Nordic Hydrology, 31 (2), 2000, 107-126

Multivariate Technique for Validating Historical Hydrometric Data with Redundant Measurements

Saad Bennis
École de technologie supérieure, Montreal, Que H3C 1K3 Can.

Narut Kang
Hydro-Québec, Montreal, Que H2Z 1A4 Can.

The aim of this research was to develop an automated methodology for validating chronological series of natural inflows to reservoirs. Theoretically, gauges located on the same reservoir should indicate the same reading. However, under the influence of meteorological and hydraulic factors, or simply because of failed measuring equipment, there may be large deviations between the various measurements. Since the calculation of historical natural inflows is directly linked to the measurement of reservoir level by the water balance equation, there will be as many series of natural inflows as there are of reservoir levels. A multivariate filtering technique is used to validate the historical natural inflow computed by each water level variation. The multi filter methodology has the advantage of balancing the water volume of natural inflows to the reservoir when applied over a relatively long period of time. As a result, the validated flood peaks are not systematically overestimated or underestimated and the validated natural inflows are nearly identical for all the gauges. The proposed technique has been incorporated into a software program called ValiDeb, which has been successfully tested on-site on the Gatineau River in Quebec.

Introduction

At Hydro-Québec, a major Canadian hydroelectric company, engineers and specialists manage reservoirs with a capacity of 200,000 hm^3, which produce 95% of the energy consumed in Quebec. In order to optimize hydroelectric production, hydrologists require an accurate estimate of the energy stored as water in the reservoirs.

These data are then used to plan and manage the hydroelectric plants according to four time horizons: very short, short, medium, and long term. Hydrometric data currently available are insufficient to manage the system efficiently. Of course, it is necessary to anticipate how phenomena will evolve in the future. Hydrologists have shown that erroneous and missing data can cause misleading forecasts, thereby compromising the optimal management of water resources (WMO 1985). The daily net basin supply of the large reservoirs is currently computed by the water balance equation, which is specifically based on gauge measurements. The number of gauges and their location heavily depends on simplified hypotheses that generate many errors in the computation of the net basin supply. These problems greatly affect Hydro-Québec's real time management of water resources. Unfortunately, the validation of data, especially hydrometeorological data, is very poorly documented. The classical algorithms used for finding anomalies make it possible to automatically detect measurements that are very clearly defective, such as those that exceed a given threshold, but frequently fail because they do not pick up less obvious anomalies (Perreault *et al.* 1991). Furthermore, these techniques are often based on univariate laws of probability and do not take into account either variables or cross correlations between the different variables.

An effective way to ensure precise measurements is to increase the number of gauging stations on large reservoirs. A reservoir may then be equipped with two stations (duplex system), three stations (triplex system), or more (multiplex system). This material redundancy, by opposition to analytic redundancy, makes it possible to detect the deficient gauge (Bath 1985). Following this principle, Hydro-Québec recently decided to add specially designed gauging stations at their large reservoirs in barren areas and harsh climates. Theoretically, gauges installed at different locations on the same reservoir should show the same readings. Unfortunately, depending on the hydrometeorological parameters and deficiencies in the gauges, a significant difference may occur between the various readings. In addition to the isolated disturbances on some gauges, there is a scale factor that varies over time between the readings obtained from different gauges. The existence of only one very inaccurate value may alter the mean reservoir level, thereby falsifying the amount of stored volume and reservoir inflows. On the other hand, the use of the one station deemed most representative does not makes optimal use of all the available information.

One must therefore establish a rule of decision to set aside the erroneous readings before averaging the accepted values in some way. Considering a triplex system that measures, at time t, the reservoir levels L_{1t}, L_{2t}, and L_{3t}, a simple error diagnosis would involve calculating the three residuals

$$\begin{aligned} r_t(1,2) &= L_{1t} - L_{2t} \\ r_t(1,3) &= L_{1t} - L_{3t} \\ r_t(2,3) &= L_{2t} - L_{3t} \end{aligned} \qquad (1)$$

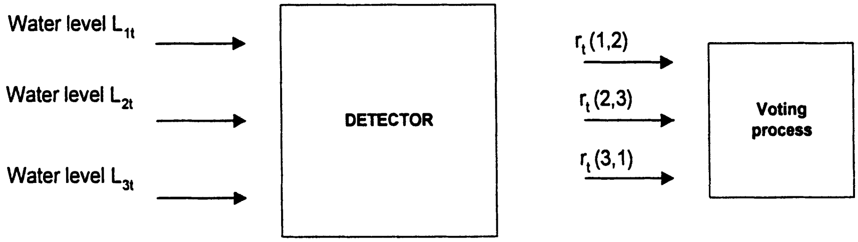

Fig. 1. Simple voting process for diagonosis.

The deficient gauge may then be detected by using a logical voting system that compares these residuals $r_t\ (i,j)$ to a given threshold, based on the specified tolerance for the computed hydrometric data. The principle of this method is illustrated in Fig. 1.

As in the univariate case, the drawback of this method resides in the constraint to impose an arbitrary threshold to detect the erroneous values. In fact, the errors are not always of large amplitude and the information may be scattered between the various gauges.

Recently, Roy *et al.* (1992) proposed the virtual-level method, which makes it possible to correct many of the anomalies observed in natural inflows. However, many negative natural inflows can be observed following the virtual-level validation exercise. Also, while this technique does improve the quality of natural inflows computed on the basis of virtual levels, it does not permit validation of the level series of reservoirs.

This article proposes a complete methodology for validating all hydrometric data. The validation methodology will be carried out in three steps. The first step involves calculating the natural inflows by applying the water balance equation individually to readings from each gauge. Secondly, these series of natural inflows will be validated by the proposed methodology. Finally, the validated natural inflows will be used to re-estimate the levels of the reservoirs, along with all the hydrometric variables derived from them.

Method for Calculating Net Natural Inflows

In the most general case, for an incompressible fluid such as water, the change in storage volume consists of the difference between inflows and outflows

$$\frac{dV}{dt} = Q_I - Q_O \tag{2}$$

where

Q_I (m^3/s) is the sum of all inflows to the system;
Q_o (m^3/s) is the sum of all outflows from the system;
V(m^3) is the volume of water contained in the defined space.

