

国設大気測定網 (NASN) 浮遊粒子状物質観測データに  
おける欠損値の補完

関根嘉香、比嘉宏之、益田文教

〒259-1291 神奈川県平塚市北金目 1117  
東海大学 理学部 化学科  
E-mail: sekine@keyaki.cc.u-tokai.ac.jp

Complement of Missing Values in Atmospheric Concentration Data on  
Suspended Particulate Matter Operated by National Air Surveillance  
Network, Japan

By

Yoshika Sekine, Hiroyuki Higa and Fuminori Masuda

Department of Chemistry, School of Science, Tokai University  
Hiratsuka, Kanagawa, 259-1292, Japan  
E - mail: sekine@keyaki.cc.u-tokai.ac.jp

**Abstract**

Incomplete data set, which has missing values in the matrix, is a problem that is repeatedly encountered in environmental researches. To cope with the problem, this study aims to provide useful methods to complement missing values in air quality data sets. As a case study, authors have tested some methods to fill in the missing values in compiled SPM and multi-elemental concentration data by National Air Surveillance

Network, Japan, using simulated missing data patterns. Missing rows in the matrices were filled with mean of surrounding values (MSV), mean of values before and after 12 months of the missing data (MV12) and linear interpolation (LIN). Performance of those methods were evaluated by statistical indices such as Index of Agreement developed by Willmott, root mean square squared error (RMSE) and mean absolute error (MAE). Whereas MSV and LIN were strong tools to fill in the short length of gap, MV12 could be an alternative way to interpolate the longer length of gap. Linear regression analysis was applied to complement the missing values in column. This method was found useful when a correlation coefficient was greater than 0.7 between concentrations of element with missing values and predictor element. Based on the test results, full complement was demonstrated on the incomplete data matrix of Kawasaki site from April 1974 to March 1997.

## 1 . はじめに

環境情報の解析において、測定データの欠損値、すなわち不完全なデータマトリックスの存在とその扱いは重要な問題である。欠損値は、サンプリングや分析操作における試料の損失、測定器具の不良などによってもたらされる。特に時系列的な予測に測定データを利用する場合、欠損値による不連続性は深刻な障害になる。

欠損値に対して系列平均値で置換する方法が一般的に推奨され、市販の多くの統計ソフトウェアにもデフォルト機能として使用されている。しかし本法は、測定データ固有の構造を著しく乱し、共分散または相関分析における誤差要因となり、統計モデルの性能を大きく低下させてしまうため[1]、欠損値の補完に関する新しい方法あるいはアルゴリズムの開発が行われている[2,3]。

Junninen et al.[1]は、フィンランド国ヘルシンキおよびベルファストの大気質測定データ (NO<sub>x</sub>, NO<sub>2</sub>, O<sub>3</sub>, PM<sub>10</sub>, SO<sub>2</sub>, CO, 風速, 風向, 温度および湿度の1時間値の時系列データ) における欠損値の補完方法として、種々の単変量補間法 (最近接値置換、直線補間、曲線補間)、多変量補間法 (回帰分析法、最近接値置換法、Self-organization map、Multi-layer perception) あるいは直線相関と各多変量補間法の組み合わせ (ハイブリッドモデル) を検討し、各手法の有用性を比較している。しかしながら適用すべき補完方法は、対象とするデータマトリックスの特性、欠損値の現れ方によって検討されねばならない。

環境省 (旧環境庁) では国設大気測定網 (National Air Surveillance Network, NASN) の一環として、1974年4月から1997年3月まで全国の測定点において浮遊粒子状物質 (Suspended Particulate Matter, SPM) およびその元素濃度の観測

を実施した[4]。この観測データには、過去 23 年間の SPM 濃度および 31 元素濃度の地域分布、経時変化が記録されており、東アジア地域における微量元素の長距離輸送の解明[5,6]、黄砂現象の検討[7]などに利用されている。しかしながら観測データにはいくつかの欠損値があり、不完全なデータマトリックスとなっている。そこで本研究では、NASN の浮遊粒子状物質観測データの欠損値を補完することを目的に、観測データの特性に適した補完方法の検討を行った。

## 2. 方法

### 2.1 用いたデータ

NASN の浮遊粒子状物質の観測は 1974 年度より開始された。その後、順次観測点が追加され、札幌、新潟、仙台、東京、川崎、名古屋、大阪、尼崎、松江、倉敷、宇部、大牟田、野幌、籠岳、京都八幡および筑後小郡の 16 地点において、SPM 濃度および 31 元素濃度の月平均値の観測が実施された。SPM 試料は 10 $\mu$ m カット式ローボリューム・エアー・サンプラー（新宅製 SK、捕集口径：100mm $\phi$ 、吸引流量：20 $\ell$ /分）を用い、ニトロセルロース製メンブランフィルター（ポアサイズ 3 $\mu$ m）上に毎月 1 日 10 時から 26 日 10 時までの 25 日間連続して吸引捕集された。バックアップフィルターにはポリフロン製フィルター（ポアサイズ 10 $\mu$ m）が用いられた。大気中 SPM 濃度は重量法、元素濃度は中性子放射化分析法および蛍光 X 線分析法（Cd, Ni および Pb）により定量された[8]。

