Contribution to the planning of the station network for measuring the chemical composition of precipitation

By I. L. KAROL and L. T. MYATCH, *Institute of Experimental Meteorology, Obninsk, USSR*

(Manuscript received April 26, 1972)

ABSTRACT

The data on the chemical composition of precipitation obtained during the period from 1958 through 1965 at 28 stations situated on the European Territory of the USSR (ETU) are used to calculate correlation functions of $Ca^{\cdot\cdot}$, $Na^{\cdot}$, SO_4'' ion and ion sum concentrations in precipitation and deposition. Correlation distances for various ions are found to be in the range of 90 to 450 km. The mean square random errors in ion composition measurements (35–45%) are estimated using calculated correlation functions. It is shown that the mean station separation of about 100 km gives the error of optimum interpolation to the point for the total ion concentration in precipitation equal to the random error of observations. The optimum averaging of the deposition value over some area can give the same error, 2–3 stations being in the square of 300 × 300 km².

Ever increasing human activities have resulted in considerable global air pollution from industrial and agricultural sources in the last few decades. Further increase in pollution can lead to direct danger to people and environment and to inadvertent weather and climate modification (SCEP Report, 1970).

In this connection the World Meteorological Organization recommends that the existing stations should be supplemented with new stations, and all of them should constitute an international network. This network should include "baseline air pollution stations" and "regional air pollution stations". Baseline stations are aimed at measuring, to a high accuracy, as many minor constituents as possible (and mainly CO_2) to determine long-term changes in their content. Regional stations are intended for the conduct of simpler and rather well elaborated measurements of atmospheric turbidity and chemical composition of precipitation to document changes in the atmospheric composition caused by local conditions. Location and performance of both baseline and regional stations should be unaffected by local sources of pollution (WMO, 1971).

The International Commission for Atmospheric Chemistry and Radioactivity (CACR) recommends the establishment about 10 baseline stations. To measure more variable constituents partly of a regional origin it is necessary to have a more dense network of regional stations, but the criteria for the required density of this network have not been established yet. CACR recommends (CACR, 1970), without evident objective grounds, about 100 regional stations situated at a distance of about 1 000 km from one another. At the present time, in meteorology, some methods already exist to determine the optimum density of station networks using the statistical structure of fields measured (Gandin, 1963, 1968). Such data can be obtained, having a large volume of observations from a rather dense network.

In the network of stations covering large territories, of all the data on the global air pollution, a great volume of data is available only on the chemical composition of precipitation (Drozdova et al., 1964; Ericksson, 1970; Selezneva, 1970).

The present paper gives the results of pilot studies of the statistical structure of the fields of the mean-monthly concentration and deposition with precipitation of some chemical constituents using data from the ETU stations for the period of 1958–1965 (Drozdova et al., 1964; Selezneva, 1970). The objective estimates have been obtained for the minimum number

of stations to measure fields of contaminant deposition with precipitation which would satisfy some practical requirements (interpolation of data to any point and representation of total constituent deposition over a given area with a specified accuracy).

For the statistical analysis we use the mean-monthly concentrations in precipitation (q) of SO_4'' anions, $Ca^{\cdot\cdot}$ and $Na^{\cdot}$ cations as well as the concentration of all the ion sum and their monthly deposition (s) at 28 stations situated at the ETU (Petrenchuk & Selezneva, 1970). These constituents are chosen as representative of the main sources of air pollution: SO_4'' mostly characterizes the pollution resulting from fossil fuel combusion and $Ca^{\cdot\cdot}$ and $Na^{\cdot}$ point mainly to the natural aerosol of continental and maritime origin, respectively (Drozdova et al., 1964; Ericksson, 1970; SCEP Report, 1970; WMO, 1971). Monthly data from the ETU stations for the above four constituents for the warm half-year (April–September) of 1958–1965 comprise 48 cases at 11 stations (measurements have been conducted since January, 1958) and 24 cases at an additional 17 stations (measurements have been carried out since 1962) (Drozdova et al., 1964; Petrenchuk & Selezneva, 1970).