If the system consists of a reservoir supplied by a river, the inflows and outflows may be written as

$$Q_I = I + R + P + FIN \tag{3}$$

$$Q_O = O + E + T + FOUT \tag{4}$$

where I is the flow released from upstream reservoir which may be obtained by a flood routing type model. This quantity is generally available;
R is the excess rainfall entering the reservoir. This quantity also includes the inflow from ungauged rivers and streams that supply the reservoir. This quantity is difficult to measure directly;
P is the precipitation falling directly on the reservoir. This quantity may be measured directly but is not always available;
FIN is the subsurface flow to the reservoir. This quantity is difficult to ascertain directly;
O is the reservoir outflow. This value includes discharged flow, spilled flow, etc. All of these components are usually known;
E is evaporation from the reservoir surface. This quantity may be estimated theoretically but the necessary parameters are not always available;
T is transpiration by the vegetation cover. This quantity is difficult to ascertain directly;
$FOUT$ is the infiltration from the reservoir to the groundwater system. This quantity is difficult to ascertain directly.

Substituting Eq. (3) and Eq. (4) into Eq. (2) yields

$$\frac{dV}{dt} = I - O + NI \tag{5}$$

where

$$NI = R + P + FIN - FOUT - E - T \tag{6}$$

NI is the net natural inflow (Fig. 2)

During flood periods, the first term on the right-hand side of Eq. (6) greatly dominates the subsequent terms, resulting in a positive net natural inflow. During drought periods, the net natural inflow term may become negative due to marked evaporation or sublimation losses. In such a case, it is necessary to ensure that the amplitude of these negative inflows remains within a given interval, which corresponds to that of the potential evaporation losses, estimated from meteorological data. In all cases, the terms on the right-hand side of the equations that contribute to

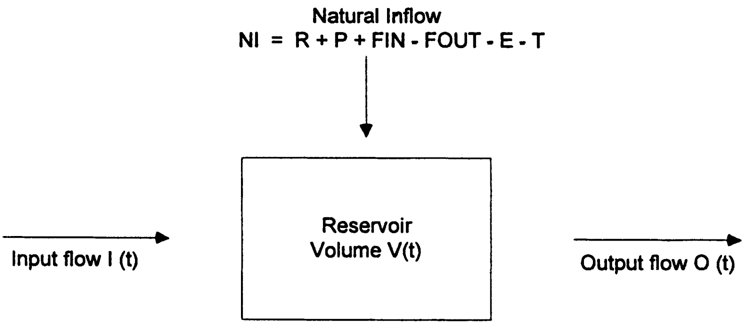

Fig. 2. Water balance equation.

natural inflow demonstrate a particular persistence in their daily variations. This persistence is more evident as the area of the drainage basin and of the reservoir increases. This natural smoothing behavior will be reproduced by mathematical validation models in order to correct the natural inflows, which follow a sawtooth pattern of great amplitude, with alternating positive and negative signs. The error diagnosis in the latter case systematically casts doubt on the validity of the reservoir's level reading.

In practice, the reservoir volume is estimated based on the measured level and the pondage equation specific to the reservoir. The calculation of natural inflows is therefore directly related to the measurement of the reservoir level. The amount of flows discharged and spilled is also estimated on the basis of reservoir levels. Since the measured water level is often not representative, and is also noisy, the quality of the hydrometric data may be doubtful, resulting in natural inflows that are often negative and water balances that do not agree. Table 1 illustrates the consequences of an error of 1 cm in the level reading of the main Hydro-Québec reservoirs on the esti-

Table 1 – Consequence of 1 cm error in level readings on the calculation of natural inflow to the main Hydro-Québec reservoir.

Reservoir	Equivalent natural inflow
Manic – 5	228 m^3/s
Caniapiscau	595 m^3/s
LG-2	325 m^3/s
Lac Cassé	94 m^3/s
Outardes 4	72 m^3/s
Mattawin	14 m^3/s
Gouin	224 m^3/s
Baskatong	38 m^3/s
Cabonga	51 m^3/s

mation of natural inflows obtained by the water balance equation. There is clearly a great deal of imprecision in the determination of natural inflows. Consequently, reservoir levels must be very carefully measured, and these readings must be validated.

Validation Methodology

Two different approaches are used to validate and correct the level time series according to the number of level gauges installed on the reservoir. Where the reservoir is equipped with only one level gauge, univariate filtering is used to smooth the series (Berrada *et al.* 1996). For example, frequency filtering will systematically eliminate all variations corresponding to frequencies higher than the predetermined threshold frequency (Van Den Enden and Verhoeckx 1992). In the most favorable case, where there are several gauging stations on the same reservoir, multivariate filtering proposed here is used to validate the level time series at each of the gauges. The underlying principle of this methodology involves relying more heavily on gauges that give readings more consistent with previous and subsequent validated values within a given series. Thus, the isolated, positive or negative, variations observed within that series will be eliminated in the case where corresponding values at other gauges exhibit more consistent variations. In order to evaluate the degree of persistence specific to a given reading, one may compare it to the predicted value obtained from an autoregressive model, with calibration based on the previous validated reading. As a result, this filtering technique constitutes a more "intelligent" alternative to the above-mentioned frequency filtering method. In more practical terms, this method compares, on the one hand, the deviation between the value predicted by an autoregressive model and the measured historical value, and on the other hand, the deviation between the value predicted by the same autoregressive model and the value estimated by a regressive model at other stations on the same reservoir. Among the values measured and estimated by the regressive model, that which is closest to the value predicted by the autoregressive model will be retained.

One feature of validation of historical hydrometric data is the availability of data recorded before and after the date of the value to be validated. In other words, if there are M successive values to validate the value of the level $L_{i\,t}$, we use not only the values $L_{i\,t-1}, L_{i\,t-2}, \ldots L_{i1}$, but also the values $L_{i\,t+1}, L_{i\,t+2}, L_{i\,t+M} \ldots$, both at the station to be validated and at neighboring stations. In order to exploit this feature, the process of validation by the multi filter method as described above will be executed twice, *i.e.* forward and backward in time. In the forward direction, the filtering process is executed from day 1 to day t. In the backward direction, the process is applied by starting with the last day, M, and ending with day t. To correct only values that are truly atypical, a historical value is rejected only if it has been set aside in both the forward and backward validation processes.

