Nordic Hydrology 5, 1974, 213–228
Published by Munksgaard, Copenhagen, Denmark

ADAPTATION AND APPLICATION OF THE KARAZEV METHOD TO THE RATIONALIZATION OF QUEBEC'S HYDROMETRIC BASIC NETWORK

M. LECLERC, B. BOBÉE and J. P. VILLENEUVE

Institut National de la Recherche Scientifique, Québec, Canada

In a research program carried out by I.N.R.S.-Eau, certain principles inherent to the rational development of hydrometric networks were applied to the Quebec network. This study describes an application of a method specific to the development of a basic network. The method used to establish the number of stations required is based on the actual knowledge of the spatial distribution and time variability of the mean annual runoff. A modified version of Karazev's (1968) method applied originally to a study of the hydrographic basins in the USSR was found to be particularly suitable to regions of little hydrologic information and particularly to Quebec.

In almost all countries, governments take inventories of available water resources. To take an inventory, stations are placed in different locations representative of the characteristics of the surface water regime. These networks of hydrometric stations are more or less ramified or diversified according to the needs and development of each region.

In a research program carried out by I.N.R.S.–Eau for Environment Canada (Villeneuve et al. 1972, 1973), certain principles inherent to the rational development of hydrometric networks were applied to the Quebec network (base, management, project stations . . .). This study describes an application of a method specific to the development of basic networks.

The objectives defined for this type of network are the supply of a minimum level of knowledge of the water resources at a given precision, and the establishment, with the help of permanent stations, of the long term hydrologic tendencies.

The method used to establish the number of stations required to reach these objectives is based on the actual knowledge of the spatial distribution and time variability of the mean annual runoff. The mean annual runoff is the least variable hydrologic characteristic, but the network established according to this data will also take into account the variability of the other characteristics. This method, published by Karazev (1968) in a study of the hydrographic basins in the USSR and adapted by the authors, is particularly suitable to regions with little hydrologic information. It is based on three points:

i) The gauged basins must be large enough for the measured runoff to be representative of the regional tendency;

ii) The regime stations must be sufficiently far apart for the specific runoff measured at two adjoining stations to be significantly different, so that repetition of information is avoided and costs are thus minimized.

iii) The stations must be close enough, however, that the runoff can be interpolated. The precision of the interpolation depends mainly on regional synchronisms of runoffs expressed as correlation functions, and on the distance (center to center) between the basins. The synchronism depends on the variability of both physiographic and meteorologic characteristics.

THEORY

Regional characteristics

This method rests essentially on the variability of the annual specific runoff ($ft^3/s/mi^2$ or $m^3/s/km^2$) which eliminates the influence of the area of the drainage basins on the runoff. This value is defined annually for each station (basin) in operation. The mean annual flow and cartography for the territory concerned must also be known (for Quebec, see Fig. 1).

A certain number of the principal regional characteristics are then calculated as follows:

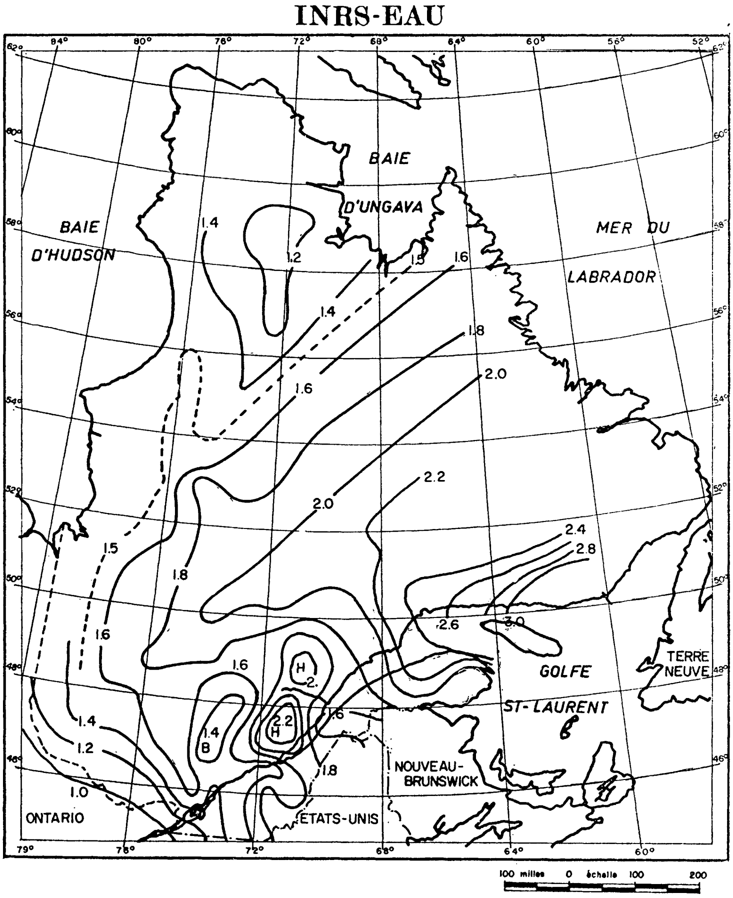

Fig. 1.
Map of mean specific annual runoff (ft ^{3}/s/mi^2).

i) the mean annual runoff

$$q_o^{*} \text{ reg.} = \frac{1}{A_{\text{reg.}}} \sum_{1}^{n} q_{oj} \cdot A_j \tag{1}$$

---

* See List of Symbols.

ii) the average gradient of runoff

$$\nabla \text{ reg.} = \frac{1}{A_{\text{reg.}}} \sum_{1}^{n} \frac{q_{o,j+1} - q_{oj}}{l_{j,j+1}} \tag{2}$$

iii) the coefficient of variation of the annual runoff

$$Cv = \frac{\sigma q}{q_o} \quad \begin{array}{l}\text{relatively constant} \\ \text{within the same} \\ \text{hydrologic region}\end{array} \tag{3}$$

iv) regional functions of runoff

$$q_i(\xi) \equiv q_o(\xi) + f_i(\xi) \ ; i \equiv 1, 2 \ldots . N \tag{4}$$

$f_i(\xi)$ expresses the random annual component of the runoff. The variability of the annual flow is usually represented by a normal law.

Within a homogeneous hydrologic region and for a basin not too sensitive to local conditions, the random component can be considered as almost homogeneous, i. e.

The arithmetic average is zero

$m_f = 0$

The standard deviation is constant within the region

$\sigma_f = \sigma_q = \text{cte}$

v) regional function of correlation

The fluctuations of the annual runoff are not always well synchronized from one basin to another. In fact, the degree of synchronism between the annual flow of two basins (expressed by the coefficient of correlation) depends on the distance between them. A good approximation of this relation can be obtained by the following expression verified for each region:

$$\varrho_q(l) \equiv 1 - al \equiv 1 - \frac{l}{L_o} \tag{5}$$

L_o is the approximate average distance for which the synchronism between two regional basins is zero. The correlation coefficient is calculated as follows:

$$\varrho_q(l) = \frac{1}{\sigma_q^2} \cdot \frac{1}{N} \cdot \sum_{1}^{N} f_i(\xi) \cdot f_i(\xi + l) \tag{6}$$

$$= \frac{\sigma_q^2}{\text{cov.}(l)} \tag{7}$$

Regionalization

Regionalization is defined on the basis of the homogeneity of a certain number of characteristics representative of the variability of the specific annual runoff within a certain part of the territory. A territory with homogeneous characteristics is classified as a region.

The characteristics considered are:

i) the spatial variability given by the gradient of runoff: ∇

ii) the time variability estimated by the coefficient of variation of the annual discharge: Cv

iii) the synchronism of the runoff of the different rivers depending on the average distance between them. This characteristic is expressed by the relation

$$\varrho_q = 1 - al$$

The criterion of minimum representative area: A_{min}

Drainage basins are chosen to have an area such that the specific annual runoffs are independent of the area and consequently dependent only on the physiographic and meteorologic characteristics of the basins. It is possible that the partition limit of the surface water of a basin does not coincide exactly with the partition limit of the underground water. As the area of basins is determined from topographic maps, this fact is not considered and consequently leads to errors in the calculation of the specific runoff. The error increases with decreasing size of the basin.