The mean annual distribution of total ion concentrations in precipitation over the territory of the USSR from the data of network considered is given by Petrenchuk & Selezneva (1970). At the ETU the latitudinal distribution is predominant and concentrations increase from North to South. Concentrations of individual ions in precipitation also increase from North to South. For ions of maritime origin alone q increases in seaside regions as well (Drozdova et al., 1964).

The correlation functions have been calculated according to the procedure given in the paper by Gandin & Boltenkov (1964), which is especially convenient for a limited number of measurements.

For the calculations the data of ion concentrations in precipitation and of their deposition per km^2 (s) per warm half-year are used. To obtain s, concentration q values have been multiplied by the amount of precipitation (h). Calculations of variance at all the stations show that variance of $Ca^{\cdot\cdot}$ and SO_4'' concentrations changes from station to station by a factor of 5 and the total ion concentration by a factor of 7. Variance of deposition changes by a factor of 10 and 20 for SO_4'' and the sum of ions, respectively.

Thus, fields of q and s are greatly variable as well as fields of other meteorological elements (Gandin, 1963). Therefore, just as with other elements, normalized deviations of q and s are analysed statistically

$$\varphi_{ij} = \frac{f_{ij} - \bar{f}_j}{\sigma_j} \qquad \text{(I)}$$

where f_{ij} is the individual value of an element at the jth station, $\bar{f}_j$ is the value averaged over all the data available for the station and σ_j is the mean-square deviation. The fields of normalized deviations are more homogeneous because their variance is equal to unity everywhere. Values of the correlation function for some distance are determined by the averaging of φ_{ij} for all the pairs of stations falling into a given range of distances (in our case this range increment is equal to 200 km). In all the ranges, except the first (0–200 km), 350–500 cases are averaged.

For the first range the averaging is performed over 60–70 cases since on the ETU there are few stations with a relative distance less than 200 km.

Random observational errors are determined by extrapolation of correlation functions to zero point by the procedure described by Gandin et al (1968) and Guschina et al. (1967). These values are then subtracted from the calculated correlation functions. The random observational errors (defined as the ratio of random errors squared to the element variance) prove to be approximately equal to 0.15–0.20. The rms relative random error constitutes about 35–45 % respectively, which is much higher than the instrumental error (10–15 %) (Drozdova, et al., 1964). High values of random errors are explained by great variability of fields considered.

Fig 1 gives some of the calculated correlation functions $\mu_q(z)$ and $\mu_s(z)$, the random observational errors having been subtracted from them. The correlation function of the monthly liquid precipitation amount which is calculated from the same data is also given for comparison. Analysis of calculated functions shows that the correlation functions $\mu(z)$ of various contaminants decrease rapidly as the distance z in-

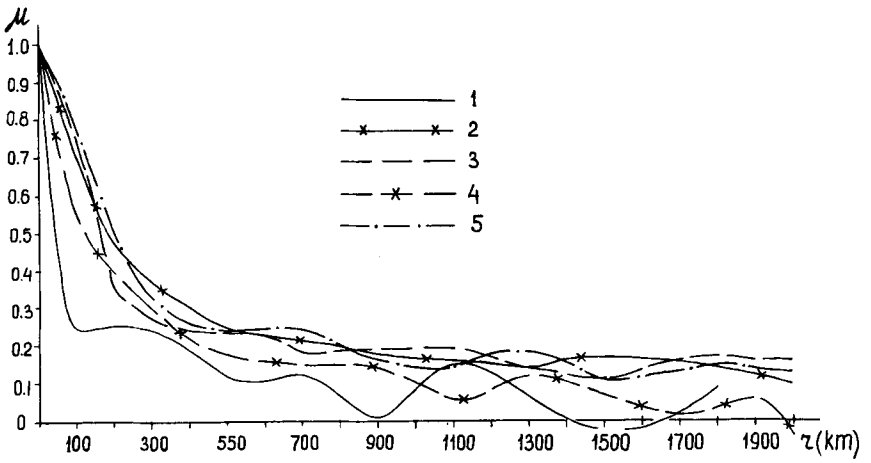

Fig. 1. Correlation functions of mean-monthly ion concentration (1, Ca; 2, Na); correlation functions of monthly deposition of ions with precipitation (3, SO_4; 4, Ca); correlation function of monthly sums of precipitation amount (5).

creases. This refers especially to the correlation function of $Na^{\cdot}$ ions, characterizing sea aerosol. The difference between the correlation functions of $Ca^{\cdot\cdot}$ and SO_4'' is only observed for distances of about 100 km. The correlation functions of the total ion concentration in precipitation is close to that of precipitation amount.