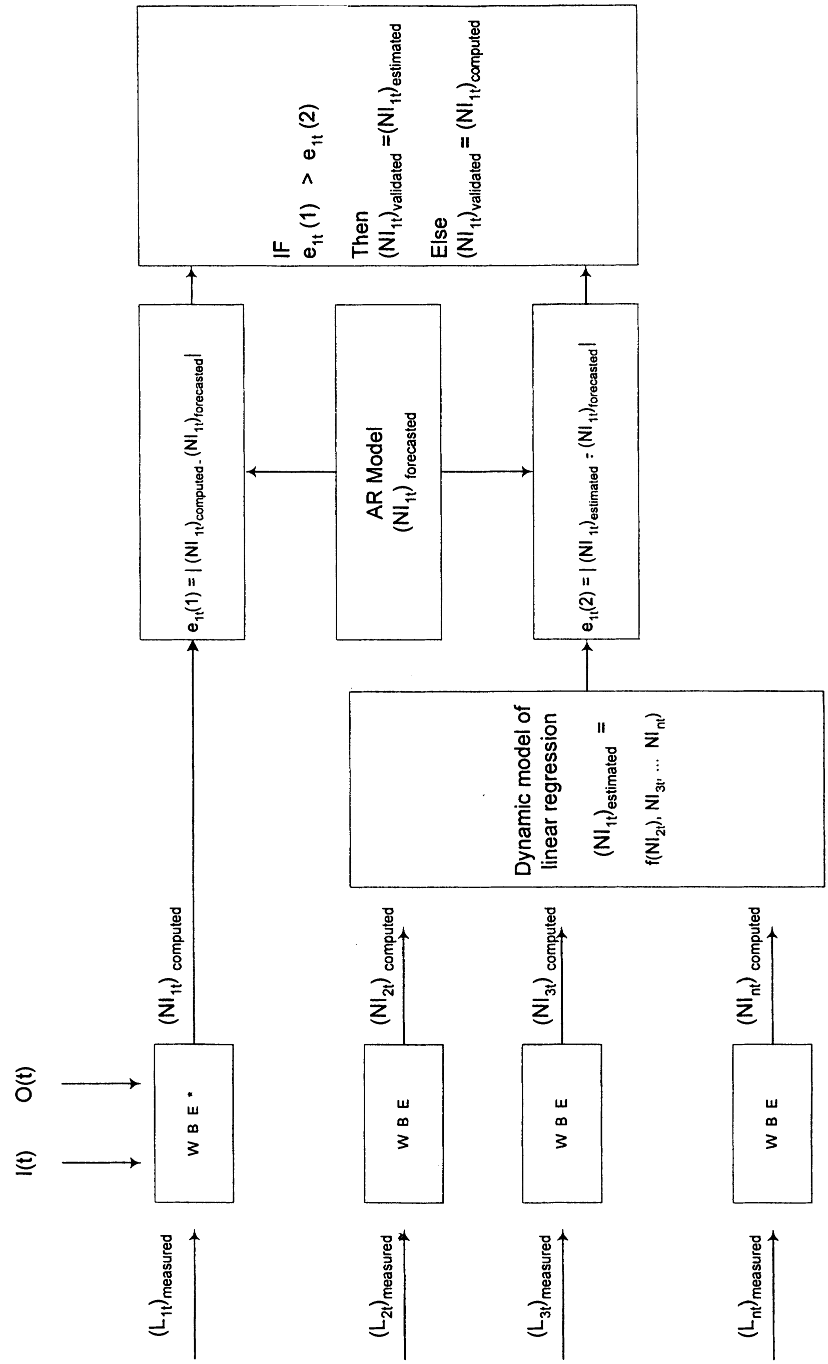

Fig. 3. Proposed voting process for diagnosis associated with gauge 1. *WBE - Water Balance Equation.

From a physical point of view, it may seem absurd to write that the level on day "t" depends on the level on days t+1, t+2,.... However, we must bear in mind that level fluctuations occur due to natural as well as artificial phenomena. Fluctuations resulting from natural inflows may well be represented by means of an autoregressive model in the forward direction of time, as long as the time interval used is reasonably small. On the other hand, artificial or sudden fluctuations, which result from the filling or emptying of the reservoir or from a rapid change in meteorological conditions, require the introduction of levels t+1, t+2, *etc*. For this reason, an autoregressive model in the backward direction of time is also considered.

The validation process described above may be applied in two ways. The first approach consists of applying the process directly to the time series of water levels $L_{i\,t}$ measured by different gauges situated on the same reservoir. Validated natural inflows and other hydrometric data may then be obtained by the water balance equation.

The second approach consists of applying the water balance equation separately to the water level measured by each gauge to obtain as many time series of natural inflows NI_{it} as time series of water level $L_{i\,t}$. The validation process is then applied to the series of natural inflows. Theoretically, the latter approach should yield more accurate results. In fact, the autoregressive model used to compute the trend of variation in the series before and after the data under validation is more appropriate for natural inflows than for water level elevation, which may be influenced by sudden variations in the evacuated flow. Fig. 3 illustrates the process of validation.

The methodology proposed here is nothing but an "intelligent" smoothing process. With univariate filtering, one operates a systematic cut of all fluctuations exceeding a threshold frequency in the series to be validated. In the proposed multifilter approach, fluctuations in readings from a gauge are only eliminated if the trend of variation in readings from other gauges is maintained.

In order to consider non-stationary models and to avoid the well-known bias of the least squares method, the Kalman filter will be used to identify the parameters of the retained models (Bennis and Bruneau 1993). In particular, the use of this filter makes it possible to improve the accuracy of the estimated model parameters and to obtain independent residuals (Fitch and Mc Bean 1991).

The Forecasting and Regression Models

The Forecasting Model AR – An ARMA model of type AR(2) is used to simulate and forecast variations in natural inflows (Box and Jenkins 1976). The systematic use of the AR(2) model is justified by its simplicity and the opportunity of using linear methods for estimating its parameters (Bennis and Assaf 1994). This model provides an accurate estimate of variations in natural inflows. The optimal order may be found for each case by using the Akaike Criterion (1974). As this model is used only to indicate the trend of variation in the natural inflows, and since we are concerned with automated methodology, we have arbitrarily set the order at 2. The AR formulation is adapted to the Kalman algorithm used to filter the model and measurement noises and to identify autoregressive parameters (Kalman 1960). The AR(2) model used as state equation is

$$X_t = \phi_{t/t-1} X_{t-1} + W_{t-1} \tag{7}$$

where

$$X_t = \begin{bmatrix} (NI^F_{it})_{\text{Forecasted}} \\ (NI_{it-1})_{\text{Validated}} \end{bmatrix}$$

$$X_{t-1} \equiv \begin{bmatrix} (NI_{it-1})_{\text{Validated}} \\ (NI_{it-2})_{\text{Validated}} \end{bmatrix}$$

and

$$\phi_{t/t-1} = \begin{bmatrix} \theta_1 & \theta_2 \\ 1 & 0 \end{bmatrix}$$ is the state transition matrix

$$W_t = \begin{bmatrix} W_t \\ 0 \end{bmatrix}$$ is a random process (input noise) accounting for the overall inaccuracy of the AR representation ($E(W_t) = \bar{W}$).

