A data validation method for real-time control of sewer systems.

Eric Piatyszek - Société RHEA.
Centre SIMADE, Département Ingéniérie de L'Environnement
Ecole des Mines de Saint-Etienne
158 cours Fauriel - 42023 Saint Etienne, Cedex 2. FRANCE.
tel. : 0477420117 - Fax : 0477426666 - e-mail piatysze@emse.fr
Pierre Voignier - Société RHEA.
11 rue du vieux Pont - 92000 Nanterre. FRANCE.
Tel. : 0477420134 - Fax : 0477426666 - e-mail : voignier@emse.fr
Didier Graillot - Centre SIMADE, Département Ingéniérie de L'Environnement
Ecole des Mines de Saint-Etienne
158 cours Fauriel- 42023 Saint Etienne, Cedex 2. FRANCE
tel. : 0477420130 - Fax : 0477426666 - e-mail graillot@emse.fr

Abstract. Urban sewer networks include more and more sensors and automatic devices; the new regulations of 1992 set new objectives and as a consequence, the task of operators who manage and supervise the urban sewer networks is more and more complex. Real-time control systems can provide them with attractive support, especially in order to reduce the damage caused by pollutant discharges and floods. However, one of the limitations of such real-time acquisition systems is that they require the efficient and rapid interpretation of the data by the operators who have to make the best decision at the best time. This paper presents the development of a real-time system for fault-detection and data-validation which aims to improve the quality of information and on-line decision assistance to the operator.

Key Words. Real-time control, Urban sewer netwoks, fault-detection, data-validation, Kalman-Filter.

I - INTRODUCTION.

Mainly invisible at first glance, Urban Sewer Networks (USN) play nevertheless a major role in our quality of life. Their goals have been increasing with time, and especially, with the growth of cities. Primarily limited to sanitary role during the XVIII th century, the boom of urbanization, at the time of industrial revolution made new demands upon them : the disposal of rainwater and the reduction of magnitude and frequency of damage caused by floods. As a result, until the middle of XX th century, the main aim of USN was the disposal of waste-waters and rainwaters. But for the last 50 years, population increase and industrial development have lead to an increase in pollutant discharges into the receiving waters. This is the reason why, according to the new regulation, purification of wastewaters and more recently even rainwaters before discharge, became a priority. It is worth noticing that this objective has been partially achieved. Indeed, in France, in 1993, the efficiency of the sewage treatment plant was 69%, the waste-water collected rate was about 65%, and the national depollution rate was 45% [RNDE, 1995].

Thus, the task of operators who monitor the USN, has became more complex. This job is even more complicated since new regulations were passed in 1992. As a consequence, the sanitation facilities are judged, not only by users, town councils, ecologists, but also by state control services such as 'Agences de l'Eau'. Now, the operators who run the USNs have to meet these various targets with respect to quality cost and reliability.

A means to facilitate the management of USN is based on real time control or supervised systems. Such systems use sensors, actuators, means of transmission, calculators and automatic devices. Real time control systems can be 'static' or 'dynamic'. The former correspond to supervised sewer systems. In the latter, the network is adapted, in real time, to a new situation due to a random phenomenom [Lancelot, 1986]. The actions on the network, adjust its configuration, help regulation by automatic devices or remote control and aim at defining strategies to avoid pollution and/or floods. This is carried out by means of measurement of water levels and water flows at several key points in the network. These sensors have to provide the most relevant and most realistic views of the real state network. The measurements taken are tele-connected to a control centre. Based on the information from these measurements and the knowledge of the network, the operator decides which actions to undertake (i.e. : lock gate, start pump, ...).

Thus, the measurements form the keystone to any decision making. The information provided by the sensors has to be the most representative of the real phenomena taking place in the network, in order to settle the best management strategy.

Since 1989, RHEA has studied and performed data-processing methods to validate the information provided by sensors. This validation seeks to check the coherence of measurements by detecting inconsistent situations. In addition it can provide values to missing data caused by a failure of measurement (sensor breakdown, cut of the transmission line).

In the first section, this paper will describe, one of these fault-detection and validation methods and in the second part an example will be developed.

II - Fault-detection and validation Method.

II - 1 General presentation.

As above-mentionned, we want to control the global coherence of the measurement network. This data-validation is based on information-redundancy [Cassar et al., 1996]. This redundancy can be a hardware redundancy (several sensors measure the same variable), or rely on the knowledge of the system. In the case of USN, the installation of sensors is tricky (difficult access, hostile surroundings...). That is why we can not systematically use hardware redundancy. The operators however, usually know their networks very well. They have excellent information on rainfalls, from the analysis of the radar data (RHEA patent number FX9208545, registered the 09/07/92), fitted with the measurements on rain gauges [Cunge et al., 1992]. This knowledge of the system (networks architecture, estimate of the rainfall) makes it possible to build up a model. It is essential to know with accuracy (spatial distribution, precipitation depth, duration) rainfalls in order to achieve a relevant and effective validation.

