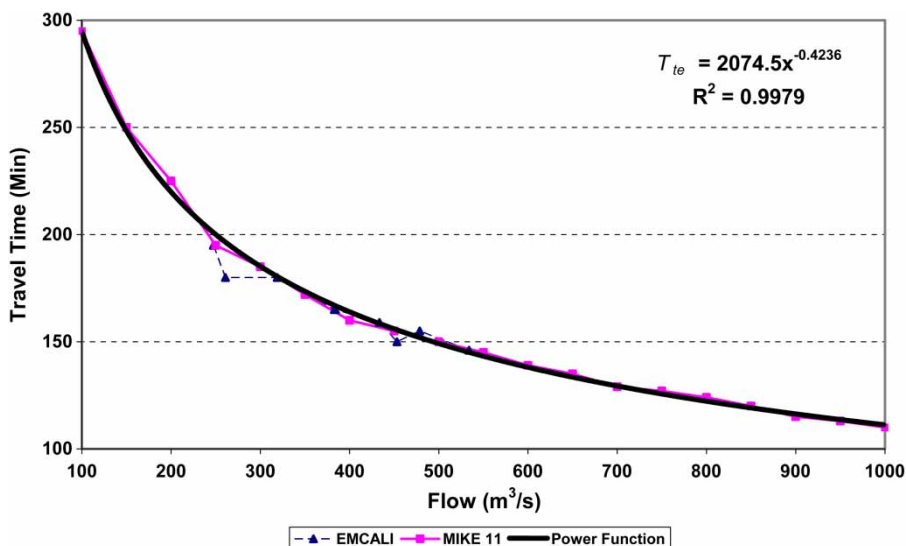

$$T_{te} = 2074.5x^{-0.4236}$$
$$R^2 = 0.9979$$

Figure 9 │ Pollution travel time as a function of river flow.

The analysis of the series shows that during an event, the minimum concentration of oxygen detected at Milan is related to the minimum concentration measured in the PMTP. Plotting the minimum measured on each side, a first relation was found. Following similar criteria for each event, the same number of data were selected before the minimum and after the minimum. In this way the same number of data are used for each series. After this, the values of oxygen at Milan are plotted against the values of oxygen at the intake of the PMTP. The data showed a linear relation for every event. The slopes of the lines suggested a further classification according to the travel time.

The model developed is a decision tree that considers the concentration of oxygen at the monitoring station upstream (Milan), the hydraulic conditions (e.g., water level, discharge) of the river and the travel time to estimate the expected concentration of oxygen at the intake of the PMTP, using 70% of the events as a training set. Figure 10 shows the resulting model tree to forecast the value of DO and Figure 11 shows the validation results of the model tree for two particular pollution events with different travel times.