$(NI^F_{it})_{\text{Forecasted}}$ represents the forecasted natural inflow from gauge i at time t by an autoregressive model operating in the forward (exponent F) direction of time. $(NI_{it})_{\text{Validated}}$ represents the corresponding validated natural inflow.

To account for errors in water level measurements and then in natural inflow calculations, a second equation is used as measurement equation

$$Z_t \equiv H_t X_t + V_t \tag{8}$$

where Z_t is the computed natural inflow using the measured water level and V_t is a random process (output measurement noise) accounting for errors in estimating water lever measurements, inflows, and outflows ($E(V_t) = \bar{V}$).

The standard assumption in the foregoing formulation is that W_t and V_t are normally distributed white noise sequences with the following covariance matrices

$$E\left((W_t-\bar{W})(W_{t'}-\bar{W})^T\right) \equiv Q_t \hat{\delta}_{t't}$$

$$E\left((\bar{V}_t=V)(V_{t'}=\bar{V})^T\right) \equiv R_t \hat{\delta}_{t't} \tag{9}$$

where $E[.]$ designates the expectation, Q_t and R_t denote the covariance matrices of modeling noise and measurement noise respectively. The symbol $\hat{\delta}_{t't}$ stands for the Kroenecker delta and T denotes the transpose of a matrix. In addition, the two noises are assumed to be mutually independent, as expressed in the following equation

$$E\left((V_t-\bar{V})(W_i-\bar{W})^T\right) = 0 \tag{10}$$

$H_t = [\ 1\ 0\ ]$ is a scale matrix.

When validating NI_{it} in the backward direction of time, a similar model to Eq. (7) is used, although t-1, t-2, and $(NI_{it}^F)_{\text{Forecasted}}$ are respectively replaced by t+1, t+2 and $(NI_{it}^B)_{\text{Forecasted}}$.

The implementation of Kalman Filtering algorithm can be found in Bennis and Assaf (1994).

The Regression Model

To obtain the best estimation of natural inflows, one can use the linear regression model with variable coefficients described by Bennis *et al.* (1997). The natural inflow, NI_{it}, based on the water level measured at gauge i at time t, may be estimated by the following regression model

$$(NI_{it}^F)_{\text{Estimated}} \equiv \sum_{\substack{\ell=1 \\ \ell \neq i}}^{n} a_\ell^F NI_{\ell t} + e_t^F \tag{11}$$

Where $NI_{\ell t}$ represents the natural inflows based on gauge ℓ at time t. a_ℓ^F is the coefficient of regression associated with gauge ℓ and estimated by Kalman filtering when validating NI_{it} in the forward direction of time. These coefficients are calibrated recursively at each time step using validated natural inflow $NI_{\ell t-1}^F$... . e_t^F is the model noise at time t. Similarly, the regression model in the backward direction of time is

$$(NI_{it}^B)_{\text{Estimated}} \equiv \sum_{\substack{\ell=1 \\ \ell \neq i}}^{n} a^B NI_{\ell t} + e_t^B \tag{12}$$

where exponent B stands for backward direction of time.

Voting Process for Diagnosis

Assuming that there are M values of natural inflows computed from water level measured at gauge i, the validation of each value $(NI_{it})_{\text{Computed}}$; $1 \leq t \leq M$ is accomplished in two stages. In the forward direction of time, the validation of NI_{it} is executed after the validation of NI_{it-1}, Ni_{it-2},... *etc.* An autoregressive model calibrated on these validated values makes it possible to predict the natural inflow for time t to obtain $(NI_{it}^F)_{\text{Forecasted}}$. Let $(NI_{it}^F)_{\text{Estimated}}$ be the natural inflow obtained by Eq. (11)

and $(NI_{it}^F)_{\text{Validated}}$ be the validated natural inflow in the forward direction of time. Then the logical voting is

if

$$|(NI_{it}^F)_{\text{Forecasted}} - (NI_{it}^F)_{\text{Estimated}}| \leq |(NI_{it}^F)_{\text{Forecasted}} - (NI_{it}^F)_{\text{Computed}}|$$

then

$$(NI_{it}^F)_{\text{Validated}} \equiv (NI_{it}^F)_{\text{Estimated}}$$

else

$$(NI_{it}^F)_{\text{Validated}} \equiv (NI_{it}^F)_{\text{Computed}}$$

In the backward direction of time, the validation of NI_{it} is executed after the validation of NI_{it+1}, NI_{it+2}, ... *etc.* An autoregressive model calibrated on these validated values makes it possible to predict the natural inflow for time t, to obtain $(NI_{it}^B)_{\text{Forecasted}}$. Let $(NI_{it}^B)_{\text{Estimated}}$ be the natural inflow obtained by Eq. (12) and $(NI_{it}^B)_{\text{Validated}}$ be the validated natural inflow in the backward direction of time; then the logical voting is similar to the forward direction of time, although exponent F is replaced by exponent B.

The historical computed value $(NI_{it})_{\text{Computed}}$ based on the measured level L_{it} is not set aside unless it has been rejected by both the forward and backward validation processes. In that case, $(NI_{it}^F)_{\text{Estimated}}$ and $(NI_{it}^B)_{\text{Estimated}}$ are combined to give the optimal estimation of $(NI_{it})_{\text{Estimated}}$ by the equation (Bennis *et al.* 1997)

$$(NI_{it})_{\text{Estimated}} = W_1 (NI_{it}^F)_{\text{Estimated}} + W_2 (NI_{it}^B)_{\text{Estimated}} \tag{13}$$

where

$$W_i \equiv \sum_{j=1}^{2} \alpha_{ij} / \sum_{i=1}^{2} \sum_{j=1}^{2} \alpha_{ij} \tag{14}$$

where the α_{ij} terms are the elements of Σ^{-1}. Σ is the covariance matrix of forecast errors whose elements are of the form

$$\Sigma_{ij} = \text{Covariance}[e_t^i, e_t^j] \qquad \text{for } i \neq j$$

and

$$\Sigma_{ii} = \text{Variance}[e_t^i]$$

The matrix Σ is computed at each time step by a Kalman Filter (Bennis *et al.* 1996).