Thus the basins to be chosen as representative of the region must be sufficiently large that this situation does not arise. Consequently, a threshold of representativity written as a minimum area to be gauged

$$A > A_{min} \tag{8}$$

is calculated by a process of "step-wise" regression as follows:

$$q_o = k \cdot x_1^{a_1} \cdot x_2^{a_2} \cdots\cdots x_p^{a_p} \tag{9}$$

Concomitant data of the mean specific annual runoff in as many actual stations as possible are entered in the left-hand term of the regression. The parameters $x_1, x_2, \ldots. x_p$ correspond to the areas, the mean altitudes, the mean longitudes, the mean latitudes, the percentage area of lakes, the percentage area of forests, etc., of the basins respectively, these being the most important physiographic characteristics. This regression is calculated a certain number of times, sequentially eliminating basins smaller than 50, 100, 150, 200 .... square miles.

At each step, the additional percentage of variance explained by the inclusion of each x_i is noted in addition to the exponent a_i. To evaluate whether the specific runoff is independent of the surface of a basin, the variance explained by the inclusion of this variable must be zero or the exponent corresponding to the regression must be zero.

x_1 being the area, then

$a_1 \cong 0$

For this type of operation, basins are chosen within the same homogeneous hydrologic region in order to obtain a minimum threshold of representativity for each region.

Table 1 and Fig. 2 show the summary results of this regression applied to 78 hydrographic basins in Quebec. Because of the scarcity of data on small basins, the regression process has been applied to all regions together. However, when more complete regional data are available, the regression can again be applied to regionalize the value of A_{min}.

Table 1.

Results of the regression concerning the influence of the surface of the basin..

Minimum area	Variance explained by A	a_1: exponent of A or x_1	Number of basins	Standard error on a_1
mi	%			
0	0,025	–0,036	78	0,017
50	0,035	–0,042	76	0,018
100	0,0185	–0,033	71	0,020
150	0,0017	–0,11	68	0,022
200	0.0031	–0,015	66	0,022
250	0,0012	–0,0098	61	0,024
300	0,00078	–0,0078	54	0,025
350	0,00026	–0,0046	47	0,025

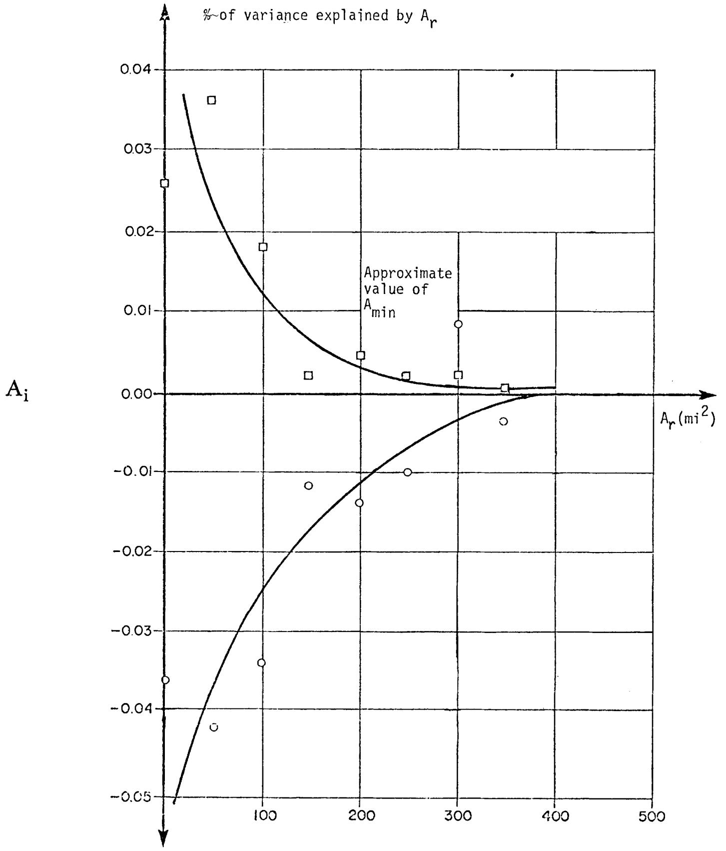

Fig. 2.
Results of each step of evaluation of the minimum representative area (A_{min}).