この NASN の観測データにはいくつかの欠損値が存在する。欠損値の発生パターンは 1) 観測月の全測定項目が欠損している場合（データマトリックスの行成分が欠損していることから、以下、行欠損と略記する）、2) 特定の元素濃度のみが欠損している場合（データマトリックスの列成分中の一部が欠損していることから、以下、列欠損と略記する）の二通りがある。図 1 に行欠損、列欠損の一例を示す。

行欠損は主としてサンプリングのエラー（フィルター破損、試料の汚染など）に基づくと考えられる。各測定地点における行欠損の発生率を表 1 に示す。発生率は 0（籠岳）～5.8%（大牟田）、平均で 2.3%であった。また欠損長は 1 ヶ月が最も多く全地点合計で 71 回、2 ヶ月が 7 回、8 ヶ月が 1 回であった。

列欠損は、主として分析上の問題に起因すると考えられる。各測定点における測定項目毎の列欠損の発生率を表 2 に示す。尚、列欠損発生率には行欠損を含めていない。

### 2.2 行欠損の補完方法

行欠損の補完方法には、周囲値平均(Mean of surrounding values, 以下 MSV と略記する)、前後年平均(Mean of values before and after 12 months of the missing value, 以下 MV12 と略記する)および線形補間(Linear interpolation, 以下 LIN と略記する)

を用いた。

MSVは欠損値の補完方法として広く利用されており、ここでは欠損している値 $n_i$ の前後2ヶ月、計4ヶ月の値の平均から求めた。

$$MSV = (n_{i-2} + n_{i-1} + n_{i+1} + n_{i+2}) / 4 \quad \dots$$

MV12は欠損月 $i$ の12ヶ月前( $i-12$ )の値 $n_{i-12}$ と12ヶ月後( $i+12$ )の値 $n_{i+12}$ の平均値と定義した。式に定義式を示す。これは、大気中元素濃度に顕著な季節変化があり、欠損月の前後月よりも前年または後年同月の値の方が観測値の類似性が高い場合に有効である。尚、補完に用いる値も欠損している場合、さらに12ヶ月先の値を使用した。

$$MV12 = (n_{i-12} + n_{i+12}) / 2 \quad \dots$$

LINは、行欠損値 $n_i$ の前後の値 $n_{i'}$ および $n_{i''}$ を直線で結び、欠損部分を補完する方法であり、式によって求められる。本法は、大気中元素濃度が一定の増加または減少傾向を示す場合に有効と考えられる。

$$LIN = n_{i'} + k(i - i') \quad \dots$$

ここで、 $k = (n_{i''} - n_{i'}) / (i'' - i')$ 、 $i' < i < i''$  である。

### 2.3 列欠損の補完方法

発生源が同一で大気中での挙動が類似する元素間では、大気中濃度の経時変化が類似し、相関分析において有意に高い相関係数を示すことがある。この性質を利用して、列欠損を有する元素濃度の経時変化に対して、もっとも相関の高い経時変化を示す他の元素濃度を用いて欠損値の補完を行った。列欠損を有する元素濃度を従属変数 $y$ 、列欠損を持たない他の元素濃度を独立変数 $x$ として回帰分析を行い、相関係数 $r$ が有意に高い元素を選択し、その元素濃度 $x$ から一次回帰式により欠損部分を補完した。

$$y = ax + b \quad \dots$$

### 2.4 補完方法の有効性の判定

上記の補完方法の有効性を検証するために、行欠損・列欠損いずれも存在し

ない籠岳測定点のデータマトリックス(行:192ヶ月、列:SPM+31元素濃度)において、意図的に欠損値を発生させて、観測値と補完値を比較した。なお、検出下限以下のデータは検出下限値を用いた。

行欠損:SPM濃度を対象に試行した。欠損率は大牟田の最大欠損率5.8%を踏まえて5.8%に設定した。欠損長は1、2、5および8ヶ月の場合をそれぞれ設定し、乱数表を用いてランダムに欠損値を発生させた。各欠損長においてそれぞれ5パターンの欠損マトリックスを作成し、観測値と補完値の比較を行った。

列欠損:はじめに籠岳のデータマトリックス(192ヶ月分)を用いて元素濃度間の相関行列を求めた。その結果、Fe濃度が他の元素濃度に対して比較的相関が高かったため、従属変数yはFe濃度を用いることにした。次に意図的に10%(19ヶ月)、20%(38ヶ月)および30%(58ヶ月)をデータの始めから列欠損させ、Fe濃度に対して各元素濃度で回帰分析を行い、欠損値の補完を試みた。

補完した値の有効性の判定は、Willmottの一致の指標(Index of Agreement,以下dと略記する)、平方根平均二乗誤差(Root Mean Squared Error,以下RMSEと略記する)および平均絶対値誤差(Mean Absolute Error,以下MAEと略記する)の3種類の統計的基準を用いた。

Willmottのdは観測値とモデル推定値の比較の指標として開発されたものであり、式によって定義される[9,10]。これはdが1に近いほど予測値と実測値との差が小さいことを示している。なお、ここではk=1とした。

$$d = 1 - \left[ \frac{\sum_{i=1}^N (P_i - O_i)^k}{\sum_{i=1}^N (|P_i - \bar{O}| + |O_i - \bar{O}|)^k} \right] \quad \dots$$