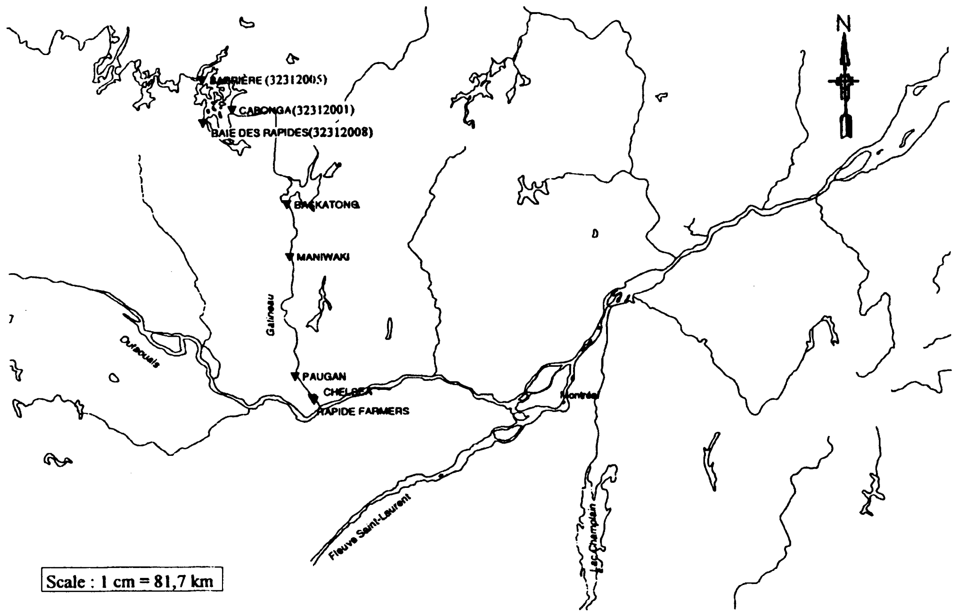

Fig. 4. Schematic of the Gatineau Complex.

Results

In order to test the ability of ValiDeb to validate the daily natural inflow series, Hydro-Québec suggested using the Gatineau River basin. This basin was chosen partly due to the fact that the Cabonga reservoir is equipped with three level-gauging stations, which makes it possible to test the multifilter model. This complex has already been used in previous validation attempts. The results obtained by automated validation using ValiDeb can therefore be compared with other techniques (Bérubé *et al.* 1987). Fig. 4 illustrates a schematic of the Gatineau complex and the location of the gauging stations. The three level-gauging stations of the Cabonga reservoir are located as follows: the first at the outflow near Baskatong (32312001), the second near Dozois (32312205), and the third in Baie des Rapides (32312008). The three level-gauging stations located on the same reservoir should theoretically give the same readings. However, under the influence of hydraulic variables and meteorological factors, or simply due to failed measuring equipment, there may be large deviations between the various measurements. Fig. 5 provides a comparison of the historical level variations of the Cabonga reservoir recorded by the three level-gauging stations. In addition to the unrelated fluctuations observed at some stations, there is a time varying scaling factor among the measurement series of the various gauges.

Fig. 6 illustrates the results of applying the proposed validation methodology directly to the series of water levels, which now present nearly the same variation and thereby the same validated natural inflows.

As shown in Fig. 5, one gauge (32312001) gave wrong measurements over most of the period from 12/11/78 to 3/21/79. The validated data shown in Fig. 6 have improved dramatically for measurements from the gauge. One may observe that some data have also improved for the other two gauges. For example, in the middle of Fig. 5, the disturbance in gauge 32312005, indicated by the arrow, has been corrected in Fig. 6. This indicates that even if two gauges fail, the third can be used to correct the observation.

Fig. 7 and Fig. 8 respectively compare historical level variations and computed natural inflows for the same reservoir for another period. Fig. 9 to Fig.11 compare natural historical computed inflow values for each gauge station using the water balance equation with the natural inflows validated by the technique described in pthe previous section. The proposed methodology of natural inflow validation eliminates virtually all unrelated fluctuations and undesirable negative values. To appreciate this rationally, the values of autocorrelation function were calculated for each natural inflow series before and after the validation process for different *K* lags by using the formula

$$R(k) = \frac{\sum_{t=1}^{N-k} \left((NI_t - \overline{NI})(NI_{t+k} - \overline{NI}) \right)}{\sum_{t=1}^{N} (NI - \overline{NI})^2} \tag{15}$$

Tables 2 to 4 show clearly that the values of the autocorrelation function increased significantly for different lags after the validation process. It is important to note that this significant noise reduction does not result from an arbitrary smoothing of data as in the univariate case (Berrada *et al.* 1996). It results naturally from an intelligent voting process (Fig. 3) for deciding which of the values provided by different gauges is the most representative of the natural inflow at each time *t*. This increase in the autocorrelation function is more consistent with the variation of natural inflow into a large reservoir. In addition, it will significantly reduce the forecasting variance by stabilising the numerical algorithm used for parameter identification.

The multifilter methodology has the advantage of balancing the water volume of natural inflow to the reservoir when applied over a relatively long period of time. In the case studied here, the ratios between the volume of historical natural inflows computed by the water balance equation at each gauge and the volume of validated natural inflow are respectively 0.997, 1 and 1.004. It is interesting to note that the flood peaks are not systematically overestimated or underestimated after the validation. In the case of the natural inflow hydrogram based on gauge 32312008, the validated peak flow is greater than the historical computed value. The opposite is true

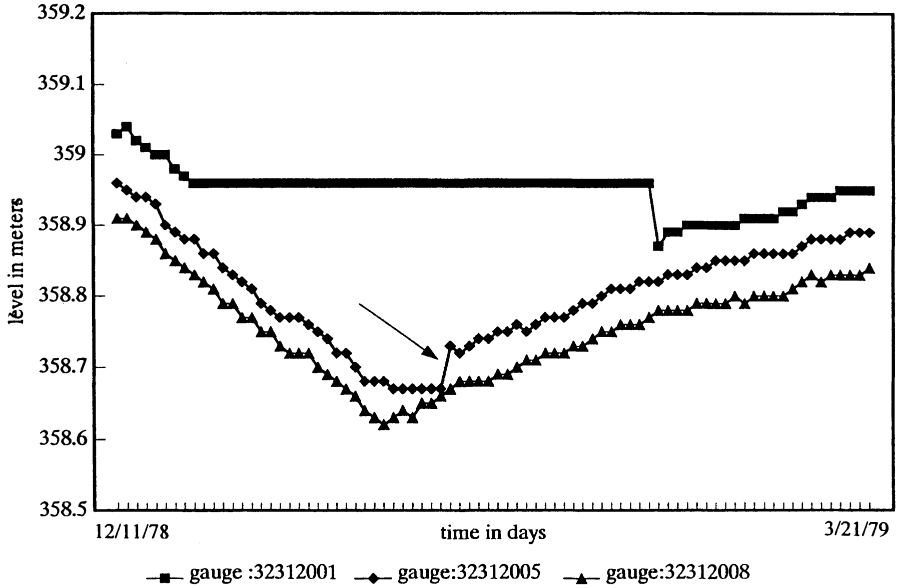

Fig. 5. Comparison between measured levels by gauges situated on the same reservoir.