Communication of warnings

A crucial component of an EWS is the generation and delivery of warning messages to the PMTP staff. The dissemination of the warnings is based on visual signals on the user interface of the web-based application and via text messages (SMS). The message includes the current warning level, the concentration values at the monitoring point, the value of the parameters (C_{ave} and S_{ave}) and the estimated time at which the pollution is expected at the intake. The purpose of the message is to inform the operator about the state of the system and this is done through SMS in order to guarantee that the operator is always informed. The rules for sending these messages are described in Figure 12 and, in principle, it considers sending alerts when the warning levels are between 1 and 4. It also classifies the type of messages according to the recipient. In this way, only warning levels 1 and 2 (imminent intake closure) are sent to the PMTP Director, in charge of making the decision whether to close it or not, while warning levels from 1 to 4 are sent to the PMTP Operator to warn him/her to be on alert. Although in principle warning levels from 1 to 7 have an associated alert component, it is not convenient to send
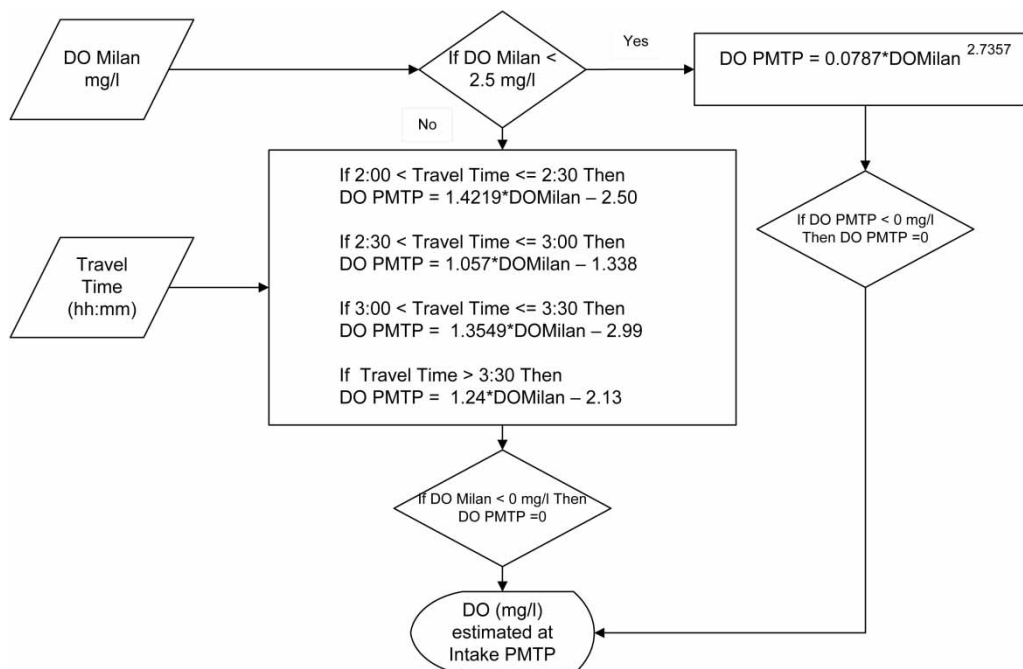

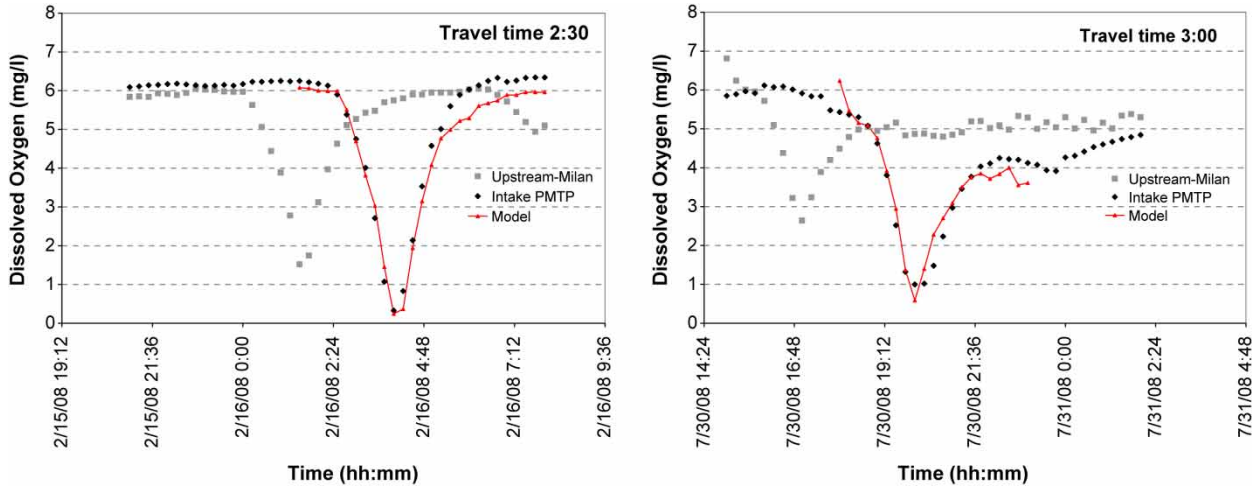**Figure 10** | Flowchart of data-driven model to forecast DO at the intake of PMTP.