RMSEは、観測値と補完値の残差を累積するものであり、補完方法の平均誤差を表す。

$$RMSE = \left( \frac{1}{N} \sum_{i=1}^N [P_i - O_i]^2 \right)^{\frac{1}{2}} \quad \dots$$

MAEは式を用いて算出した。これは観測値と補完値の差の絶対値が大きくなる程、値が大きくなる。

$$MAE = \frac{1}{N} \sum_{i=1}^N |P_i - O_i| \quad \dots$$

### 3. 結果および考察

#### 3.1 行欠損の補完

籠岳測定点の SPM 濃度データに対して意図的に行欠損を与え、各手法により補完した場合の有効性を検討した。図 2 は Willmott の  $d$ 、図 3 は RMSE、および図 4 は MAE によって評価した結果である。 $d$  は 1 に近いほど、MAE、RMSE は 0 に近いほど誤差が少なく、有効性も高いことを表す。図 2、3、4 を見ると、欠損長が短いほど MSV、LIN が MV12 より有効であるが、欠損長が 8 ヶ月になると両方法ともに著しく補完性能が低下し、相対的に MV12 が優れる結果となった。これは、大気中 SPM 濃度の変動は周辺月の経時傾向により支配されていることを表しており、数ヶ月程度の欠損の場合は前後月の濃度レベルが欠損値の濃度レベルに近接することを示している。したがって MV12 法は必ずしも適当な補完方法ではないが、長期に行欠損が発生する場合には MSV や LIN の代用法として利用できると考えられる。以上のことから、短い欠損長には LIN または MSV を用い、長い欠損長には MV12 を使用することが適当と考えられた。

#### 3.2 列欠損の補完

籠岳の Fe 濃度データに対して意図的に列欠損を与え、各手法により補完した場合の有効性を検討した。図 5 は Willmott の  $d$ 、図 6 は RMSE、および図 7 は MAE によって評価した結果である。補完方法の性能は、補完に用いる元素との相関係数の高さによって変化するため、相関係数  $r$  を横軸、縦軸に有効性の指標をとった。尚、図 5、図 6 および図 7 において、10%欠損時は Fe 濃度を Co 濃度( $r=0.601$ )、Mn 濃度( $r=0.698$ )、Ce 濃度( $r=0.815$ )および Sm 濃度( $r=0.90$ )で補完した。また 20%欠損時は、Fe 濃度を Co 濃度( $r=0.613$ )、Mn 濃度( $r=0.696$ )、Ce 濃度( $r=0.808$ )および Sm 濃度( $r=0.904$ )により補完した。さらに 30%欠損時は、Fe 濃度を Na 濃度( $r=0.588$ )、Mn 濃度( $r=0.704$ )、Ti 濃度( $r=0.781$ )および Sm 濃度( $r=0.916$ )により補完した。

図 5 の Willmott の  $d$  では欠損率 10%、20%、30%のいずれにおいても相関係数  $r=0.6\sim 0.7$  を境目に大きく増加していることがわかる。また  $r=0.7$  以上では  $d=0.9$  以上となり良好な結果となった。この傾向は図 6 の RMSE でも見られた。したがって、列欠損を回帰分析によって補完する場合、予測に用いる元素濃度との相関は、相関係数として 0.7 以上が望ましい。

#### 3.3 川崎測定点データの補完

以上の検討をもとに川崎測定点のデータ（行欠損：欠損長 1 ヶ月が 5 箇所、列欠損：最大で 13%）に対して、欠損値の補完を試みた。行欠損は LIN により補完し、列欠損は回帰分析により補完した。結果の 1 例として、大気中 As, Sc

よび Ti 濃度に対する結果を図 8 ~ 10 に示す。As に対しては Mn ( $r=0.724$ )、Sc に対し Al( $r=0.864$ )、Ti に対し Al ( $r=0.748$ ) を用いて回帰分析により列欠損を補完した。また補完値に対して 95%信頼区間をエラーバーで表示した。これらの方法によって補完したデータマトリックスは、大気汚染状況の解析により有用と思われる。

