Application of a Scientific Climatological Approach and Statistical Method for EANET Network Optimization

Elena V. Gritsan[1)]* and Sergey A. Gromov[2)]

1) Institute of Global Climate and Ecology, Roshydromet and the Russian Academy of Science (RAS)

**Contact address: 20-b Glebovskaya str., Moscow, 107258, Russia [e-mail: gritsan@gmail.com]*

2) Acid Deposition and Oxidant Research Center (ADORC)

Precipitation chemistry measurements of the EANET network were analyzed using a statistical approach. Spatial representativeness, which is a condition of data spreading throughout the site surrounding area with the certain accuracy, was defined for each station. Areas that are not covered with the EANET network observations were evaluated on the basis of the obtained results.

1. Introduction

One of the main goals of an atmospheric monitoring network is to obtain data for appropriate and adequate evaluation of the deposition process under different geographical or climate conditions. From this point of view, information on measurements' representativeness is of great importance. Representativeness is a condition of spreading measurements at a single station throughout its surrounding area with the certain accuracy. In this paper, an attempt to identify such areas for each station of the EANET network was made. Approach applied in the study reveals "measurement gaps", i.e. territories to which measurements of any station can not be spread. These territories should be regarded as the first-priority areas under organization of new stations.

2. Description of the method

The most vital question in measurement representativeness estimation is choice of a representativeness criterion. One of the wide spread approaches is assumption an interpolation error or an error of spatial average value estimation at one station over a region as the representativeness criterion. An interpolation error is used for parameters with relatively low temporal and spatial variability. In case of precipitation chemistry we have a discrete field in time and space, so we should consider spatial averaged values and error of spatial average value estimation at one station over a region should be used as the representativeness criterion.

Assuming the exponential form of the correlation function, the error of estimation of the spatial average value at one station over region with area S can be calculated by the observations from the following expression (Kagan and Gushchina, 1966).

$$E = \sigma(1 - r(0) + 0.23\, r(0)\frac{\sqrt{S}}{l_0})^{\frac{1}{2}} \qquad (1),$$

where σ is the standard deviation, $r(0)$ and l_0 are parameters of the correlation function r(l), which reflects the dependence of the correlation coefficient from distance, and S is an area to which measurements at a station can be spread. $r(0)$ is a value of the correlation coefficient under zero distance. Theoretically, $r(0)$ should equal to 1 as it is correlation of observations at a station with themselves. However, it is smaller then 1 in practice due to the measurement error and influence of some other factors. Thus, a value of 1-$r(0)$ characterizes an input of the stochastic errors to element variability. l_0 is a correlation radius, i.e. it is a distance at which the correlation coefficient decreases in e times

Expression (1) is written in absolute values, but it is more convenient to deal with non-dimensional values. Both parts of (1) must be divided by the multiyear mean concentration (C) to obtain the non-dimensional form

$$\frac{E}{C} = \frac{\sigma(1 - r(0) + 0.23 r(0)\frac{\sqrt{S}}{l_0})^{\frac{1}{2}}}{C} \quad (2).$$

E/C varies from 0 to 1. It's a natural requirement E/C to be as small as possible. However, in practice it can not be so. In this investigation, E/C was accepted to be equal to the DQO (0.15).

Transforming (2), we obtain S. After it, we should decide what shape to choose for approximation of the area S. Generally, it can be any shape with the area S. In this paper, it was assumed to be a circle with the radius (S/π)1/2. So, the final expression for calculation of the representativeness is as following

$$R = \frac{l_0}{0.23 r(0)\pi^{\frac{1}{2}}}\left[\left(\frac{0.15C}{\sigma}\right)^2 - 1 + r(0)\right] \quad (3).$$

Mapping of the calculated radiuses shows the representativeness areas for all stations and makes possible to see "measurement gaps".

So, the method is based on calculating correlation function characteristics for the precipitation chemistry fields. Then, obtained results are used to find areas of representativeness for each station.

3. Data and procedure

The monthly averaged data on precipitation chemistry in 2000-2005 were used for the first approach and the weekly averaged data were used for more detailed investigation. The weekly averaged data were regarded separately for the climatic regions and only for the summer period (June-August) as data completeness during this time is very high. The set of stations, which data were used, are shown in Annex 1. nss-SO_4^{2-}, NO_3^-, NH_4^+, nss-Ca^{2+}, H^+ were chosen among all the measured components as they are priority pollutants of the rain water.