Figure 11 │ Validation results of model tree shown in Figure 10.
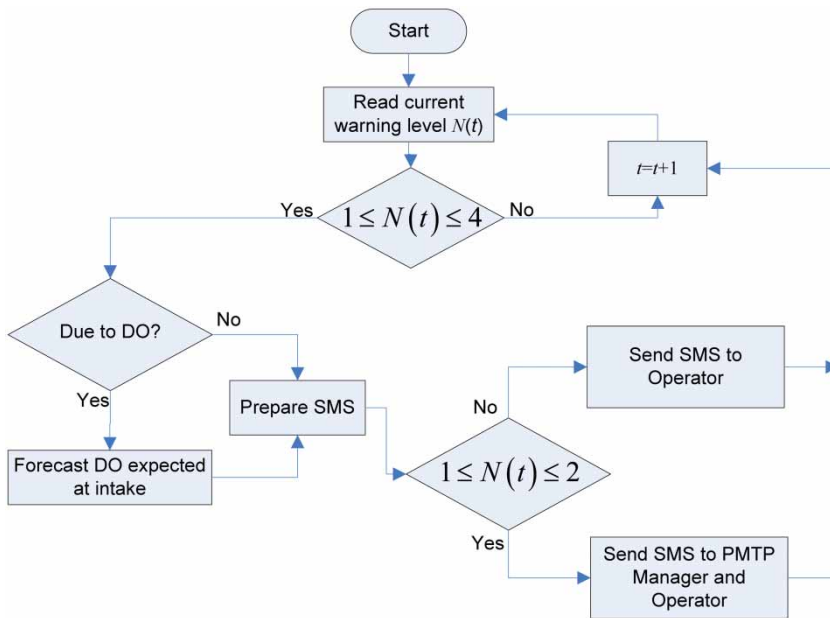

Figure 12 │ Flowchart with the rules for SMS delivery.

them all since messages from 5 to 8 may be treated as spam messages with the consequent lack of attention in the long term.

The technology used by Centinela to send warning messages utilizes a Clickatell Mobile Gateway, which allows SMS to be sent via the internet. In brief, when a warning is generated Centinela prepares and sends the message via the internet to that gateway, which transfers it to the predefined mobile ID through the local mobile network.

SUMMARY OF CENTINELA'S CONCEPTUAL FRAMEWORK

The conceptual framework of the application is presented as a flowchart in Figure 13. The process includes tasks that are repeated in a loop with a frequency determined by the sampling time step Δt of the data from the database. The input data are the water quality variables at Milan and PMTP (q_A and q_B, respectively) and the water levels (h) in
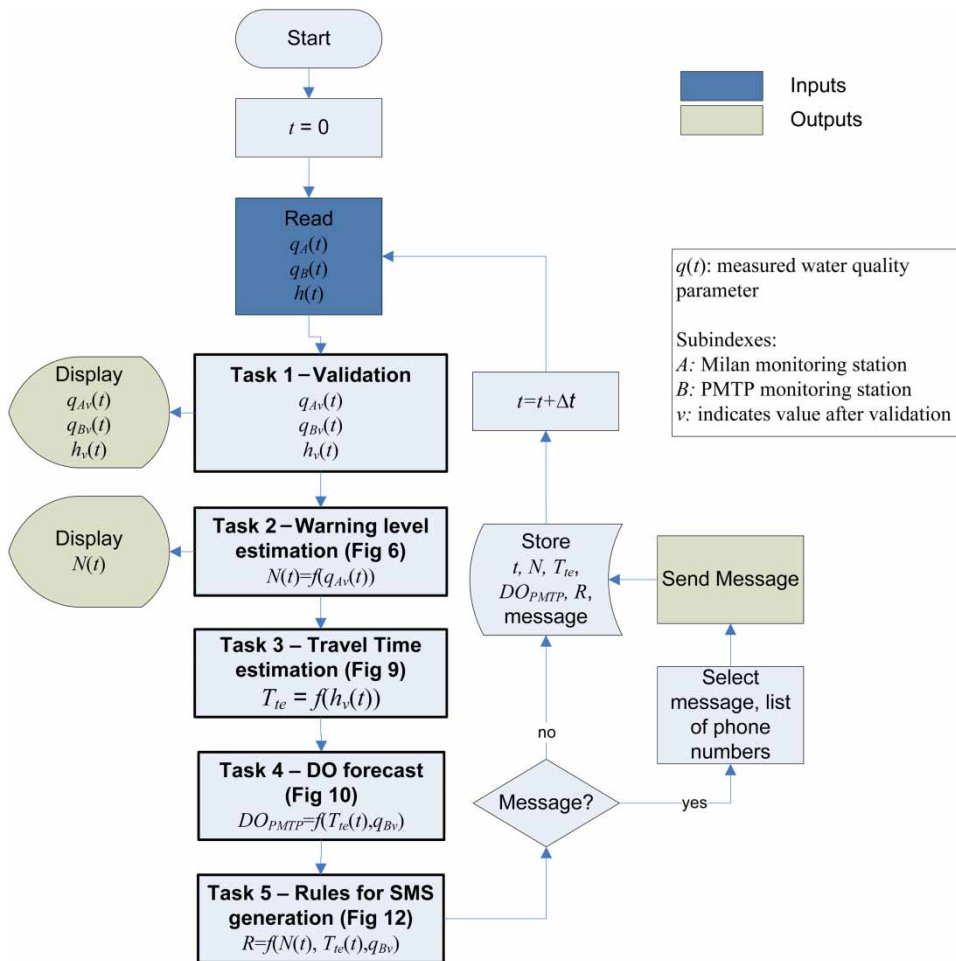
Figure 13 | Conceptual framework of Centinela, WQ-EWS for the Cauca River.