For all the contaminants the correlation functions of concentration $\mu_q(z)$ and deposition $\mu_s(z)$ are close to each other up to the distances of about 100–200 km. At greater distances $\mu_q(z)$ decreases less rapidly than $\mu_s(z)$. $\mu_s(z)$ and $\mu_q(z)$ of the sum of ions differ both at small and at great distances.

These conclusions are confirmed by the obtained correlation distance as well (Table 1).

The greatest change of $\mu_q(z)$ and $\mu_s(z)$ is observed in the 0–300 km range where their values are less reliable because, as mentioned above, fewer cases are available for this range. Therefore the conclusions drawn here and below can be improved by using data from a more dense network of stations for statistical analysis.

Table 1. *Correlation distances (in km) for concentration of various constituents in precipitation and their deposition*

	$Ca^{\cdot\cdot}$	So_4''	$Na^{\cdot}$	Sum of ions	Precipitation amount
R_q	270	300	120	450	430
R_s	240	250	90	330	

The obtained correlation functions and estimates of observational errors have been used to evaluate errors of optimum interpolation ε^2 of q values measured at n stations to a given point ($j = 0$) from the known formulas (Gandin, 1963):

$$\varepsilon^2 = 1 - \sum_{j=1}^{n} p_j \mu_{j0}, \quad j = 1, 2, \ldots, n \tag{2}$$

where optimum weights of stations p_j yield the set of equations:

$$\sum_{i,\,j=1}^{n} \mu_{ij} p_i + \eta p_j = \mu_{j0} \tag{3}$$

Here μ is the normalized correlation function, η is the measure of an observation error.

Fig. 2 gives results of error calculations at optimum interpolation to the centre provided that 1 to 8 stations are uniformly distributed along the circumferences of various radii (from 30 to 57 km). With the help of this Figure one can determine what the distance should be between the centre (control point) and the circumference with the stations to obtain a given interpolation error. An interpolation error approximately equal to the observational error (0.20) is obtained if the interpolation is carried out for 2, 3, 4 and 6 stations situated at a dis-

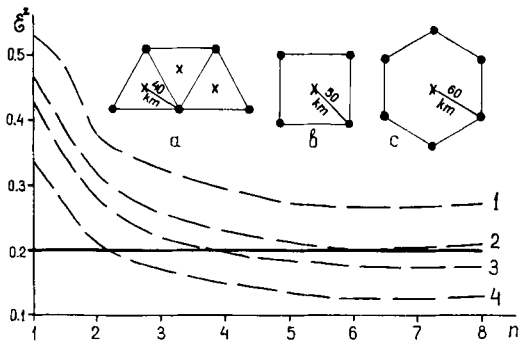

Fig. 2. Dependence of the error of the optimum interpolation to the point (ε^2) of the mean-monthly concentrations on the number (n) of the interpolation stations uniformly distributed along the circumferences of various radii: (1) 75 km; (2) 60 km; (3) 50 km; (4) 30 km and variants of the station's location in the triangular, rectangular and hexagonal grid points.

tance of 30, 40, 50 and 60 km, respectively, from the control point (Fig. 2). The use of more than 6 stations for interpolation does not lead to a noticeable decrease in the interpolation error.

The mean distance between the stations in the second, third and fourth cases (in the order given above) is about 50, 85 and 100 km, respectively. Thus, 100 km is the optimum distance between the stations at which the contaminant concentration in precipitation is determined at any point of the territory with the error not greater than the observational error.