On the first step, correlation matrices were calculated for the every pair of stations and for the each parameter on the basis of the monthly and weakly averaged data. On the second step, dependences of the correlation coefficients from a distance were found and approximated by analytical expressions, i.e. the correlation functions were obtained. The correlation coefficients were averaged per distance's gradations beforehand. As a result, the shape of function was defined more evidently. It's necessary to note that only correlation coefficients for stations outlying from each other for no more than 1000 km were used for investigation. This is the scale of the regional atmospheric processes and along it correlation can be considered as nonrandom one.

4. Results

The correlation functions and its parameters $r(0)$ and l_0, which were obtained on the first step, are presented in Table 1.

The minimum value of l_0 is typical for NO_3^-. This fact probably is connected with the NO_3^- formation process. It is formed during the reactions of oxidation of NOx, which passes relatively fast, so NO_3^- spreads to smaller distances. Some statistical characteristics of the radiuses, which were calculated according to (3), are presented in Table 2.

Table 1 Correlation function and its parameters, calculated on the base of the monthly data

	Correlation function	$r(0)$	l_0, km
nss-SO_4^{2-}	$y=0.5398e^{-0.0009x}$	0.5398	1111
NO_3^-	$y=0.5263e^{-0.002x}$	0.5263	500
NH_4^+	$y=0.4101e^{-0.0011x}$	0.4101	909
nss-Ca^{2+}	$y=0.2759e^{-0.0015x}$	0.2759	667
H^+	$y=0.3853e^{-0.0011x}$	0.3853	909

Table 2 Statistical characteristics of radiuses for the monthly averaged data

	nss-SO_4^{2-}	NO_3^-	NH_4^+	nss-Ca^{2+}	H^+
Mean value, km	2112	1030	3025	4198	3303
Standard deviation, km	122	50	98	60	201
Maximum value, km	2281 (Jiwozi)	1097 (Kototabang)	3151 (Tappi)	4278 (Tanah Rata)	3539 (Weishuiyan)
Minimum value, km	1734 (Ijira)	833 (Ijira)	4051 (Listvyanka)	2794 (Primorskaya)	2878 (Tappi)

This result lets us see that statistics of precipitation chemistry from the existent stations may be applied for the first-order evaluation of acid deposition throughout the whole territory for all considered parameters from the certain point of application under the condition of 15%-error concentration determination. However, such high values of the representativeness radiuses are due to the suggestion about similarity of the correlation functions for the whole territory and for the whole year.

The second-order approach, which is based on the weekly averaged data, is more accurate definition of the correlation functions. Three climatic regions were separated on the investigated territory, which are, namely, the middle latitudes region (30 – 55° N), the subtropical region (22 – 30° N), and the tropic region (10° S – 22° N). The regions are presented in Fig.1. Only 2 stations are situated in the subtropical region, so none calculations can be done for it.

The calculated parameters of the correlation functions and the values of $r(0)$ and l_0 are shown in Table 3. In some cases, available data didn’t let to disclose the dependence of the correlation coefficient from distance.

The radiuses calculated on the base of the weekly averaged data for the summer period are smaller then radiuses calculated on the base of the monthly averaged ones. It means that correlation of the weekly averaged concentrations decreases faster with distance, so weekly measurement data of sites can be spread to smaller area than monthly ones.

The statistical characteristics of the calculated radiuses are presented in Table 4.