Criterion of spatial variability: A_{∇}

In the choice of the rivers to be gauged, the measured flows must be significantly different from one station to the other. Present knowledge of the distribution of the interannual specific runoff is utilized to establish the maximum density of stations in a region, i. e. in a region of area A_{reg}, not more than A_{reg}/A_{∇} stations are necessary.

The increase of the average annual runoff (in a direction normal to the isolines) from one point to the other on a region must be beyond, at a given probability, the average sampling error of the measured values at these two points.

The increase is expressed as follows:

$$\triangle\, q_0(\xi) \equiv \frac{dq_0(\xi)}{d\xi} \cdot \triangle\, \xi \tag{10}$$

$$\triangle q_0(\xi) = \nabla \cdot \triangle \xi \tag{11}$$

It can also be expressed as:

$$\triangle q_0(\xi) = q_0(\xi) - q_0(\xi + \triangle \xi) \tag{12}$$

If the dependence of the events $q(\xi)$, $q(\xi + \triangle\xi)$ is negligible, then

$$\sigma^2_{\triangle q_0} = \sigma^2_{q_0(\xi)} + \sigma^2_{q_0(\xi + \triangle \xi)} \tag{13}$$

As the hydrologic regions are defined according to criteria of homogeneity (σ_f, Cv, $\nabla, \ldots,$ N), then:

$$\sigma_q{}^2(\xi) \cong \sigma_q{}^2(\xi + \triangle \xi) \cong \sigma_q{}^2 \text{ reg.} \tag{14}$$

For samples of the same size, N

$$\sigma^2_{q_0(\xi)} \cong \sigma^2_{q_0(\xi + \triangle\xi)} \cong \sigma^2_{q_0 \text{ reg.}} = \frac{\sigma_q{}^2 \text{ reg.}}{N} \tag{15}$$

and the following relation is obtained:

$$\sigma^2_{\triangle q_0} = 2\, \sigma^2_{q_0 \text{ reg.}} \tag{16}$$

In order to respect the criterion outlined above:

$$\triangle q_0 = \nabla_{\text{reg.}} \cdot \triangle \xi \geqq k_1 \cdot \sigma_{\triangle q_0} \tag{17}$$

then:

$$\triangle \xi \geqq k_1 \cdot \frac{\sqrt{2 \cdot \sigma^2_{q_0 \text{ reg.}}}}{\nabla_{\text{reg.}}} \tag{18}$$

where k_1 expresses a standard deviation factor, which, in the case of a normal distribution, is related to the confidence interval.

Stations should be placed in such a way that the center of the gauged basins will form a quadrilateral with an average side of $\triangle\xi$.

$$A_\nabla \cong (\triangle \xi)^2 \geqq \frac{k_1{}^2 \cdot \sigma^2_{\triangle q_0}}{\nabla^2 \text{reg.}} \tag{19}$$

$$A_{\nabla} \equiv 2 \cdot k_1^2 : \frac{e^2_{q_0} \cdot q_0^2 \text{ reg.}}{\nabla^2 \text{reg.}} \tag{20}$$

and

$$A_{op.} > A_{\nabla}$$

In most cases, a value of 5 % is recommended, at each station, as an objective for the relative sampling standard error, and

$$A_{\nabla} \overset{=}{=} 2 \cdot k_1^2 \cdot \frac{(.05)^2 \cdot q_0^2 \text{reg.}}{\nabla^2 \text{reg.}} \tag{21}$$

Criterion of correlation: $A_{cor.}$

The criterion of spatial variability determines the density ($1/A_{cor.}$) which can yield reliable interpolation. However, a maximum distance between basins (center to center) must be established in order that the error of interpolation between the specific annual flows does not become too important. This error depends evidently upon the synchronism of the flows of different basins of a given region, which is expressed by a regional function of the correlation coefficient.