尚、今回試行した方法により補完したデータマトリックスは、インターネット web 上で公開する予定である。

### 参考文献

- [1] Junninen, H., Niska, H., Ruuskanen, J., Kolehmainen, M., Tuppurainen, K., Methods for imputation of missing values in air quality data sets, *Atmos. Environ.*, 38(18), 2895-2907 (2004)
- [2] Little, R.J.A., Rubin, D.B., *Statistical analysis with missing data*, Wiley, New York (1987)
- [3] Schafer, J.L., *Analysis of incomplete multivariate data*, Monographs on Statistics and Applied Probability, No.72, Chapman & Hall, London (1997)
- [4] 日本環境衛生センター, 国設大気測定網 (NASN) 浮遊ふんじんおよび浮遊粒子状物質分析結果報告書 (1975-1997)
- [5] Var, F., Narita, Y., Tanaka, S., The concentration, trend and seasonal variation of metals in the atmosphere in 16 Japanese cities shown by the results of National Air Surveillance Network(NASN) from 1974 to 1996, *Atmos. Environ.*, 34(17), 2755-2770 (2000)
- [6] Sekine, Y., Hashimoto, Y., Nakamura, T., Chen, Z.L., Mitsuzawa, S., Characterization of Atmospheric Aerosol Components at Beijing, China, *大気汚染学会誌*, 27(5)237-245 (1992)
- [7] 田中茂, 田村定義, 橋本芳一, 黄砂現象によるアジア大陸からの土壌粒子の移動とわが国に及ぼす影響-NASN データによる考察, *大気汚染学会誌*, 18(3), 263-270 (1983)
- [8] 戸村健児, 国設大気監視網と環境大気調査, *ふんせき*, No.5, 292-300 (1986)
- [9] Willmott, C.J., Some comments on the evaluation of model performance, *Bulletin of the American Meteorol. Soc.*, 63, 1309-1313(1982)
- [10] Willmott, C.J., Ackleson, S.G., Davis, R.E., Feddema, J.J., Klink, K.M., Legates, D.R., O'Donnell, J., Rowe, C.M., Statistics for the evaluation and comparison of models, *J. Geophys. Res.*, 90(C5), 8995-9005(1985)

## Figure Captions

Fig.1 Typical missing rows and missing values in column, commonly found in the SPM and multi-elemental concentration data by NASN.

Fig.2 Performance of different complement methods as a function of length of gap, shown by index of agreement,  $d$ . The tests were performed using simulated missing data pattern of SPM concentration at Nonotake site.

Fig.3 Performance of different complement methods as a function of length of gap, shown by RMSE. The tests were performed using simulated missing data pattern of SPM concentration at Nonotake site.

Fig.4 Performance of different complement methods as a function of length of gap, shown by MAE. The tests were performed using simulated missing data pattern of SPM concentration at Nonotake site.

Fig.5 Performance of linear regression complement method as a function of correlation coefficient,  $r$ , shown by index of agreement,  $d$ . The tests were performed using simulated missing data pattern of Fe concentration at Nonotake site.

Fig.6 Performance of linear regression complement method as a function of correlation coefficient,  $r$ , shown by RMSE. The tests were performed using simulated missing data pattern of Fe concentration at Nonotake site.

Fig.7 Performance of linear regression complement method as a function of correlation coefficient,  $r$ , shown by MAE. The tests were performed using simulated missing data pattern of Fe concentration at Nonotake site.

Fig.8 Variation of atmospheric concentrations of arsenic in Kawasaki, a) observed data with missing values, b) complemented data using as manganese as predictor element.

Fig.9 Variation of atmospheric concentrations of scandium in Kawasaki, a) observed data with missing values, b) complemented data using as aluminum as predictor element.

Fig.10 Variation of atmospheric concentrations of titanium in Kawasaki, a) observed data with missing values, b) complemented data using as aluminum as predictor element.



year	month	SPM ( $\mu\text{g}/\text{m}^3$ )	Element( $\text{ng}/\text{m}^3$ )				
			Ag	Al	As	Ba	Br
1974	4	59		1820		6.0	18
	5			1260		4.4	17
	6	55		1000		5.3	19
	7	40		600		4.6	15
	8	46		1190		6.7	15
	9	37		1210		1.3	26
	10	Missing row					
	11	65		1490		4.1	38
	12	44		950		2.8	30
	1975	1	38		780		0.98
2		32		790		1.8	18
3		39		1370		1.8	18
	4	42		1800		4.9	15
	5	36		940		8.0	14

Missing row
  Missing value in column

Fig.1 Typical missing rows and missing values in column, commonly found in the SPM and multi-elemental concentration data by NASN.

Table 1 Missing rows in the atmospheric SPM and multi-elemental concentration data by NASN

monitoring site	Period		Number of			Number of gap length		
	start	end	rows	missing rows	%	1 month	2 months	8 months
Ube	Apr.1981	Mar.1997	192	7	3.6	7	0	0
Kyoto	Apr.1981	Mar.1997	192	4	2.1	4	0	0
Sapporo	Apr.1974	Mar.1997	276	8	2.9	4	2	0
Matsue	Apr.1974	Mar.1997	276	7	2.5	5	1	0
Niigata	Oct.1975	Mar.1997	258	12	4.7	10	1	0
Sendai	Apr.1974	Mar.1997	276	4	1.4	4	0	0
Kawasaki	Apr.1974	Mar.1997	276	5	1.8	5	0	0
Osaka	Apr.1974	Mar.1997	276	4	1.4	4	0	0
Kurashiki	Apr.1979	Mar.1997	216	2	0.9	2	0	0
Ohmuta	Apr.1974	Mar.1997	276	16	5.8	6	1	1
Chikugo-ogori	Apr.1979	Mar.1997	216	6	2.8	6	0	0
Tokyo	Apr.1974	Mar.1997	276	6	2.2	4	1	0
Amagasaki	Oct.1976	Mar.1997	246	6	2.4	6	0	0
Nonotake	Apr.1981	Mar.1997	192	0	0.0	0	0	0
Nagoya	Oct.1976	Mar.1997	246	4	1.6	4	0	0
Nohoro	Apr.1979	Mar.1997	216	2	0.9	0	1	0