This model defines a fixed context for our network. That is to say, the working state of our system, linked with a particular configuration of our networks (opened or closed gate, load-shedding...). For the most part, this model represents the 'normal' operating state of our network : i.e. the working state usually observed. The coherence of both information coming from the sensors and the information made by our model is tested so as to criticize, validate the measurements and should the occasion arise, to detect mismatching. Such mismatches are mainly caused by divergences between the observed state and the predicted state of our model. They express an inconsistent situation, which is due to either sensor faults or to component faults (gate opened too much or closed, surface yield too important, obturated collector). The sensors' faults have a local effect, limited to the disturbed sensor. As far as component faults are concerned, their effects are global : a component fault affects the information provided by several sensors. Thus, we can distinguish origins of different faults.

We have performed several approaches which allow us to test the global coherence of the measurement network, by means of a model. This described in this article offers the advantage to provide values to missing data. It distinguish from the others methods by the way we build the model and we test the coherence. One can find a description of the others methods in [Szafnicki et al., 1994] and in the patent number FR9313799.

II - 2 Analytical approach based on a state estimator (figure 1).

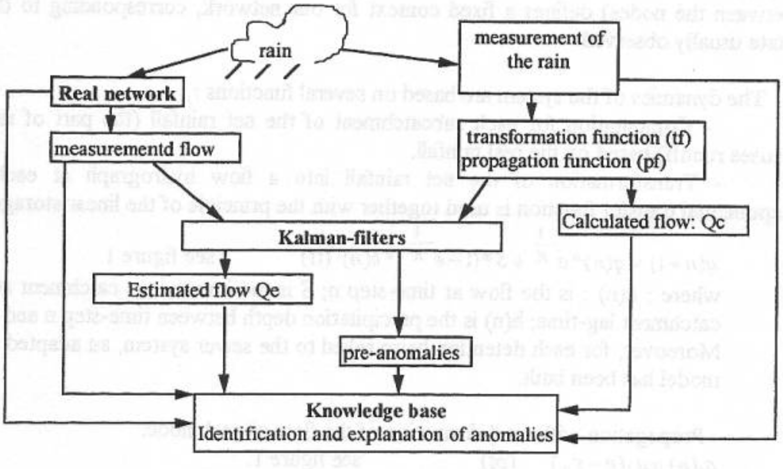

Figure 1. scheme of our detection/validation method.

Choice of a model.

The physical laws which describe the dynamics of our system are known and thus we have a deterministic representation of our networks : for example detailed information about the flow within the pipe based on a surface runoff function (tf), and a transport function (pf) is available. That is why, from this deterministic representation, we can build up a state estimator, for the estimation of flows. However, these estimates may be biaised. Firstly, because of the measurement accuracy which is, in some cas, about 30% [Deleu, 1990] and, secondly, because of the physical laws which do not exactly describe the flow complexity (open-channel flow, pressure-conduit flow, non linear phenomena, hydrograph deformation ...). This uncertainty which is partly due to measurement and partly due to the model cannot be ignored. To this end we have chosen a Kalman-filter rather than a state estimation observer [Deleu, 1990][Chen, 1993]. Indeed, the discrete-time Kalman filtering algorithm uses a stochastic state-space form of our system [Chen, 1993]. The estimate produced by this algorithm, integrates the uncertainties due to unmodelled disturbances, noises and model mismatch and sensors faults.

Building of a Kalman-filters bank.

From the real architecture of our network, a simplified model has been built, which has got as many nodes as points where a flow is measured. Each node is associated with one subcatchment which models the contribution of all the surfaces that are directly connected to that node. Connections are established between the various nodes and define a privileged flow course. This structure of the model (number of nodes, limits of the catchments, connections between the nodes) defines a fixed context for our network, corresponding to the operating state usually observed.

The dynamics of the system are based on several functions :

- Computation for each subcatchment of the net rainfall (the part of rainfall which causes runoff) based on the real rainfall.

- Transformation of the net rainfall into a flow hydrograph at each node. An exponential transfer function is used together with the principle of the linear storage (tf) :

$$q(n+1) = q(n) * e^{-\frac{1}{K}} + S * (1 - e^{-\frac{1}{K}}) * h(n) \quad \text{(tf)} \qquad \text{see figure 1}$$

where : q(n) : is the flow at time-step n; S is the impervious catchment area; K is the catchment lag-time; h(n) is the precipitation depth between time-step n and n+1.
Moreover, for each detention basin relied to the sewer system, an adapted hydrological model has been built.