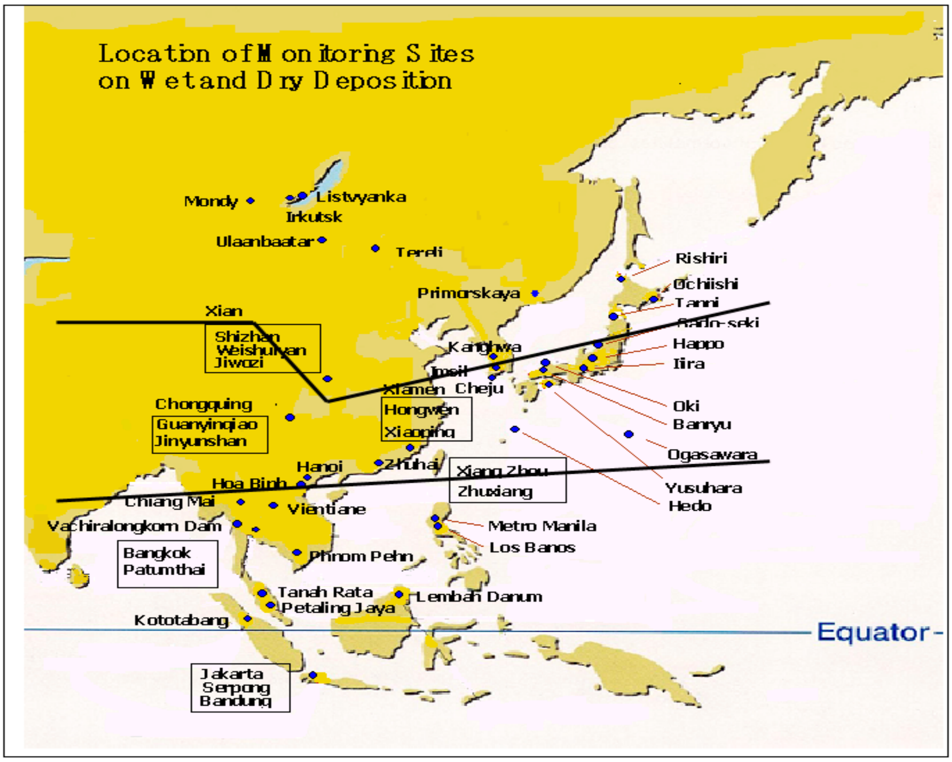

Figure 1 **Locations of the middle latitudes region (30 – 55° N), the subtropical region (22 – 30° N), and the tropic region (10° S – 22° N).**

Table 3 **Correlation function and its parameters, calculated on the base of weekly data**

	Correlation function	$r(0)$	l_0, km
Middle latitudes region			
nss-SO_4^{2-}	$y=0.619e^{-0.0016x}$	0.619	625
NO_3^-	$y=0.6137e^{-0.0021x}$	0.6137	476
nss-Ca^{2+}	$y=0.4829e^{-0.0025x}$	0.2759	400
Tropical region			
nss-SO_4^{2-}	$y=0.6536e^{-0.0011x}$	0.6536	909
NO_3^-	$y=0.6626e^{-0.0016x}$	0.6626	625
NH_4^+	$y=0.3872e^{-0.002x}$	0.3872	500

Table 4 Statistical characteristics of radiuses for the weekly averaged data

	nss-SO_4^{2-}	NO_3^-	NH_4^+	nss-Ca^{2+}	H^+
Middle latitudes region					
Mean value, km	828	646		3013	
Standard deviation, km	70	54		91	
Maximum value, km	923 (Jiwozi)	719 (Jiwozi)		3151 (Happo)	
Minimum value, km	654 (Ijira)	513 (Ijira)		2800 (Kanghwa)	
Tropical region					
Mean value, km	1061	853	1344		594
Standard deviation, km	55	60	43		44
Maximum value, km	1137 (Chiang Mai)	923 (Kototabang)	1403 (Chiang Mai)		655 (Chiang Mai)
Minimum value, km	972 (Kototabang)	774 (Hoa Binh)	1275 (Kototabang)		523 (Kototabang)

The smallest radiuses are typical for NO_3^- in the middle latitudes region and for H^+ in the tropical region. Radiuses for NO_3^- in the tropical region are also not so large, thus it confirms the fact that the densest network is required for adequate evaluation of the deposition process for this compound.

The maps of the representativeness areas for nss-SO_4^{2-}, NO_3^-, H^+, which were obtained on the base of the weekly averaged data, are presented in Figs. 2-4.