$$\varrho_q = 1 - a l = 1 - \frac{l}{L_0} \tag{22}$$

If a linear interpolation between the flows is carried out, it has been proved that this error will be maximum halfway between the geometrical centers of the two basins (Villeneuve et al. 1972) and can be expressed by the following relation:

$$\sigma_0^2 \text{int.} = \frac{1}{2} (Cv^2 : a \cdot l + e_0^2) \tag{23}$$

Now, this error should be less than a value expressed as a factor k_2 of the standard deviation e_0 of annual flow measurement error (generally, it is considered that the annual flows are measured with a relative precision of 5 %). Thus:

$$\sigma_0 \text{ int.} \leqq k_2 \cdot e_0 \tag{24}$$

combining (23) and (24)

$$l \leqq \frac{(2 k_2^2 - 1)^2 \cdot e_0^2}{a \cdot Cv^2} \tag{25}$$

and since $A \cong l^2$

$$A_{op.} \leqq A_{cor.} = \frac{(2\,k_2^2 - 1)^2 \cdot e_0^4}{a^2 \cdot Cv^4} \tag{26}$$

Finally, three limit areas are obtained that should be applied to optimize the territorial distribution of basic network stations.

$$A_{min} \leqq A_{\nabla} \leqq A_{op.} \leqq A_{cor.}$$

LOCATIONS OF STATIONS

The number of stations to be established within a given region can be determined by dividing the total surface of the region by the optimal surface ($A_{op.}$). As much as possible, these stations should be spread uniformly throughout the region.

Every regional network should include a certain number of stations placed on large basins; these basins should be one and a half times the size of the optimal surface. Some stations should be placed on small basins; it is suggested that the number of stations on small basins should represent 15 % of the total number of stations.

Finally, certain stations should be planned to operate continuously. Their purpose will be to specify the long term tendencies of the hydrologic regime. The other stations of the representative network should be in operation during a common period of at least 10 years as a first step. The constraints of rationalization could then be re-evaluated periodically, allowing a continual adaptation and improvement of the method described.

RESULTS

For the purposes of this study, Quebec has been divided into 18 hydrologic regions. Fig. 3 shows the delimitation of these regions and Table 2 shows the corresponding values of the parameters of regionalization. The first objective recommended is that the error of interpolation on the annual flow shall have a maximum value of 6 %, and that the differences between the measured flows, at two consecutive stations, must be significant with a confidence interval less