Table 2 Percentage of missing values in columns of atmospheric SPM and multi-elemental concentration data by NASN (Unit: %)

monitoring site	No. of rows	SPM	Ag	Al	As	Ba	Br	Ca	Cd	Ce	Cl	Co	Cr	Cs	Cu	Fe	Hf	K	La	Lu	Mn	Na	Ni	Pb	Sb	Sc	Se	Sm	Th	Ti	V	W	Zn				
Ube	192	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		
Kyoto	192	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
Sapporo	276	0	12	0	12	0	0	12	7	0	0	0	0	0	3	0	1	12	0	1	0	0	1	0	0	12	0	0	0	12	0	1	0	0	0	0	
Matsue	258	0	13	0	36	0	0	13	4	1	5	0	0	0	2	0	3	13	0	2	0	0	0	0	0	13	0	2	2	13	0	0	0	0	0	0	
Niigata	246	0	13	0	37	0	0	13	4	0	5	0	0	0	2	0	3	13	0	2	0	0	0	0	0	13	0	2	2	13	0	0	0	0	0	0	
Sendai	276	0	13	0	13	0	0	13	5	0	4	0	0	1	0	0	3	13	0	3	0	0	0	0	0	13	0	2	2	13	0	0	0	0	0	0	0
Kawasaki	276	0	13	0	13	0	0	13	4	0	0	0	0	0	0	0	1	13	0	2	0	0	0	0	0	13	0	0	0	13	0	0	0	0	0	0	0
Osaka	276	0	13	0	13	0	0	13	5	0	0	0	0	0	0	0	0	13	0	2	0	0	0	0	0	13	0	0	0	13	0	0	0	0	0	0	0
Kurashiki	216	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Ohmuta	276	0	12	0	12	0	0	12	4	0	5	0	0	0	0	0	3	12	0	2	0	0	0	0	0	12	0	1	1	12	0	0	0	0	0	0	0
Chikugo-ogori	216	0	0	0	0	0	0	0	0	0	0	0	0	0	0	5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Tokyo	276	0	13	0	13	0	0	13	5	0	4	0	0	1	0	0	4	13	0	2	0	0	0	0	0	13	0	0	2	13	0	0	0	0	0	0	0
Amagasaki	246	0	2	0	2	0	0	2	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	2	0	2	0	2	0	0	0	0	0	0	0
Nonotake	192	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Nagoya	246	0	2	0	2	0	0	2	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	2	0	0	2	0	0	0	0	0	0	0	0
Nohoro	216	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

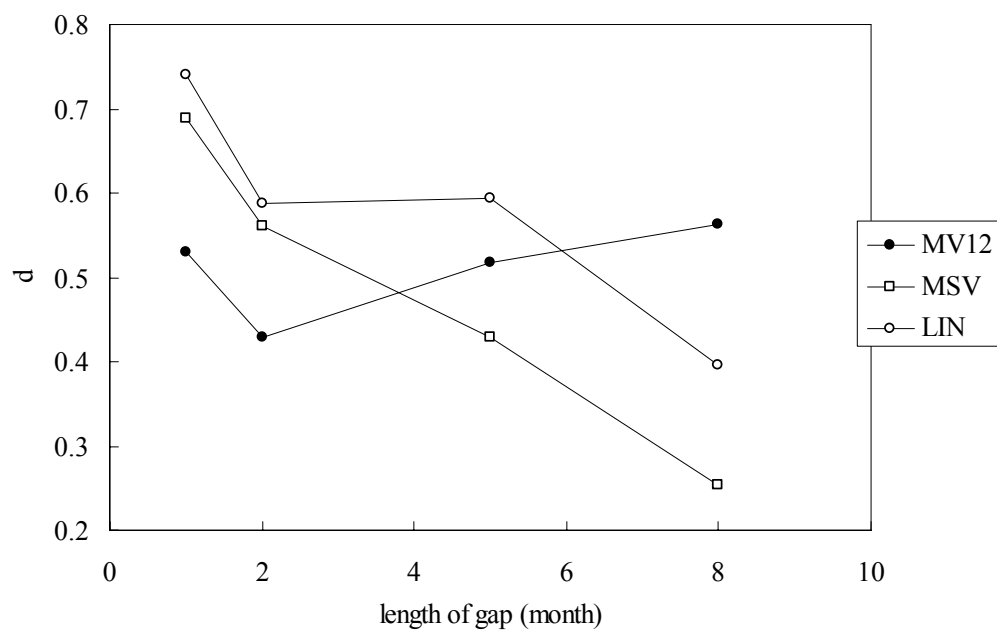


Fig.2 Performance of different complement methods as a function of length of gap, shown by index of agreement,  $d$ . The tests were performed using simulated missing data pattern of SPM concentration at Nonotake site.