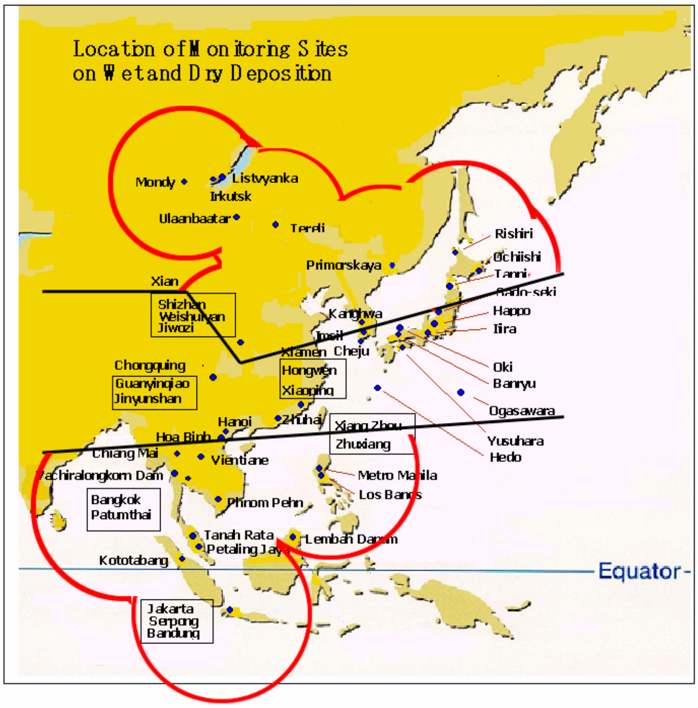

Figure 2 The representativeness areas for nss-SO_4^{2-}, obtained on the base of weekly data.

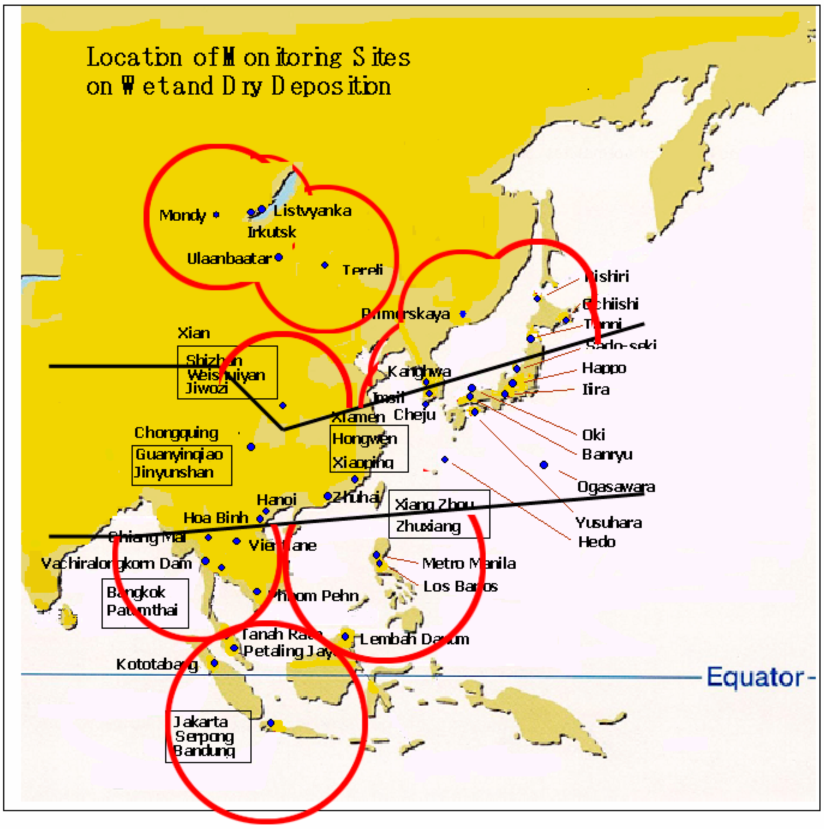

Figure 3 The representativeness areas for NO_3^-, obtained on the base of weekly data.