- Propagation without deformation of the flow of each node.

$$q_i(n) = q_j(n - \tau_{ij}) \quad \text{(pf)} \qquad \text{see figure 1.}$$

where : q_i is the flow at node i, coming from node j; q_j is the flow at node j; τ_{ij} is the time propagation between node i and j.

The structure and the dynamics of our model make it possible to build a state-estimator, which is a Kalman-filter [Deleu, 1990]. More precisely, these functions define the matrix of the dynamics and measurement equations of the Kalman-Filter [Chen, 1993].

As mentionned aboved, one Kalman-filter describes, by means of its two equations, one particular operating context of our network. The starting idea might be to build as many Kalman-Filters as implied by the number of different contexts. In that case, detection and diagnosis are simultaneous. Indeed, detection determines the most likely filter and, since each filter represents one inconsistent situation, characterizes the context of the network. But, this remains utopian, in the face of the great number of different working situations and the feasibility of modelising each of them. Considering that sensors faults and component faults also might produce either an increase, or a decrease in the observed flow, we have, therefore, only kept three Kalman-filters. The first one describes a common working context of the networks, the two others a positive or a negative bias on the measurement. These filters are separated enough so as not to be in competition on a same anomaly and are, for a short period, the manifestation of hydraulic situations often met (opening or closing of gates, surface yield too important ...). In that approach, we successively make the detection and the diagnosis.

Detection and diagnosis.

The detection module enables determination of whether a node is in a normal state (observed situation consistent with the situation predicted by the model) or in an abnormal state. In the case of an abnormal situation (filter with positive bias or negative bias is the most likely), it must be known whether this anomaly is due either to sensors faults or to component faults. This choice is carried out within the diagnostic module by means of rules [Framling, 1992].

Thus, for each node, and at each time-step, a label is given to the measurement, which indicates whether the measurement is consistent or not with the a priori information of our system. Moreover, in case of lack of measurement, an estimate of the flow can be provided to the operator.

III - An on site test example, with real data.

The test has been realized with data from the USN of the Seine Saint Denis county.

Our validation method has been applied to a stormsewer network, for a rain storm.This network has been modelised with 4 nodes located at strategic key points of the network (i.e. inputs and outputs of tank, main collectors) (figure 2).

The radar data of this rain storm show us a homogeneous space distribution of the rain event on the upstream part of the network (return period between 2 and 5 years) and a more intense rainfall near the node 4 (return period of 10 years).

We have analyzed the measurements, with our method described in section II. We haven't detected any mismatch on the upstream nodes of the network (node 1, 2 and 3). This is the reason why, this example is focused only on node 4.

The figure 3 shows the inflows at node 4 together with the hyetograph of the studied rainstorm event.

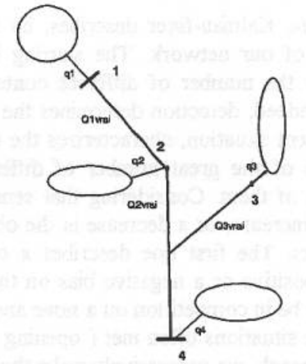

Figure 2 Scheme of the modelised network

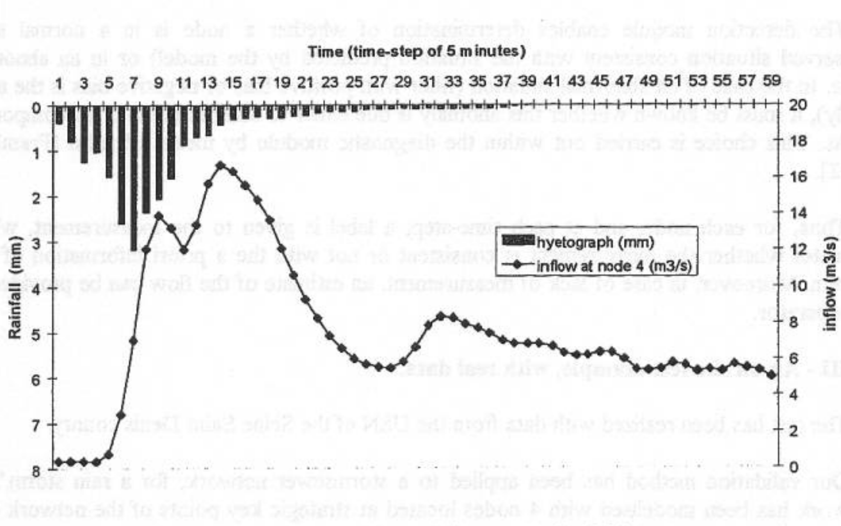

Figure 3. Inflow at node 4 and hyetograph of this event.