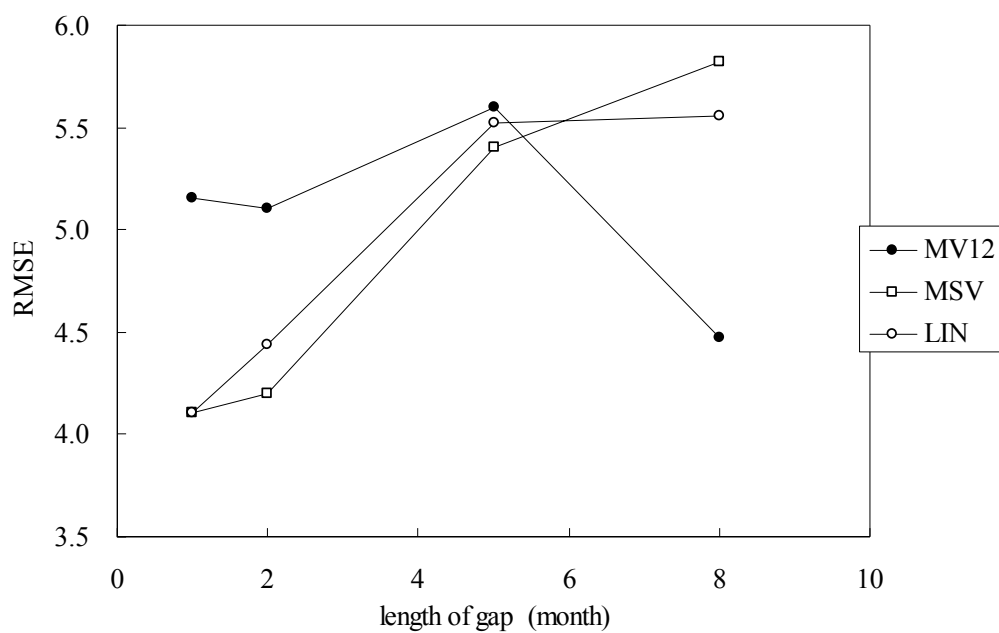


Fig.3 Performance of different complement methods as a function of length of gap, shown by RMSE. The tests were performed using simulated missing data pattern of SPM concentration at Nonotake site.

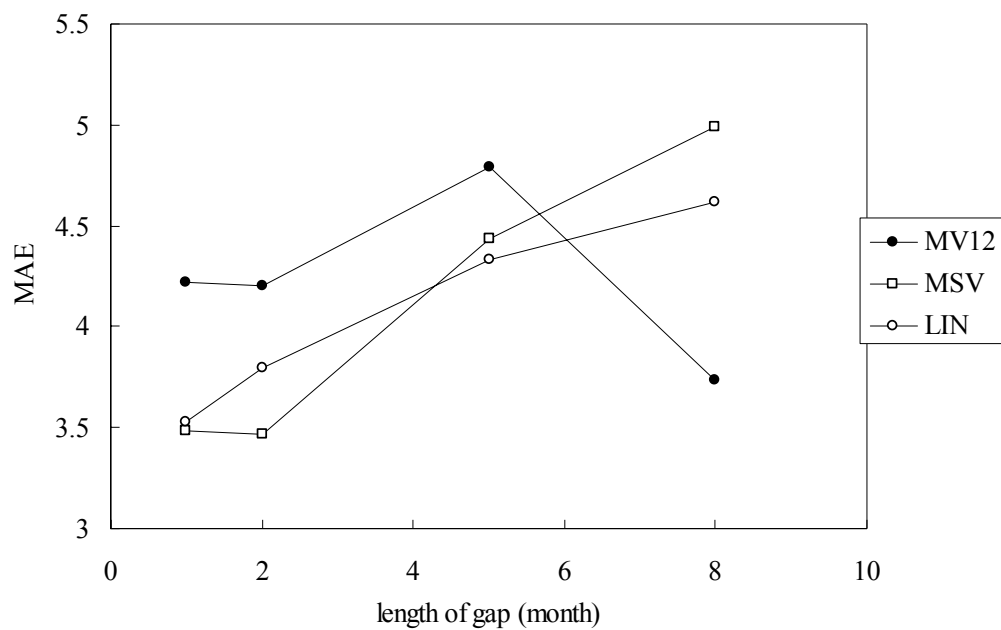


Fig.4 Performance of different complement methods as a function of length of gap, shown by MAE. The tests were performed using simulated missing data pattern of SPM concentration at Nonotake site.

