AN ANALYSIS OF AIR POLLUTION DATA
IN KAWASAKI CITY, JAPAN

by

YASUTOSHI WASHIO¹
Keio University, Japan
and Stanford University

WORKING PAPER #2

This draft is intended for limited circulation

September 17, 1976

NOTICE
This report was prepared as an account of work sponsored by the United States Government. Neither the United States nor the United States Energy Research and Development Administration, nor any of their employees, nor any of their contractors, subcontractors, or their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness or usefulness of any information, apparatus, product or process disclosed, or represents that its use would not infringe privately owned rights.

MASTER

There is no objection from the patent point of view to the publication or dissemination of this document. (b)

Patent Title No.

BROOKLYN PATENT GROUP

12/8 1976 By epc

¹This research was supported by Grant from SIMS

Prepared For
THE U.S. ENERGY RESEARCH AND DEVELOPMENT ADMINISTRATION
UNDER CONTRACT NO. EY-76-S-02-2874.*000

DISTRIBUTION OF THIS DOCUMENT IS UNLIMITED *leg*

DISCLAIMER

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency Thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

DISCLAIMER

Portions of this document may be illegible in electronic image products. Images are produced from the best available original document.

An Analysis of Air Pollution Data
in Kawasaki City, Japan

by

Yasutoshi Washio
Keio University, Japan
and Stanford University

1. Introduction and Data.

Kawasaki City is located between Tokyo and Yokohama and is famous as one of the most polluted regions in Japan because it has a big industrial area:

Population: 1,000,000 Area: 140 km² .

The City has been systematically gathering data on air pollution; such as, SO_x, NO_x, suspended particulates, oxidants, CO, ...etc., and on weather conditions such as, temperature, humidity, wind direction and wind velocity. However, no systematic and active analysis has been performed so far. We give attention to the data for concentrations of suspended particulates and of their constituents.

Measurements

Data:

Suspended particulates (μg/m³)

Constituents (μg/m³)

SO₄²⁻

NO₃⁻

Fe

Pb

Cu

Mn

V

Cd

Ni

Cr

Weather conditions (wind direction; wind velocity)

It is said that samplers emit Cu and that the data of Cu are not reliable, therefore, the data of Cu will be removed.

Measurements are made twice a month; one is at the beginning of a month and the other is at the latter part of a month. (The dates are selected at random under the restriction mentioned above.) Sampling is made from 10:00 a.m. till 10:00 a.m. the next day (24 hours). The sampler is a high-volume air sampler. The City began to gather the data in 1972, but we use the data in 1973 and 1974 (calendar year).

Observation Stations.

There are seven observation stations. In connection with the data in Kawasaki, I have to mention NASN (National Air Surveillance Networks) in Japan. NASN has 15 observation stations now, #1 in Kawasaki (below)

is one of these stations.

Each station is gathering the same kinds of data as those in Kawasaki. To be exact, Kawasaki City followed the NASN. No systematic analysis for the data of NASN has been performed.

2. Data Analysis

(1) Evaluation of the Degrees of Pollution.

For each station, the mean concentrations of the two years 1973 and 1974 were calculated for each characteristic. By using these results we evaluate the degree of pollution for each station.

Table 2.1

Charac- teris- tic Station	S.P.	Constituents in Suspended Particulates								
		SO ₄ ²⁻	NO ₃ ⁻	Fe	Pb	Mn	V	Ca	Ni	Cr
#1	★			▲		●	▲		▲	▲
#2	★	▲		▲		▲	▲	●	▲	●
#3	⊙									
#4	⊙									
#5	⊙			△		△			△	△
#6	⊙			△		△	△	△	△	△
#7	☆	△		△	△	△	△	△	△	△

Mark	Range of Concentrations
★	more than 170 µg/m ³
⊙	150 - 170
☆	less than 150

Mark	Range of Concentrations
●	More than 1.6 times of mean con.
▲	Between 1.2-1.6 times of mean con.
blank	Between 0.8-1.2 times of mean con.
△	Between 0.4-0.8 times of mean con.
○	Less than 0.4 times of mean con.

(i) We can see that the degree of pollution at #7 is the lowest among the seven stations; in fact, #7 is located in the residential district.