the river. The outputs of the application are displayed in the user interface and/or sent via text message. The outputs displayed are: time series of input data after validation, time (t), warning level (N), predicted reaction time (T_{te}) and predicted DO value at PMTP (DO_{PMTP}).

Five main tasks are identified according to the described methodology. The first task is the validation of input data measured in the monitoring stations. The second task estimates the warning level at time, t, using the water quality variables as input. The third task consists of a prediction of travel time (arrival of pollution at the intake). The fourth task corresponds to the forecast of the DO concentrations at the intake based on Milan's DO data and the flow conditions in the river. The fifth task comprises the generation and delivery of warning messages via SMS to the treatment facility operator. In addition, the framework

includes procedures to store the messages generated when an early warning is issued, which can be used to evaluate the EWS performance.

This conceptual framework was applied in a web-based application with a user-interface that allows interaction with operators of the treatment plant. Multiple clients can connect simultaneously to the Centinela application. The users can manually input data and get the data and information from stored data for post-processing.

CONCLUSIONS

This paper describes the conceptual development of the components required for an EWS for source waters used for the water supply of the city of Cali. This includes several

algorithms and models developed to provide the services required for the EWS. On historical data, the evaluation of Centinela showed the capability of detecting 60% of critical events caused by high loads of organic matter (low DO) and 70% caused by high levels of turbidity. More information needs to be collected and pollution events recorded to fine-tune the parameters and increase the performance of the system. Some parts of the algorithms developed here have been re-programmed in the database in Oracle. However, a full implementation of the prototype will certainly allow a proper evaluation of Centinela and the use of the warning generator as operational DSS.

Within the Centinela application two potential contaminants were analyzed: DO and turbidity as indicators of organic matter and suspended sediments. As the Cauca River also receives industrial and urban wastewaters discharges that carry a wider range of contaminants, future developments of Centinela should include more pollutants. The approach to identifying the presence of contaminants is focussed on variables measured by monitoring equipment. However, other mechanisms such as self-reporting of spills, reports by the water users or other water-related institutions should also be included as part of the EWS.

Further development of Centinela may include the response capability for operating during a critical pollution event. Alternatives include closure of intake, modification of treatment processes and the use of the pre-treated water stored in the recently built reservoir.

EWS should be viewed as an integral part of the operation of a water supply system. The design described in this paper covers all the components; however, further developments should include operational responses in the treatment processes and the storage.

ACKNOWLEDGEMENTS

REFERENCES

Abebe, A. J. & Price, R. K. 2005 Decision support system for urban flood management. *Journal of Hydroinformatics* 7, 3–15.