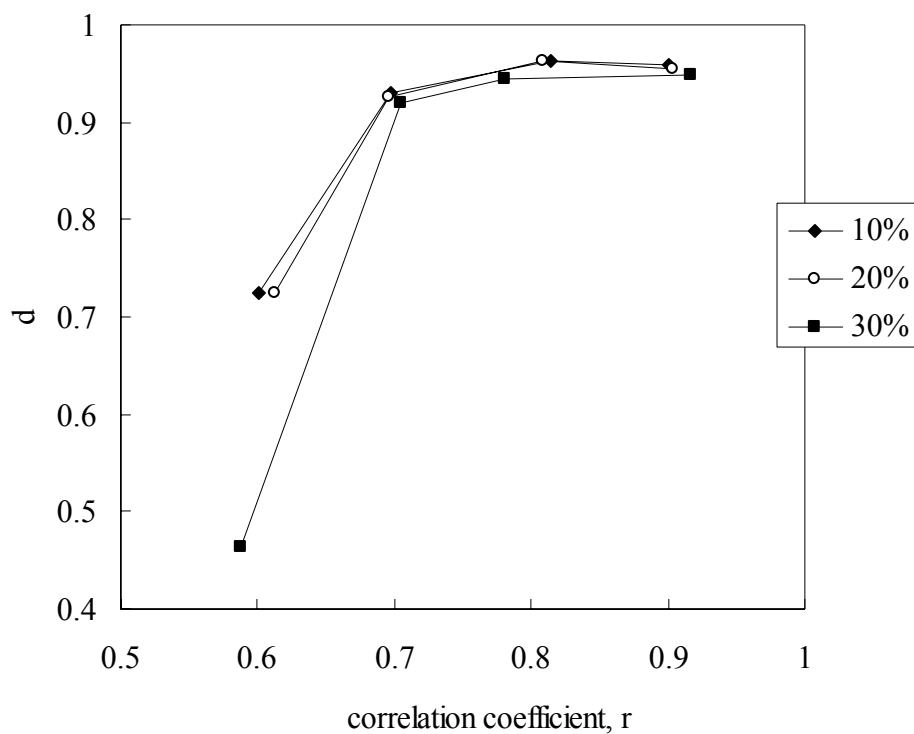


Fig.5 Performance of linear regression complement method as a function of correlation coefficient,  $r$ , shown by index of agreement,  $d$ . The tests were performed using simulated missing data pattern of Fe concentration at Nonotake site.

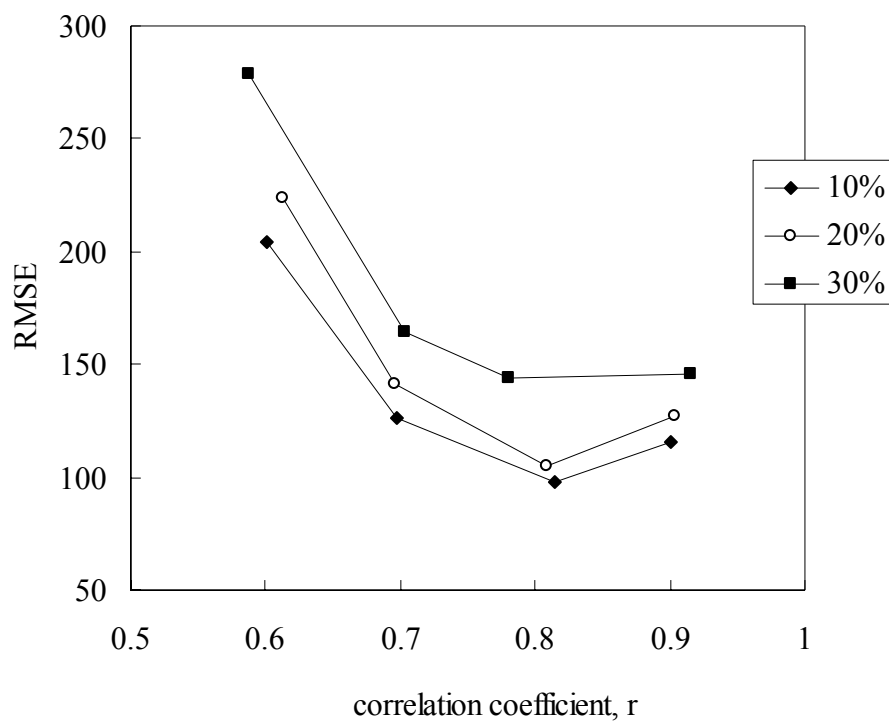


Fig.6 Performance of linear regression complement method as a function of correlation coefficient,  $r$ , shown by RMSE. The tests were performed using simulated missing data pattern of Fe concentration at Nonotake site.