(ii) In general, except NO_3^- and Pb, concentrations of pollutants are decreased as the station departs far from the seaside.

It is said that Fe and Mn come from iron industry and that V and Ni come from oil refinery industry. The results of Table 2.1 coincide with these facts.

(iii) The concentrations of NO_3^- and Pb are nearly the same over the stations. That is, the concentrations of NO_3^- and Pb are still high even in the business district and residential district.

This result coincides with the fact that NO_3^- and Pb come from automobiles.

(iv) The concentrations of Cd and Cr at #2 are abnormally high. We shall discuss this point later.

(v) In order to know the degree of pollution in Kawasaki relatively, the data in #7 were compared with another city - Matsue City, which can be considered as one of the less polluted cities in Japan.

Table 2.2 Comparison between #7 and Matsue City

Mean Concentrations ($\mu\text{g}/\text{m}^3$)

Station	S.P.	Constituents in Suspended Particulates							
		SO_4^{2-}	NO_3^-	Fe	Pb	Mn	V	Cd	Ni
#7	127.4	12.05	6.58	3.55	.305	.115	.0363	.0069	.0210
Matsue City	73.7	9.03	3.02	1.04	.054	.035	.0158	.0020	.0141

This table shows how the air in Kawasaki City is polluted, because #7 is the least polluted station in Kawasaki.

(2) Time Series of Concentrations of Pollutants.

We examined by t-test, except SO_4^{2-} and NO_3^- , whether the differences in concentration between 1973 and 1974 are significant or not. The results are given in Table 2.3.

Table 2.3

Station	Pollutant which was decreased in '74 significantly	Pollutant which was increased in '74 significantly
#1	Fe^{**} , Mn^* , V^{**}	
#2	Fe^* , V^{**}	
#3	Fe^{**} , Mn^{**}	
#4	Fe^* , Ni^*	
#5	V^*	
#6	V^*	
#7		

(*) shows 5% significance
 (**) shows 1% significance

(i) There is no pollutant which was increased in 1974 significantly. Thus, we can say that, in general, the concentrations of pollutants are of stable state or are decreasing.

(ii) At stations #1, #2 and #3 (where they are close to the industrial area), the concentrations of Fe, Mn and V have been decreased significantly in 1974. The next table will tell us one cause of this fact.

Table 2.4 Total quantity of fuels and raw materials consumed in the main plants in Kawasaki City.

Fuels R.M. Year	Crude Petroleum	Coal	Gas	Iron Ore
1969	312	176	-	456
1970	394	109	-	449
1971	385	90	-	414
1972	364	39	-	310
1973	430	1.8	423	<u>305</u>
1974	404	-	426	<u>164</u>

In fact, the quantity consumed of iron ore in 1974 has been reduced to half that in 1973. A main cause is business depression in 1974 in Japan.

(iii) As mentioned earlier, station #1 is also one of the stations of NASN in Japan, thus it has data on air pollution since 1969. The data for the concentration of "Suspended Particulates" are given by Table 2.5.

Table 2.5

Average concentration of S.P.
at Station #1 ($\mu\text{g}/\text{m}^3$)

Year	Concentration
1969	408.1
1970	334.6
1971	359.9
1972	199.5
1973	184.8
1974	160.0

This table shows that the concentrations of S.P. are decreasing from year to year. Main causes of this will be:

1. As we can see in Table 2.4, as one of the fuels, the coal has been replaced by crude petroleum or gas in 1972-1973.
2. Kawasaki City ordered the main plants to set dust collectors in 1972.

(3) Station Feature Viewed from the Type of Pollution.

Each station has nine measurements (variables). In order to clarify the feature of each station and to classify the stations, principal component analysis has been applied.

As each measurement has the same unit ($\mu\text{g}/\text{m}^3$), we may employ either the sample variance-covariance matrix or sample correlation

matrix in the theoretical point of view as an estimate of the variance-covariance matrix of nine measurements (variables). However, there is a big difference among the absolute values of their measurements. (variables). Therefore, as it seems preferable to do so, we employ the sample correlation matrix as the estimate of the variance-covariance matrix of nine variables. (The data from March 1973 through February 1974 are used for this analysis.) Table 2.6 shows the result.

Table 2.6 Coefficient vector for principal component and correlation coefficient between principal component and measurement (variable).