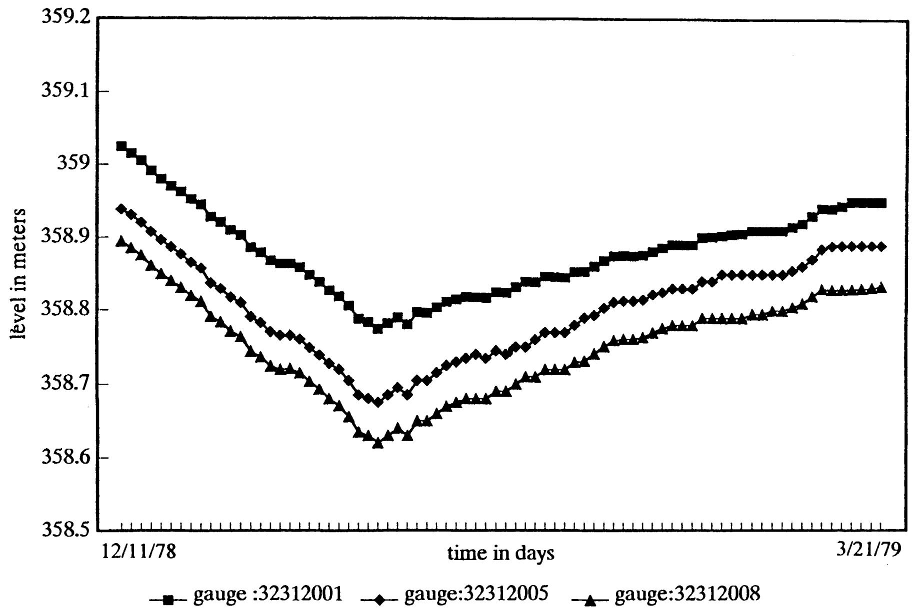

Fig. 6. Comparison between validated levels.

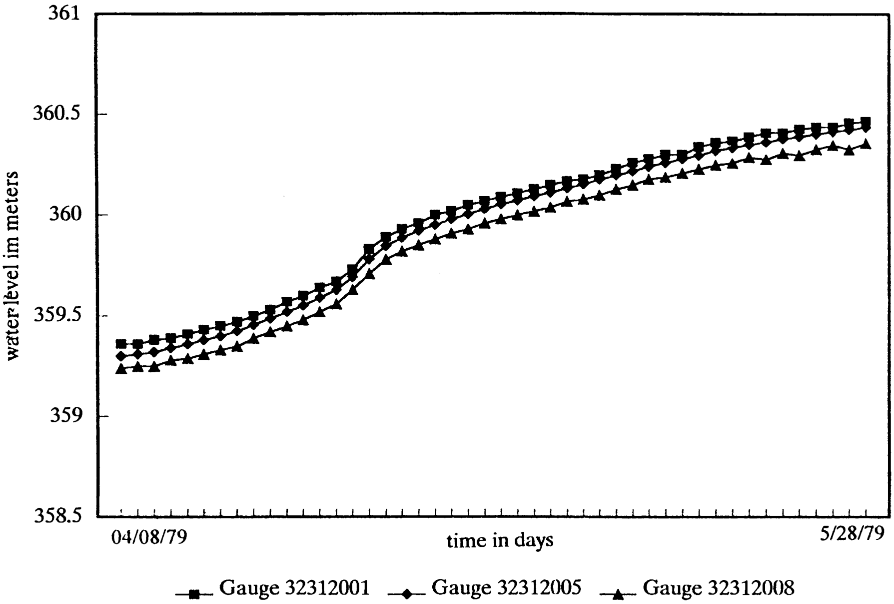

Fig. 7. Comparison between measured water levels by gauges situated on the same reservoir.

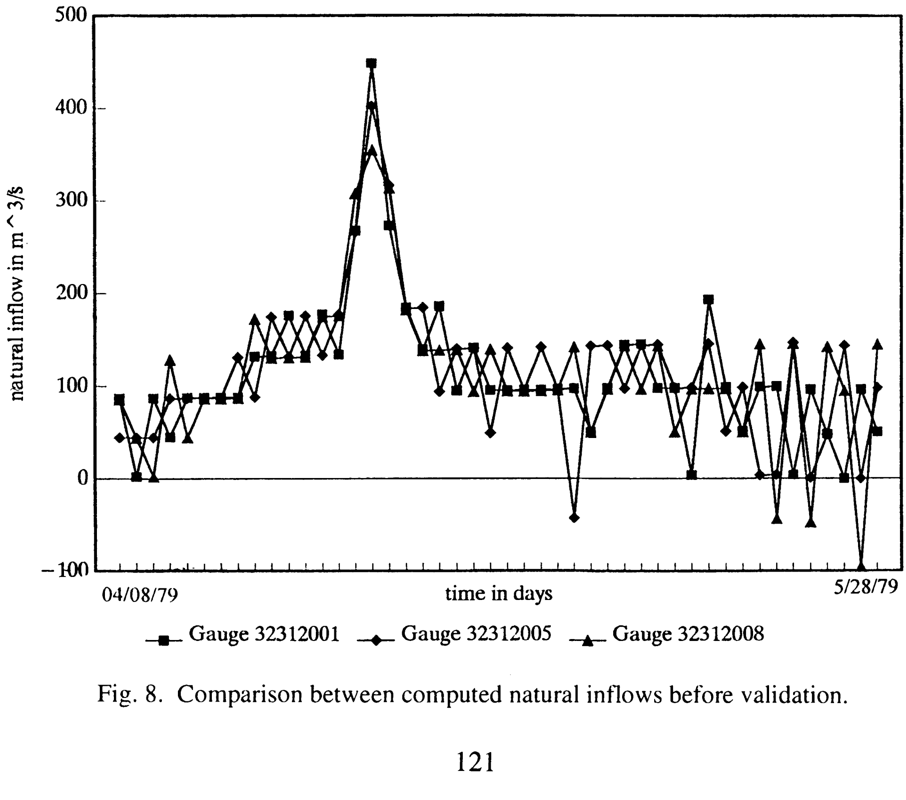

Fig. 8. Comparison between computed natural inflows before validation.