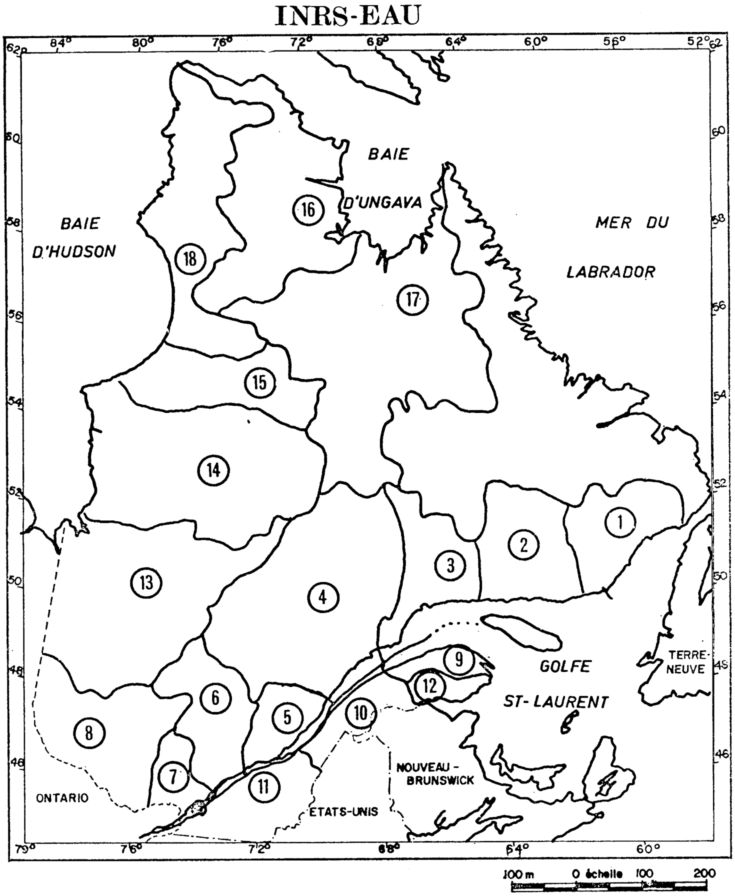

Fig. 3.
Regionalization as defined by Karazev's method.

than 68 %. If this first objective is not attainable, an error of interpolation of 7 % and a 80 % confidence interval is recommended. Table 3 shows the results of the criteria calculated according to the different objectives placed upon the confidence interval $(k_1 \cdot \sigma_{\Delta q_0})$, and on the error of interpolation between the stations $(k_2 \cdot e_0)$.

Table 2.
Regional characteristics.

Region		A	Cv	∇ rég	q_0 rég	L_0
		mi^2*	[ ]	$cfs/mi^2/mi$*	cfs/mi^2*	mi*
1	Saint-Laurant and Baie des Chaleurs tributaries	26 400	–	–	–	–
2		18 700	0,144	0,0084	2,59	275
3		21 000	0,144	0,0044	2,34	275
4		65 800	0,107	0,0043	1,96	300
5		9 360	0,145	0,026	1,98	75
6		18 300	0,133	0,0109	1,60	100
7		8 890	0,168	0,0101	1,45	325
8		32 000	0,171	0,0045	1,36	325
9		4 380	0,176	0,0086	2,16	100
10		6 220	0,21	0,0060	1,69	210
11		15 800	0,185	0,0071	1,60	150
12		6 290	–	–	2,00	–
13	James Bay, Hudson's Bay and Ungava's Bay tributaries	67 700	0,173	0,00246	1,77	430
14		68 500	0,191	0,00263	1,75	325
15		22 700	0,18	0,00201	1,58	–
16		43 900	0,132	0,00208	1,25	240
17		92 000	0,127	0,00263	1,73	240
18		43 500	–	–	–	–

* 1 mi ≡ 1,6093 km
1 mi^2 ≡ 2,59 km^2
1 $ft^3/s/mi^2$ ≡ 0,0109 $m^3/s/km^2$
1 $ft^3/s/mi^2/mi$ ≡ 0,00677 $m^3/s/km^2/km$

Table 3.
Value of the criteria of rationalization.

Region	A_{min}(mi^2)	A_{∇}(mi^2)		A_{cor}(mi^2)		A_{op}(mi^2) chosen
		$k_1 \equiv 1,0$ $\beta \equiv 68\%$	$k_1 \equiv 1,3$ $\beta \equiv 80\%$	$k_2 = 1,2$ $\sigma_{int} = 6\%$	$k_2 = 1,4$ $\sigma_{int} = 7\%$	
1	250					
2	250	500	812	3 890	11 000	3 890
3	250	1 300	2 400	3 890	11 000	3 890
4	250	1 000	1 760	15 200	>>>	15 200
5	250	<<	60	281	600	281
6	250	<<	182	2 550	6 600	2 550
7	250	<<	176	2 930	7 600	2 930
8	250	450	776	2 730	7 050	2 730
9	250	200	543	230	550	230
10	250	200	679	500	1 300	500
11	250	175	433	424	1 200	424
12	250	200	465	1 510	4 300	1 510
13	250	2 400	4 400	4 000	11 000	11 000
14	250	2 100	3 800	1 600	4 000	4 000
15	250	2 800	5 200	–	–	5 200
16	250	1 700	3 200	3 600	10 000	10 000
17	250	2 000	3 600	4 800	10 300	10 300
18	250	–	–	–	–	10 000

DISCUSSION

The large surface area of Quebec and the non-availability of sufficient hydrological data restricted the applicability of Karasev's original method to this region. Modifications of the method were thus developed to permit the use of present hydrological data on Quebec.