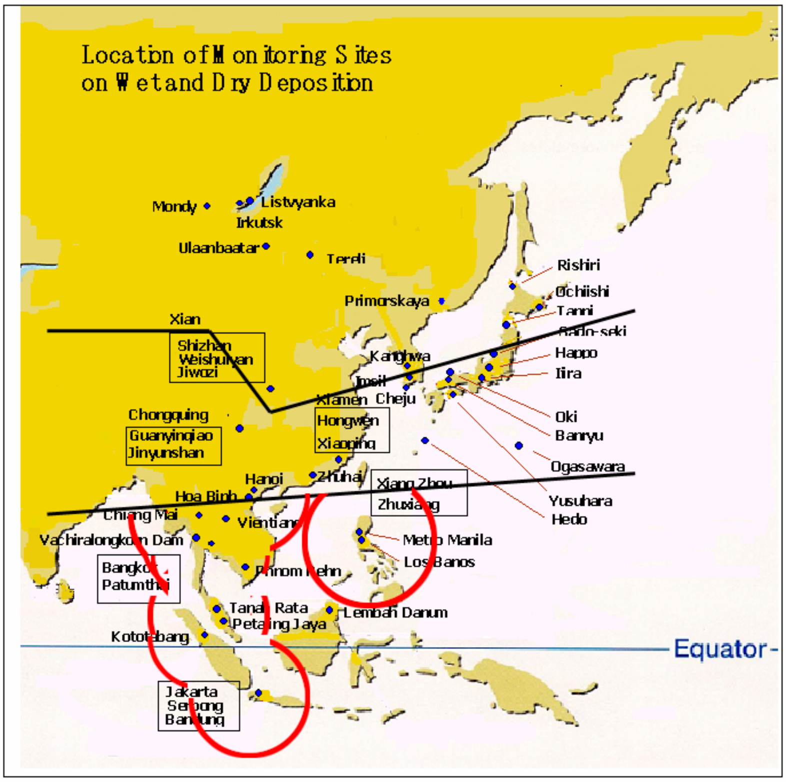

Figure 4 The representativeness areas for H^+, obtained on the base of weekly data.

Some important remarks should be done after the analysis of Figs. 2-4:

- there were no calculations for the subtropical region because only two stations belong to it (Xiaoping and Jinyunshan). Thus, it's the first area where new stations should be organized;
- for nss-SO_4^{2-}, the most part of the territory both in the middle latitudes and in the tropical region is covered by the observations. There are no suitable sites for the southwestern part of China, south part of Siberia, Russian Far East and east part of Indonesia;
- for NO_3^-, the gaps are more considerable. In addition to the same territories as for nss- SO_4^{2-}, the north and northeast parts of China are out of measurements network;
- for H^+, which was investigated only for tropical region, there are gaps in the east parts of Indonesia and Malaysia.

Summarizing this all we can say that densest network should be created for measuring NO_3^-, what is in good correspondence with the scientific conception of its cycle in the atmosphere. The gaps in the precipitation measurements are detected in the subtropical region (22 – 30° N) for all parameters; in the southwestern part of China, south part of Siberia, Russian Far East, and east part of Indonesia for nss-SO_4^{2-}; the southwestern, north and northeast parts of China, south part of Siberia, Russian Far East, and east part of Indonesia for NO_3^-; the east parts of Indonesia and Malaysia for H^+.

These results are preliminary ones. It's evident that additional data in term of spatial resolution and temporal continuance is necessary to define the correlation functions and its parameters more precise and appropriate for evaluation of the representativeness areas. Experiments with the data of modeling will be very useful for this purpose. Also additional data let to identify variability of the correlation function in the different directions, for example, to evaluate whether the representativeness areas will be an ellipse or a round circle.

References

Isaev A. A. 1988. Statistics in meteorology and climatology. Moscow University, Moscow, Russia, pp. 209-218.

Kagan R. L. 1966. Towards representativeness estimation of the precipitation data. Trudy MGO, 191: 22-34.

Kagan R. L. & Gushchina M. V. 1966. On the statistical structure of the precipitation fields. Trudy MGO, 191: 35-47.

Vitvitskii G. N., 1960. The Foreign Asia climates. Gosud. Isd. Geogr. Lit., Moscow, Russia, 399 p.

Annex 1 Sets of stations, which data were used for calculations on the first-order and second-order approaches

China	Jinyunshan Weishuiyuan Jiwozi Xiaoping	Indonesia	Serpong Kototabang
		Malaysia	Tanah Rata
		Philippines	Los Banos
Japan	Rishiri Tappi Sadoseki Happo Oki Yusuhara Ogasawara Hedo Ijira Ochiishi	Thailand	Patumthani Chiang Mai (Mae Hia) Kanchanaburi (Vachiralongkorn Dam)
Mongolia	Terelj	Viet Nam	Hoa Binh
Republic of Korea	Kanghwa Cheju Imsil	Russia	Mondy Listvyanka Primorskaya