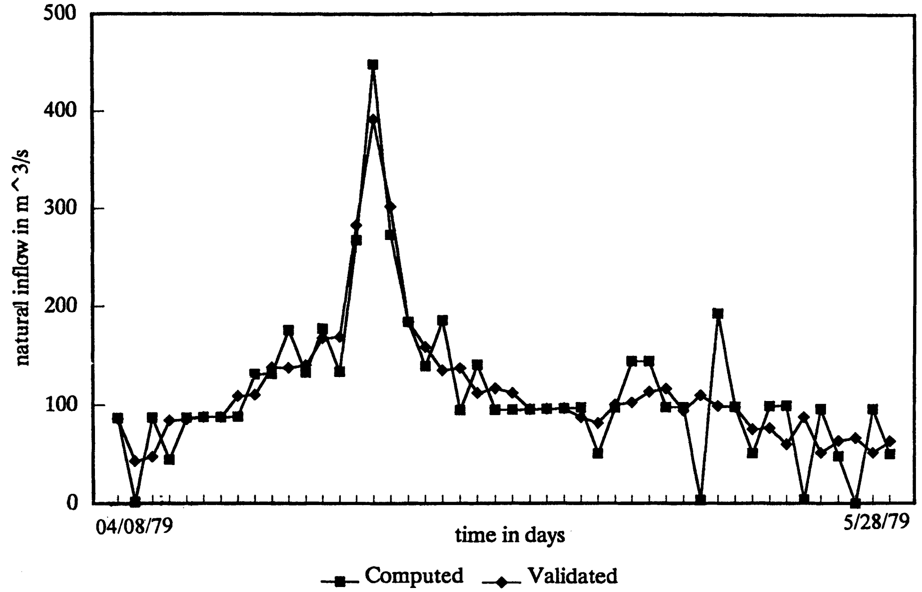

Fig. 9. Comparison between computed and validated natural inflows based on gauge 32312001.

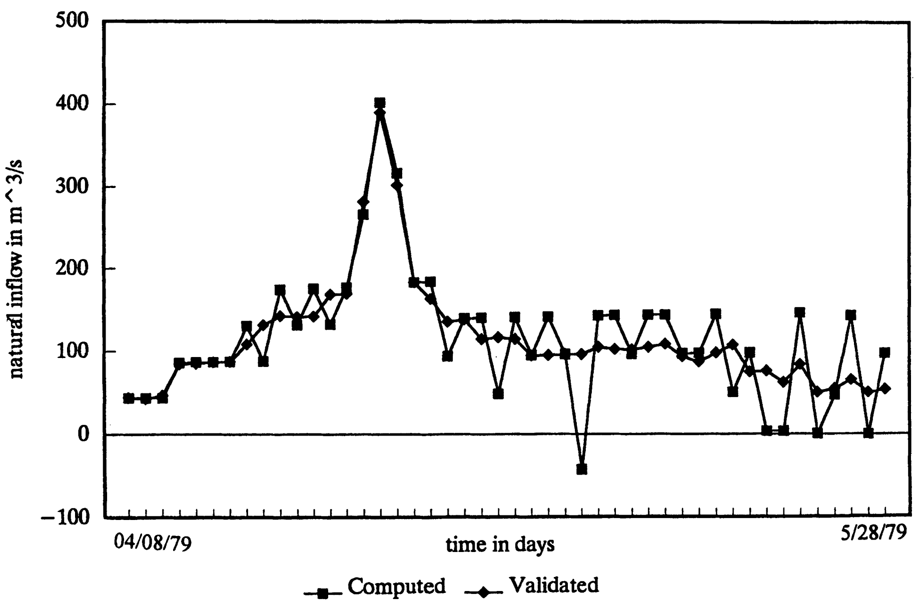

Fig. 10. Comparison between computed and validated natural inflows based on gauge 32312005.

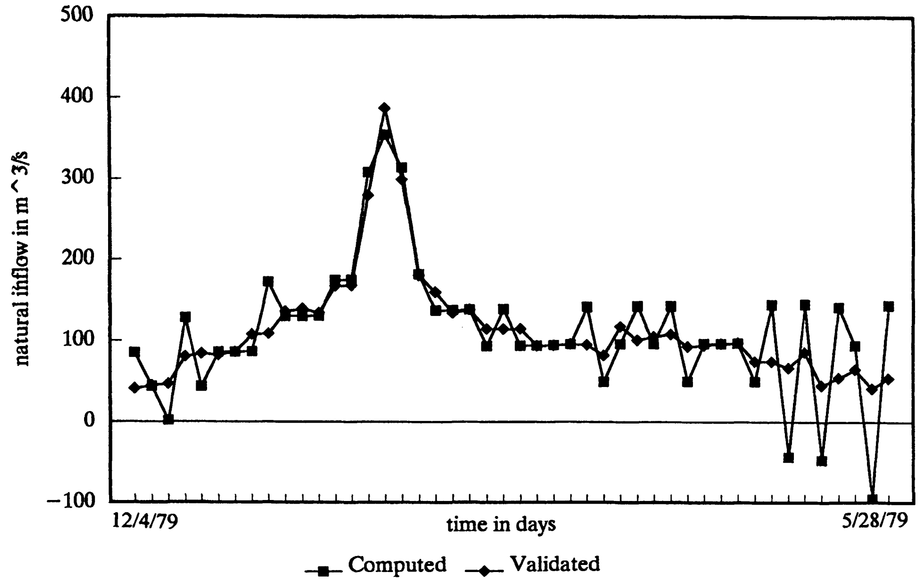

Fig. 11. Comparison between computed and validated natural inflows based on gauge 32312008.