P.C. M.	z_1		z_2		z_3		z_4	
SO_4^{2-}	.353	.764	.161	.185	-.468	-.479	.287	.232
NO_3^-	.274	.593	.520	.597	-.464	-.474	.058	.047
Fe	.368	.797	-.422	-.484	-.024	-.025	-.134	-.108
Pb	.254	.546	.527	.604	.204	.208	-.480	-.388
Mn	.364	.789	-.240	-.275	.030	.031	-.350	-.283
V	.387	.838	-.135	-.155	.142	.146	-.222	-.180
Cd	.172	.373	.364	.418	.695	.711	.310	.251
Ni	.428	.925	-.104	-.120	.050	.051	.014	.011
Cr	.328	.712	-.168	-.193	.130	.133	.631	.510
Characteristic Root	4.695		1.312		1.046		.653	
Percentage of Total Variance	52.2		14.6		11.6		7.3	
Cumulative Percentage of Variance	52.2		66.8		78.4		85.7	

(The first column shows coefficient vector and the second column shows correlation coefficient.)

Table 2.6 shows the coefficient vector for principal component and correlation coefficient between principal component and variable.

(i) All variables have nearly high positive correlations with the first principal component. Then, the first principal component appears to be a measure of the general degree of pollution.

(ii) The second principal component z_2 is highly and positively correlated with NO_3^- , Pb, and Cd and is highly and negatively correlated with Fe and Mn. Thus, z_2 may be thought of as an index which shows the relative relation between a group of NO_3^- , Pb, and Cd and a group of Fe, and Mn. It is said that Fe and Mn come mainly from heavy industry plants and that NO_3^- and Pb come mainly from cars. Therefore, the second component z_2 may be thought of as a factor which shows the type of pollution - city type or heavy industry type - depending on the values of z_2 .

(iii) The third and fourth components explain small portions of the total variance. It was difficult for us to give the meaning of them.

(iv) More than half the total variance in the nine dimensions is accounted for by component z_1 . Furthermore, about 70% of total variance is accounted for by the first two components. Hence, the first two components will be enough to explain the data.

(v) The seven stations are plotted on the $z_1 - z_2$ plane as in Figure 2.1

- a) The general degree of pollution increases as a station approaches the industrial district.
- b) Pollution in stations #1 and #2 is of the heavy industry type. Pollution in stations #3, #4, #5, and #6 is of the city type.

Figure 2.1

(vi) I think the number of stations is too small to apply principal component analysis.

(4) Relationship Between the Station Feature Viewed from Pollution Type and Weather Conditions.

In general, air pollution data will be affected by weather conditions. Above all, weather and wind direction will have a great influence on air pollution data. In order to examine this relationship, for each station, data are divided by 2x2 table as in Table 2.7 according to weather and wind direction. Three data were selected from each cell. Selection was made so that the bias due to the seasons is as small as possible. (But it was difficult to remove the bias completely.) The dates in the cells show those of the data which were selected.

Table 2.7 Grouping of the Data

Wind Direction Weather	North Wind (No)	South Wind (So)
Fine or Cloudy (Fi)	March 22 ('73) Nov. 6 ('73) Jan. 8 ('74)	July 24 ('73) Aug. 21 ('73) Nov. 18 ('73)
Rainy (Ra)	June 12 ('73) Oct. 16 ('73) Feb. 5 ('74)	May 8 ('73) May 22 ('73) Sept. 4 ('73)

The wind directions in Kawasaki City can be classified into two groups - the north and the south winds - as is shown in Figure 2.2.

Figure 2.2

For all the data selected, the values of principal components z_1 and z_2 are calculated, and the average values of z_1 and z_2 in each cell are calculated.

The four average values of components (z_1, z_2) are plotted as in Figure 2.3.

Figure 2.3 Change of pollution type due to weather conditions. (Stations #2, #3, #4, #5 and #6 are omitted.)

- Fi and No
- × Fi and So
- ▲ Ra and No
- Ra and So

(i) In general, a fine or cloudy day is more polluted than a rainy day. This fact coincides with our common sense.

(ii) At stations #1, #2, and #3, which are close to the seaside district (= industrial district), the pollution type is changed by wind direction. It is of the city type under the north wind, and of the heavy industry type under the south wind. This fact will be explained by their location.