Vinnikov & Dvorkina (1970) found 500 km to be the maximum permissible distance between the actinometric stations on the basis of analogous estimations of the interpolation accuracy for monthly sums of total radiation. Probably, the same distance will do for stations intended for measuring the atmospheric turbidity coefficient by means of actinometric devices. Thus, the density of regional air pollution stations is defined by the requirement for representativity of the chemical composition of precipitation and not that of the atmospheric turbidity. This is due to the fact that the random fields of the chemical composition of precipitation are more variable than the actinometric fields.

To study the deposition of various contaminants with precipitation it is more important to have the data on the mean total deposition over a definite area than at individual points. Thus, in this case it is more expedient to determine the necessary network density for the prescribed accuracy in the determination of mean values over area (Gandin, 1968; Kagan, 1967). This error for chemical contaminants in precipitation is estimated with the help of the optimum averaging method which is analogous to the optimum interpolation method. In this method the data from stations situated in the considered area are used, their weights being optimum in the sense of least squares. To determine these weights it is necessary to have the data on the statistical structure of not only point values of the element but also of those averaged over this area.

The autocorrelation functions of these values and cross-correlation functions of "point" and "area" values for various areas of averaging are obtained by the numerical integration of the point correlation function $\mu_s(z)$ of the sum of ions. The square with side d is taken as the area of investigation. These correlation functions of contaminant content in precipitation for small distances z prove to be a little lower than the empirical "point" correlation function. The larger the area of averaging, the greater the difference between them, but beginning with $z \geqslant d$ they practically coincide for all values of d. The mean square of the error ε^2 which arises when the actual mean deposition over some area is changed by the weighted mean deposition determined from measurements at n stations can be obtained as follows (Kagan, 1967)

$$\varepsilon^2 = l_0(d) - \sum_{j=1}^{n} p_j \omega_j(d), \quad j = 1, 2, \ldots, n \qquad (4)$$

where p_j is the optimum weights in the sense of least squares which are determined from the following set of equations:

$$\sum_{ij}^{n} \mu_{ij} p_i + \eta p_j = \omega_j(d) \qquad (5)$$

Here μ_{ij} is the correlation function constructed from the point data; $l_0(d)$ is the autocorrelation function of the values averaged over area; $\omega(d)$ is the cross-correlation function of the point values and values averaged over area.

To establish the representativity of the existing network of stations at which the total de-