Table 2 – Values of autocorrelation function of natural inflow series before and after validation for gauge 32312001

Gauge 32312001	R (1)	R (2)	R (3)	R (4)	R (5)
Before validation	0.556	0.415	0.353	0.273	0.203
After validation	0.846	0.637	0.484	0.363	0.247

Table 3 – Values of autocorrelation function of natural inflow series before and after validation for gauge 32312005

Gauge 32312005	R (1)	R (2)	R (3)	R (4)	R (5)
Before validation	0.545	0.445	0.441	0.188	0.12
After validation	0.876	0.693	0.547	0.443	0.318

Table 4 – Values of autocorrelation function of natural inflow series before and after validation for gauge 32312008

Gauge 32312001	R (1)	R (2)	R (3)	R (4)	R (5)
Before validation	0.407	0.513	0.39	0.212	0.310
After validation	0.881	0.741	0.587	0.468	0.334

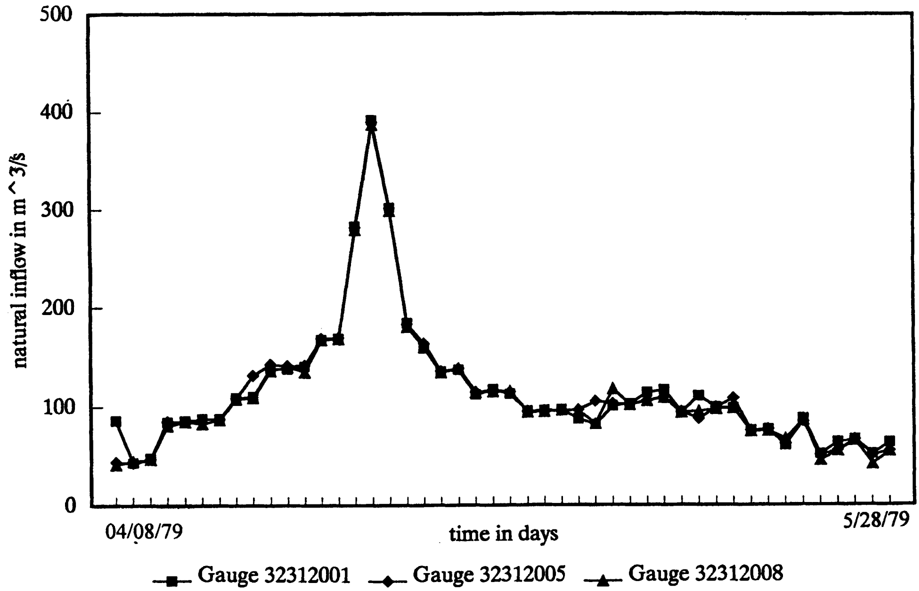

Fig. 12. Comparison between validated natural inflows.

for gauge 32312001. The validated peak flow based on gauge 32312005 is nearly identical to the historical computed value. This may be explained by the fact that the historical value of peak flow computed by this gauge is situated between the values computed by the other two gauges and corresponds approximately to the average of the three values. As illustrated in Fig. 12, which compares validated natural inflows, there are minor differences between the three hydrograms in the period of concentration and recession limb, but the peak flows are nearly identical (391.84 m^3/s, 389.94 m^3/s and 388.74 m^3/s respectively).

Conclusion

The objective of this project was to develop methodology for performing automated validation of natural inflow series. The current version of the software ValiDeb makes it possible to validate natural inflows by univariate and multivariate models. The proposed multivariate model is used when there are several level-gauging stations on the same reservoir.

The process of testing ValiDeb on the Gatineau River basin site confirmed that this software does indeed make it possible to detect and correct most of the anomalies, and in a completely automated fashion. The flood peaks are neither underestimated or overestimated, and the total volumes of natural inflows are retained.

Acknowledgements

We wish to thank the reviewers for their helpful comments and suggestions. This project was supported by a grant from the Natural Sciences and Engineering Research Council of Canada.

References

Akaike, H. (1974) A new look at the statistical model identification. IEEE Transactions on Automatic Control, AC-19, pp. 716 – 723.

Bath, L. (1985) Détection de défauts sur des capteurs redondants, application à une mesure de niveau d'un pressuriseur de centrale nucléaire, Mémoire CNAM automatique, France.

Bennis, S., and Bruneau, P. (1993) Comparaison de méthodes d'estimation des débits journaliers, *Canadian Journal of Civil Engineering, Vol. 20*, pp. 480-489.

Bennis, S., and Bruneau, P. (1993) Amélioration des méthodes d'estimation des débits journaliers, *Canadian Journal of Civil Engineering, Vol. 20*, pp. 490-499.

Bennis, S., and Assaf, G. (1994) Adaptive forecast of multi-month lake level elevations, *Canadian Journal of Civil Engineering, Vol. 21*, pp. 778-788.

Bennis, S., Berrada, F., and Kang, N. (1997) Improving single variable and multiple variable techniques for estimating missing data, *Journal of Hydrology, Vol. 191* (1-2), pp.

Berrada, F., Bennis, S., and Gagnon, L. (1996) Validation des données hydrométriques par des techniques univariées de filtrage, *Canadian Journal of Civil Engineering, Vol. 23*, pp. 872-892.

Bérubé, R., Charbonneau, R., and Dolbec, M. (1987) Validation des données hydrométriques historiques de la rivières des Outaouais, Hydro-Québec, internal report, pp.13-16.

Box, G.E.P., and Jenkins, G.M. (1976) *Time series analysis: forecasting and control*, Revised edition, Holden-Day, Oakland, pp. 46-82.

Fitch, M.A., and McGean, E.A. (1991) Multi-day flow forecasting using the Kalman Filter, *Canadian Journal of Civil Engineering, Vol. 18*, pp. 320-327.

Kalman, R.E. (1960) A new approach to linear filtering and prediction problems, Journal of Basic *Engineering (Transactions of the ASME, Series D), Vol. 82,* pp. 34-45.

Perreault, L., Roy, R., Bobee, B., and Mathier, L. (1991) Validation et estimation des apports journaliers, Scientific report No. 271, INRS Eau, Université du Québec, Canada, pp. 1-29.

Roy, R., Gauthier, R., and Hétu, M. (1992) Hydrometeorological Data Quality Improvement by Implementing and Autonomous Acquisition Network in Barren Areas, Hydraulic Engineering Software IV, Computer Techniques and Applications, Southampton Baston, pp. 351-357.

Van Den Enden, A.W.M., and Verhoeckx, N.A.M. (1992) *Traitement numérique du signal, Une introduction.* Paris, Milan, Barcelona and Bonn, pp. 222-231.

World Meteorological Organization (1985) Guidelines for Computerised Data Processing in Operational Hydrology and Land and Water Management, WMO No. 634.

Received: 14 September 1998
Revised: 8 June, 1999
Accepted: 3 December, 1999

Address:
Saad Bennis,
Départment de génie de la construction,
École de technologie supérieure,
1100 Notre-Dame Quest,
Montréal, Que H3C 1K3, Can.
Email: sbennis@ctn.etsmtl.ca

Narut Kang,
Directon ouvrages civils – Hydro-Québec,
75, Bould.rene Levesque Ouest,
Montréal, Que H2Z 1A4, Can.