(iii) At stations #6, and #7, which are located in the inland area, pollution type is not affected by wind direction. This fact is also explained by their location, because there is no wind from the east (= industrial area).

(5) The Relationship Between Concentrations of Pollutants and Weather Conditions (1).

Due to different environments, concentrations of pollutants at seven stations on the same day are different. On the other hand, concentrations of pollutants at one station vary from day to day or are changing constantly. One of the main causes of this will be the weather conditions; that is, air pollution data will be affected by weather conditions even though the source of pollution is kept constant. Thus, we can't get any information from air pollution data without knowing the weather conditions. This raises the problem of seeking the relationship between air pollution data and weather conditions.

In order to examine this relationship globally, the technique of "Control Charts" has been employed. The weather conditions at seven stations on the same day are nearly the same, and the weather conditions will vary from day to day. Then we choose the "day" as a "rational subgroup" and we draw $\bar{x}$ - R charts; that is, $n = 7$ and $k = 24$ (only the data of 1973 were used). For points (days) falling outside the control limits, we examined the weather conditions (weather, wind direction, wind velocity) and concentrations at seven stations. By doing this, we have the following: (Here we note that the following results were obtained under the assumption that the source of pollution was constant in 1973.)

S.P.:

- (i) The concentration is high on fine or cloudy days, and is low on rainy days. This fact is clear by considering that if it rains, the pollutants adhere to drops of water and fall to the ground.
- (ii) We can't find any noteworthy property about wind direction and wind velocity.

SO_4^{2-} , Fe, Mn :

- (i) The concentrations are high, especially at stations #1 and #2 when the city has the south wind. In fact, the industrial area is located south. Then we can say that they come from the plants in the industrial area.

Pb :

The concentration is high (low) when the wind velocity is low (high). There is no noteworthy property about the wind direction, and this enables us to make an inference that Pb comes from moving sources (automobiles).

Cd :

The concentration is high, especially at station #2 when the city has the north wind. Furthermore, the concentrations at #1 and #2 are extremely high comparing to other stations. This enables us to make an inference that there will be a source of Cd at just the north part of #2 and #1. In fact, we found an industrial area consisting of many small factories in the north part of #2. Thus, we can guess that Cd comes from Tokyo.

(6) Relationship Between Concentrations of Pollutants and Weather Conditions (II).

In order to analyze the relationship between concentrations of pollutants and weather conditions, the technique of regression analysis has been performed.

Dependent variables (criterion variables)

Concentrations of S.P., Fe, ..., and Ni at station #7.

Independent variables (predictor variables)

- (i) hours of sunshine
- (ii) amount of cloud
- (iii) precipitation on that day

- (iv) precipitation on the previous day
- (v) wind velocity
- (vi) temperature
- (vii) humidity
- (viii) wind direction

(The wind direction in Kawasaki City can be classified into two groups - the north and the south winds.)

Regression analysis was applied only to the data at #7. The reason (why #7 was selected) is that #7 is located far from the industrial area, and we can assume the data will not be affected by the rate of operation of the plants; that is, the source of pollution at #7 is nearly constant for a given period.

The data in 1973 were used to obtain regression equations, and the data in 1974 were used for checking the reliability of the regression equations obtained.

The results are shown in Table 2.8. The multiple correlation coefficients of concentrations of pollutants with the independent variables are between 60% and 80%.

Table 2.8

	Multiple correlation coefficient (Data in 1973 used)	Percentage of the days which the differences between predicted values and actual values are within 10% of average concentration of the year. (Data in 1974 used)
S.P.	.596	13.0
Fe	.790	21.7
Pb	.704	13.0
Mn	.661	17.4
V	.722	21.7
Cd	.735	17.4
Ni	.634	21.7

The results are not satisfactory for us. One of the reasons is a discrepancy in time between the air pollution data and some of the weather condition data.

After returning to Japan, I will get precise data on weather conditions. Then I will again try to analyze the data through regression analysis.

Air pollution data depend on the weather conditions. Even though the source of air pollution is constant, the data on air pollution vary from day to day, depending on weather conditions. Thus, when we discuss the standard on air pollution, it is necessary to know the relationship between air pollution data and weather conditions. This is the main purpose of this analysis.