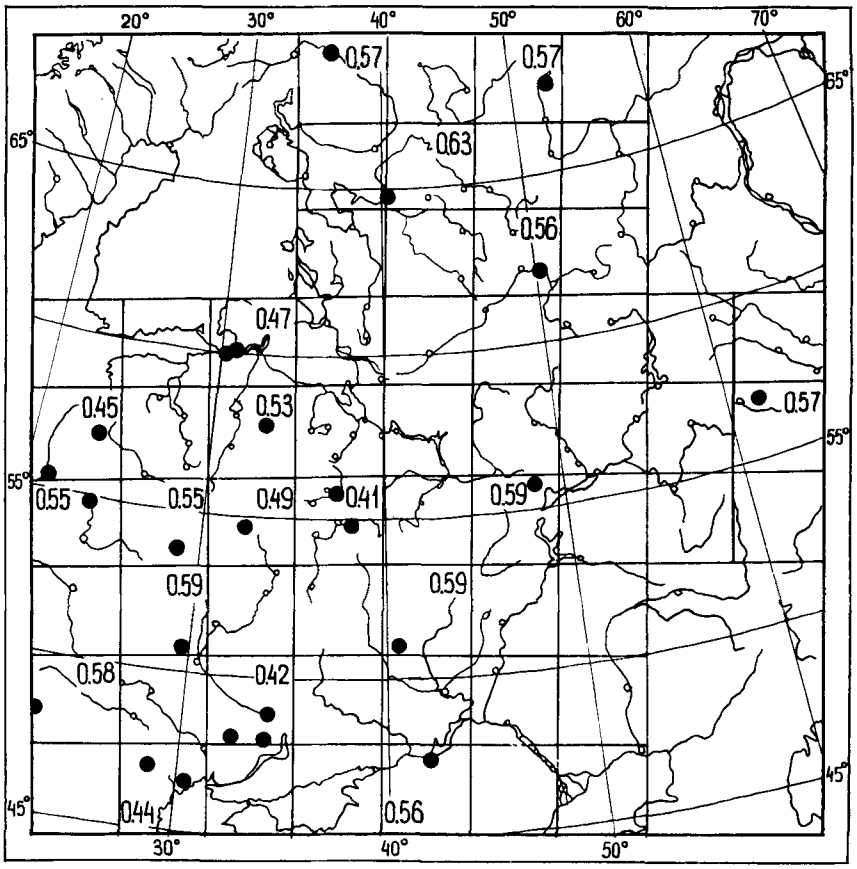

Fig. 3. The rms errors in determination of the mean-monthly deposition of the ion sum in precipitation for squares of 300 × 300 km² in the existing network of stations in the ETU.

position of the precipitation contaminants is measured, the ETU is divided into squares of 300 × 300 km² and for each of them the root mean-square error of averaging is determined (Fig. 3). The square side of 300 km is chosen because it is approximately equal to the correlation distance for deposition and at distances greater than 300 km the calculated correlation functions are not reliable. The results of calculations are given in Table 2.

As can be seen from Fig. 3 and Table 2, errors in averaging depend to a great extent, not only on the number of stations but also on their location. Thus for instance, if a station is situated near the centre of the square, the root mean-square error is 0.49 and if it is situated at the periphery of the square, the root mean-square error is 0.63, the mean value of this error being 0.57.

When a second station is added to the square, the error of averaging considerably decreases. But it is less affected by addition of one more station. This is probably caused by the non-rational location of these three stations. The network density can be considered completely adequate, if the root mean-square error of averaging over area does not exceed 0.40–0.45, the value of the random error of observation. Apparently, it is useless to make the network more dense and the error of averaging lower than the random error of observations. As can be seen from Fig. 3 and Table 2, to satisfy this

Table 2. *Dependence of the averaging error on the number of stations in the square of 300 × 300 km² in the ETU*

Number of stations in the square	Number of squares with a given number of stations		Averaging error (averaged over all the squares)	
	N	% of total number of squares	ε^2	ε
0	34	65	—	—
1	13	25	0.32	0.57
2	4	8	0.20	0.45
3	1	2	0.18	0.42
Total no. of squares	52	100	0.29[a]	0.54[a]

[a] Mean error for all the squares having stations.

criterion it is sufficient to measure the deposition of contaminants with precipitation at 2–3 stations in each square of 300 × 300 km².

Sirotenko (1969) shows that to obtain the same error while averaging the total 10-day amount of precipitation it is also necessary to have 2–3 stations in the same area. Thus, one can conclude that the variability of the contaminants in precipitation is of the same order of magnitude as that of the precipitation amount itself and is, most likely, determined by the latter. Therefore the field of contaminant concentration in the lower troposphere washed by precipitation is, probably, less variable than that of the precipitation amount.

From the requirement that the error in the point interpolation of contaminant concentrations in precipitation should not exceed the random error in concentration measurements, the maximum station separation is found to be 100 km. To obtain this mean distance between stations in the square of 300 km² it is necessary to locate about 15 stations at the apexes of hexagons 60 km in radius. In this case the ETU would have about 750 stations. Thus, a more "practicable" criterion for the network density to measure precipitation contaminants is less strict.

In conclusion, it may be said that the CACR's proposal to establish about 100 regional background air pollution stations situated at a distance of about 1 000 km from one another will result in too sparse a network for both actinometric measurements and those of the chemical composition of precipitation. However, it should be borne in mind that the criteria for the necessary density of station networks obtained in the present work and in the paper by Vinnikov & Dvorkina (1970) result from the requirement to obtain reasonably accurate information on the quantities measured from all the territory considered, without gaps. It is hardly necessary to have the full information on the chemical composition of precipitation from the whole area of oceans and all the sparsely populated regions. Therefore, probably, it is quite reasonable to meet these requirements first for some definite industrial and densely populated regions (Western Europe, eastern states of the USA) where in some countries such a network of stations already exists (Ericksson, 1970).