Alfonso, L., He, L., Lobbrecht, A. & Price, R. 2013 Information theory applied to evaluate the discharge monitoring network of the Magdalena River. *Journal of Hydroinformatics* 15, 211–228.

Balk, F., Okkerman, P. C., van Helmond, C. A. M., Noppert, F. & van der Putte, I. 1994 Biological early warning systems for surface water and industrial effluents. *Water Science & Technology* 29 (3), 211–213.

Carsell, K. M., Pingel, N. D. & Ford, D. T. 2004 Quantifying the benefit of a flood warning system. *Natural Hazards Reviews* 5, 131.

Constain, A. J., Carvajal, J. & Carvajal, A. 2008 A thermodynamic method to measure mixing lengths in natural streams. In: *River Flow 2008 International Conference*, Turkey.

DANE 2012 *Estadisticas vitales – demograficas: Censo general 2005*. Departamento Administrativo Nacional de Estadisticas, Bogota, Colombia.

Doswell, C., Davies-Jones, R. & Keller, D. 1990 On summary measures of skill in rare event forecasting based on contingency tables. *Weather and Forecasting* 5, 576–585.

EMCALI 2009 *Internal database and monitoring system*. Empresas Municipales de Cali EICE ESP, Cali, Colombia.

Galvis, A., Sabogal, L. P., Vélez, C. & Ramírez, C. 2001 Cauca river water quality assessment. In: Refresher Seminar. Water Quality Management Strategies for Sustainable Use of Water Resources, Cali, Colombia.

Grayman, W. M., Deininger, R. A. & Males, R. M. 2000 Early warning and source-water monitoring systems. In: *Water Quality Technology Conference Proceedings*.

Grayman, W. M., Deininger, R. A. & Males, R. M. (eds) 2001 *Early Warning and Source-water Monitoring Systems*. American Water Works Association, Denver, CO.

Hadihardaja, I. K., Latief, H. & Mulia, I. E. 2010 Decision support system for predicting tsunami characteristics along coastline areas based on database modelling development. *Journal of Hydroinformatics* 13, 96–109.

Holguin, J., Vélez, C., Galvis, A., Ramírez, C., Baena, L. & Duque, A. 2005 Implementing a dynamic model to study the cauca river water quality. In: *AGUA 2005 – International Seminar Integrated Management of Water Services in Nucleated Settlements*, Cali, Colombia.

Izydorczyk, K., Tarczynska, M., Jurczak, T., Mrowczynski, J. & Zalewski, M. 2005 Measurement of phycocyanin fluorescence as an online early warning system for cyanobacteria in reservoir intake water. *Environ. Toxicol.* 20, 425–430.

Kawka, M. & Cousson, R. 2012 Numerical modelling as a decision support tool for the design of an early warning system for

surface water intake. In: *10th International Conference on Hydroinformatics HIC 2012*, Hamburg, Germany.

Krzhizhanovskaya, V. V., Shirshov, G. S., Melnikova, N. B., Belleman, R. G., Rusadi, F. I., Broekhuijsen, B. J., Gouldby, B. P., Lhomme, J., Balis, B., Bubak, M., Pyayt, A. L., Mokhov, I. I., Ozhigin, A. V., Lang, B. & Meijer, R. J. 2011 Flood early warnings system: Design, implementation and computational modules. *Procedia Computer Science* **4**, 106–115.

Madera, C., Pena, J. & Quiroga, E. 2007 Resultados de Investigación en Agua Potable, Aguas Residuales y Biosolidos. Resumen Ejecutivo, Universidad del Valle y Empresas Municipales de Cali, Columbia. http://objetos.univalle.edu.co/?q=node/451 (accessed online 5th February 2014).

Mason, I. 1979 On reducing probability forecasts to Yes/No forecasts. *Monthly Weather Review* **107**, 207–211.

Mason, S. J. & Graham, N. E. 1999 Conditional probabilities, relative operating characteristics, and relative operating levels. *Weather and Forecasting* **14**, 713–725.