It is worth noticing that this inflow QE is not measured directly. The flow at the output of the detention basin of node 4 and the water level in the basin are known. Thus, based on a volumetric analysis, one can calculate QE (formula f1) :

$$QE = QS + S(Y)*\frac{\Delta Y}{\Delta t} \qquad \text{(f1)}$$

where : QS is the outflow of node 4 detention basin; Y is the measured water-level in the basin; S(Y) is the surface of the basin.

As a consequence, the parameters (Surface of the basin, measurement of the water level) must be coherent with the on site reality to calculate QE.

The results of the validation are presented in figure 4.

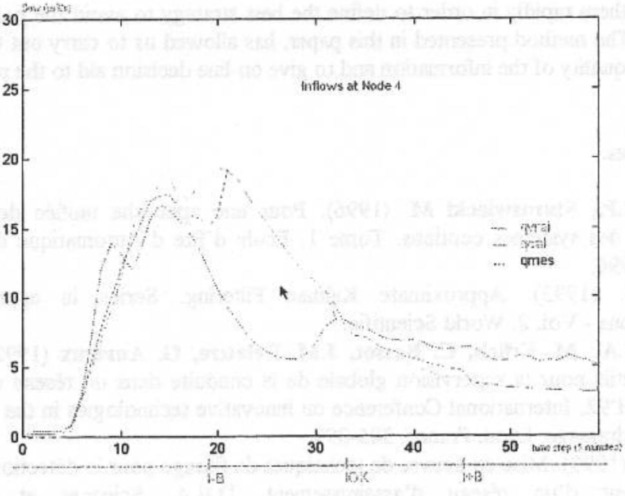

Figure 4 Inflows at node 4.

On this diagram, the dotted line represents the measured inflow, the dashed-line the calculated inflow and the solid line the likely inflow. The likely inflow is the inflow chosen by the algorithm to represent the hydraulic situation at that node. The method detects two inconsistent situations : the first one, between time-step 20 and 33, the second one between time-step 45 and 60.

The first anomaly is induced by a measured flow lower than the estimated : the measured flow is below the calculated flow which is merged in the likely flow. As the upstream nodes are in a normal state, this anomaly is caused by a local mismatch to node 4. In addition, a delayed time analysis of the parameters of formula f1, shows that the likely flow is the more realistic.

The second anomaly is induced by a modelised inflow lower than the measured : the measured flow is over the calculated flow which is merged in the likely flow. In fact, this is a component fault due to an abnormal quantity of flow not modelised.

On this example, we have detected and diagnosed two anomalies : one sensor fault, and one component fault.

IV - Conclusions and perspectives.

In order to provide a quality service, to effectively fight against pollutions and floods, the urban sewer operators equip their networks with sensors, actuators and automatic devices. As a consequence, they are overwhelmed with a great number of measurements. It is crucial to interpretate them rapidly in order to define the best strategy to avoid the pollutant discharges and floods. The method presented in this paper, has allowed us to carry out this validation, to improve the quality of the information and to give on-line decision aid to the pilot.

References.

Cassar J.P., Staroswiecki M. (1996). Pour une approche unifiée de la surveillance. Surveillance des systèmes continus. Tome 1. Ecole d'Ete d'Automatique de Grenoble, 2-6 septembre 1996.

Chen G. (1993). Approximate Kalman Filtering. Series in approximations and decompositions - Vol. 2, World Scientific.

Cunge, J.A., M. Erlich, C. Bassot, J.M. Delattre, G. Auriaux (1992). Faisabilité de nouveaux outils pour la supervision globale de la conduite dans un réseau d'assainissement. NOVATECH'92. International Conference on innovative technologies in the domain of urban storm water drainage. Lyon, France. 385-399.

Deleu B. (1990). Mise en oeuvre de techniques de filtrage pour la détection d'anomalies de fonctionnement d'un réseau d'assainissement. D.E.A. Sciences et Techniques de l'Environnement, CERGRENE, 1990.

Framling K. (1992). Connectionist techniques for detecting anomalies in urban networks. 4th European Junior Scientist course - Assessment of modeling uncertainties and measurement error in hydrology - Saint -Victor sur Loire - Saint-Etienne. 99-119.

Lancelot B. (1986). Les différents systèmes de gestion automatisée des réseaux d'assainissement existants en France - Conséquences sur l'efficacité et la fiabilité des réseaux, sur la réorganisation des services et la qualification du personnel. Eau et Informatique. Actes du colloque organisé par l'Ecole Nationale des Ponts et Chaussées, 206-215.

RNDE, (1995) L'assainissement des grandes villes. Neuilly sur Seine : Ministère de l'environnement, 1995. 20 pages.

Szafnicki K., Graillot D., Voignier P. (1994). Knowledge-based fault detection and supervision of urban sewer networks. IFAC-SAFEPROCESS'94, Espoo, Finlande, T. Ruokonen Ed., Helsinki 1994, 1 ,171-176.