As mentioned above, the results presented here are preliminary and can be improved by using the data on the chemical composition of precipitation from a more dense network for more accurate calculations of the correlation functions in the range of small distances.

REFERENCES

1. CACR (1970). Commission for Air Chemistry and Radioactivity. Report to WMO on station network for world-wide pollutants.
2. Drozdova, V. M., Petrenchuk, O. P., Selezneva, E. S., Svistov, P. F. 1964. *The chemical composition of precipitation at the ETU.* Gidrometeoizdat (in Russian).
3. Ericksson, E. 1970. The importance of in-

vestigating global background pollution. Meteorological aspects of air pollution. *WMO Technical Note 106*. Geneva.
4. Gandin, L. S. 1963. *Objective analysis of meteorological fields*. Gidrometeoizdat (in Russian). (Translation by the Israel Program for Scientific Translations, Jerusalem, 1965.)
5. Gandin, L. S. 1968. On the planning of the climatological station network. Trudy GGO, vyp. 228 (in Russian).
6. Gandin, L. S., Boltenkov, V. P. 1964. On the method for investigation of three-dimensional structure of meteorological fields. Trudy GGO, vyp. 165 (in Russian).
7. Gandin, L. S., Kagan, R. L., Tarakanova, V. P. 1968. On the rational planning of the network for air temperature observations. Trudy GGO, vyp. 228 (in Russian).
8. Guschina, M. V., Kagan, R. L., Polischuk, A. I. 1967. On the accuracy of the determination of precipitation amount averaged over area. Trudy GGO, vyp. 208 (in Russian).
9. Kagan, R. L. 1967. Some problems in interpolation of precipitation amount data. Trudy GGO, vyp. 208 (in Russian).
10. Petrenchuk, O. P., Selesneva, E. S. 1970. Chemical composition of precipitation in regions of the Soviet Union. *J. Geophys. Res. 75*, 18.
11. SCEP Report 1970. *Man's impact on the global environment*. MIT Press, Cambridge, Mass.
12. Selesneva, E. S. (ed.) 1970. Monthly data on the chemical composition of precipitation for the period of 1962–1965 (in Russian).
13. Sirotenkio, O. D. 1969. Collection and use of data on the statistical structure of precipitation amount fields in agrometeorological calculations. Trudy IEM, vyp. 8 (in Russian).
14. Vinnikov, K. J., Dvorkina, M. D. 1970. Some problems on the planning of the actinometric network. *Meteorologia i Gidrologia*, No. *10* (in Russian).
15. WMO 1971. WMO operations manual for sampling and analysis techniques for chemical constituents in air and precipitation. No. 299, Geneva.

К ПЛАНИРОВАНИЮ СЕТИ СТАНЦИЙ ПО ИЗМЕРЕНИЯМ ХИМИЧЕСКОГО СОСТАВА АТМОСФЕРНЫХ ОСАДКОВ

Данные измерения химического состава осадков в 1958–1965 гг. на 28 станциях, расположенных на ЕТС, используются для расчета корреля ционных функций концентрации в осадках и выпадений ионов $Ca^{\cdot}$; $Na^{\cdot}$, SO''_4 и суммы ионов. Радиусы корреляции для различных ионов изменяются от 90 до 450 *км*. С помощью рассчитанных корреляционных функций оценен порядок величины среднеквадратических случайных ошибок измерений содержания ионов (35–45%). Рассчитаны ошибки оптимальной интерполяции в точку для значений концентрации и ошибки оптимального осреднения для значений выпадений. Показано, что для концентрации суммы ионов ошибка интерполяции, равная случайной ошибке наблюдения, достигается при среднем расстоянии между станциями порядка 100 *км*. При оптимальном осреднении такую же ошибку можно получить при наличии 2–3 станций на площади квадрата 300×300 *км*^{2}.