Mishra, A. K. & Coulibaly, P. 2009 Developments in hydrometric network design: A review. *Reviews of Geophysics* **47** (2), 1–24.

Miyamoto, M., Sugiura, A., Okazumi, T., Tanaka, S., Nabesaka, S. & Fukami, K. 2012 Suggestion for an advanced early warning system based on flood forecasting in Bengawan Solo River basin, Indonesia. In: *Tenth International Conference on Hydroinformatics HIC 2012*, Hamburg, Germany.

Montoya, C., Loaiza, D., Torres, P., Cruz, C. H. & Escobar, J. C. 2011 Effect of increase of raw water turbidity on efficiency of conventional drinking water treatment processes. *Revista EIA. Escuela de Ingeniería de Antioquia (online)* **16**, 137–138.

Parker, D. & Fordham, M. 1996 Evaluation of flood forecasting, warning and response systems in the European Union. *Water Resource Management* **10**, 279–302.

Pyayt, A., Mokhov, I., Kozionov, A., Kusherbaeva, V., Krzhizhanovskaya, V., Broekhuijsen, J. & Meijer, R. 2012 Signal processing methods for flood early warning systems. In: *10th International Conference on Hydroinformatics HIC 2012*, Hamburg, Germany.

Quansah, J. E., Engel, B. & Rochon, G. L. 2010 Early warning systems: A review. *Journal of Terrestrial Observation* **2**, 23–44.

Roberson, J. A. & Morley, K. M. 2005 *Contamination warning systems for water: An approach for providing actionable information to decision-makers*. Report of American Water Works Association, Denver, USA.

Romang, H., Zappa, M., Hilker, N., Gerber, M., Dufour, F., Frede, V., Bérod, D., Oplatka, M., Hegg, C. & Rhyner, J. 2011 Ifkis-hydro: An early warning and information system for floods and debris flows. *Natural Hazards* **56**, 509–527.

Tryland, I., Surman, S. & Berg, J. D. 2002 Monitoring faecal contamination of the Thames estuary using a semi-automated early warning system. *Water Science & Technology* **46** (3), 25–31.

UN 2006 Developing early warning systems: A checklist. In: *EWC III Third International Conference on Early Warning From concept to action, United Nations/Inter-Agency Secretariat of the International Strategy for Disaster Reduction (UN/ISDR)*. Platform for the Promotion of Early Warning (PPEW) Bonn, Germany.

UNEP 2012 *Early warning systems a state of the art analysis and future directions*. Report of Division of Early Warning and Assessment (DEWA), United Nations Environment Programme (UNEP), Nairobi, Kenya.

UN-ISDR 2004 Terminology: Basic terms of disaster risk reduction. http://www.unisdr.org/eng/library/lib-terminology-eng%20home.htm. International Strategy for Disaster Reduction, Geneva.

Univalle 2008 *Urban water management for the city of Cali, diagnosis report*. Report of Universidad del Valle and UNESCO-IHE Institute for Water Education, Cali, Colombia.

USEPA 2005 *Technologies and techniques for early warning systems to monitor and evaluate drinking water quality: A state-of-the-art review*. Report of Environmental Protection Agency, Office of Research and Development, National Homeland Security Research Center, Cincinnati, OH.

USEPA 2007 *Water security initiative: Interim guidance on planning for contamination warning system deployment*. Report of Environmental Protection Agency, Office of Research and Development, National Homeland Security Research Center, Cincinnati, OH.

Vélez, C., Galvis, A., Ramírez, C. & Baena, L. 2006 Cauca river water quality model – hydroinformatics application in a developing country. In: *Seventh International Conference on Hydroinformatics HIC 2006* (P. Gourbesville, ed.), Research Publishing, Nice, France.

Vélez, C., Alfonso, L., Sánchez, A., Galvis, A. & Sepúlveda, G. 2012 Centinela: An early warning system for the water quality of the cauca river. In: *10th International Conference on Hydroinformatics HIC 2012*, Hamburg, Germany.

Verkade, J. S. & Werner, M. G. F. 2011 Estimating the benefits of single value and probability forecasting for flood warning. *Hydrology Earth System Science* **15**, 3751–3765.