To determine the value of the spatial representativity factor (A_{min}), mean values of physiographic characteristics at hydrometric stations have been used. Due to the scarcity of hydrological data on a sufficient number of small watersheds, this factor could not be calculated on a regional basis.

The value of the correlation parameter (A_{cor}) could be calculated for many regions only on short series of hydrological flow reports (8 years). However, due to the large number of stations utilized, the results are judged to be satisfactory and permit, in spite of the correlation coefficient errors due to the short term data, a precise determination of the factor L_0 characteristic of regional synchronism.

Karazev's method has also been adapted to include predetermined precision objective factors (k_1, k_2) related to system development costs. This modification thus permits consideration of inevitable budget restrictions.

The application of this method to the territory of Quebec allows a definition of precise standards for the expansion of the basic network within the framework of predetermined principles. A rational development of this system is an important part of a well-planned global management of the national resource.

LIST OF SYMBOLS

A:	surface area (mi^2 or km^2)
$A_{cor.}$, A_{min}, $A_{op.}$, $A_{\bigtriangledown}$:	criterion of area (mi^2 or km^2)
A_j, Q_{ij}, Q_{oj}, q_{ij}, q_{oj}:	characteristics defined for the basin or the isoline j
$A_{reg.}$, $q_{o\ reg.}$, $Cv_{reg.}$:	regional characteristics
A_r:	minimum limit of the area of the basins included in the regression process
a:	inverse of L_0 (mi^{-1} or km^{-1})
$a_1, a_2, a_3, \ldots a_n$:	exponents of $x_1, x_2, \ldots\ x_n$ in a regression
cov (*l*):	average covariance of the flow of two basins with "*l*" distance between them
Cv:	variation coefficient of the annual module

e_0:	standard deviation relative to the error of measurement of the specific measurement error
e_{q_0}:	relative standard error of the mean annual runoff (specific)
$f_i(\xi)$:	deviation for the year i of the specific annual runoff with regard to the mean
i:	year i
j:	basin or isoline j
k_1:	factor of the standard deviation $\sigma_{\triangle q_0}$
k_2:	factor of e_0
L_0:	theoretical distance where the synchronism of the annual flow of any two basins is zero (mi or km)
l:	distance between two basins' geometrical centers (mi or km)
m_f:	average of $f_i\ (\xi)$
N:	number of years of observation
n:	number of isolines
Q_0:	average annual runoff (ft^3/s, m^3/s)
Q_i:	annual runoff (ft^3/s, m^3/s)
q_0:	mean specific annual runoff ($ft^3/s/mi^2$ or $m^3/s/km^2$)
q_i:	specific annual runoff ($ft^3/s/mi^2$ or $m^3/s/km^2$)
$x_1, x_2 \ldots, x_n$:	average physiographic characteristics of a basin
ξ:	system of regional coordinates
$\varrho_q(l)$:	function of correlation between two annual series
$\sigma_{0\ \text{int.}}$:	relative standard deviation of the interpolation error
σ_f, σ_q:	standard deviation of the specific annual runoff
σ_{q_0}:	standard error on the mean annual runoff (specific)
∇:	gradient of runoff ($ft^3/s/mi^2/mi$ or $m^3/s/km^2/km$)

REFERENCES

Karazev, J. F. (1968) Principles for distribution and prospects for development of a hydrology network. *Soviet Hydrology 6,* 560–568.

Villeneuve, J. P., Bobée, B., St.-Martin, H. & Slivitzky, M. (1973) Méthodologie de rationalisation des réseaux hydrométriques. *Bulletin de l'AISH 18,* (4).

Villeneuve, J. P., Leclerc, M. et al. (1972) Rationalisation du réseau hydrométrique du Québec. Etude par INRS-Eau pour Environment Canada.

Received 2 February 1974.

Address:
Institut National de la Recherche Scientifique, (INRS-Eau),
P. O. Box 7500,
Québec 10, Qué,
Canada.