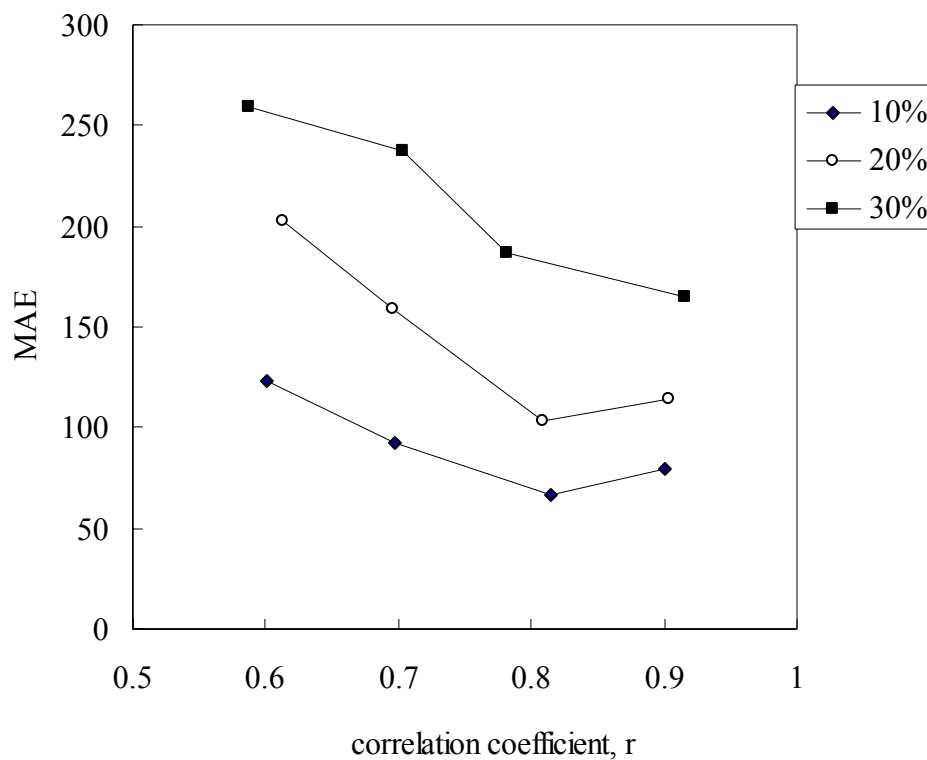


Fig.7 Performance of linear regression complement method as a function of correlation coefficient,  $r$ , shown by MAE. The tests were performed using simulated missing data pattern of Fe concentration at Nonotake site.

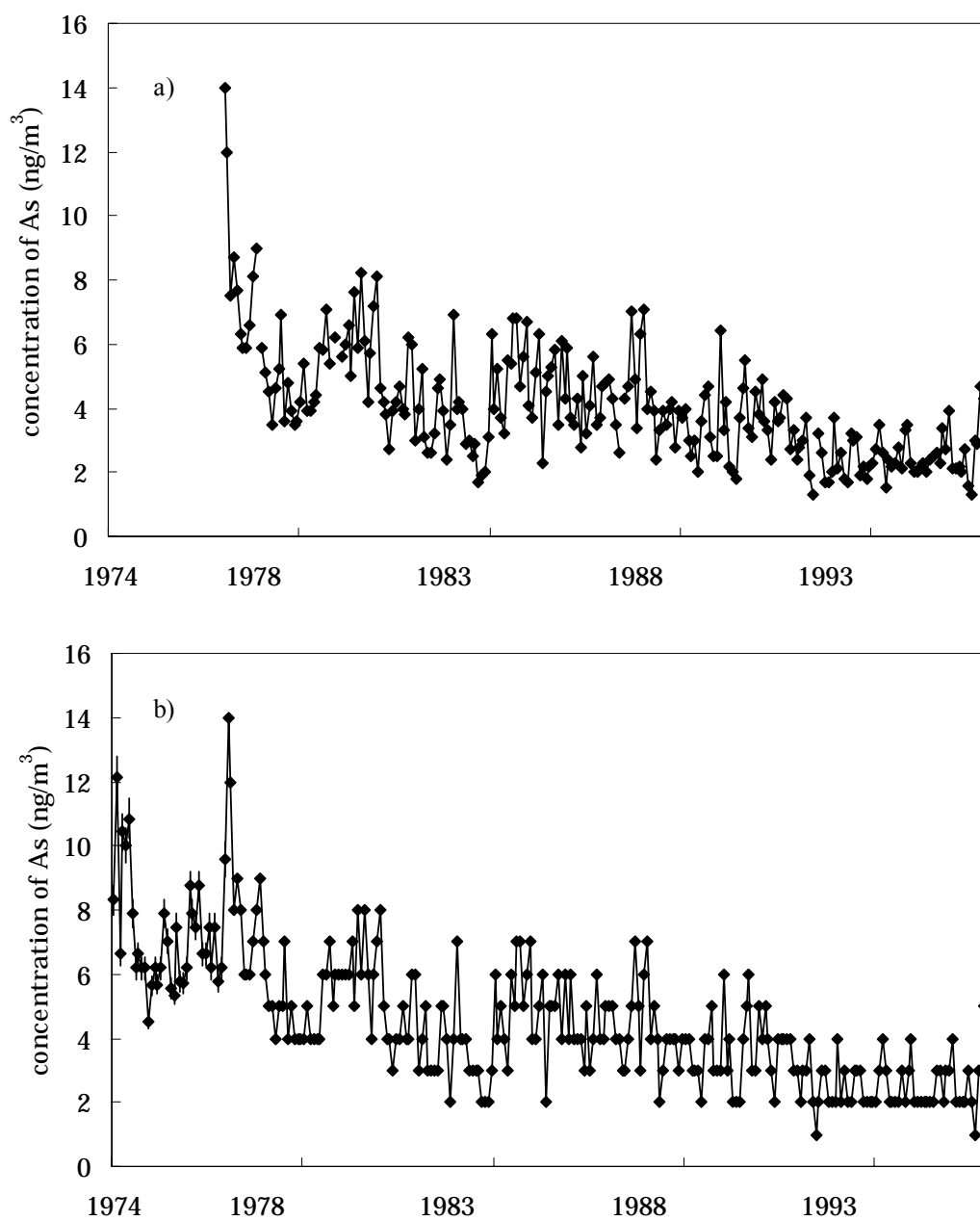


Fig.8 Variation of atmospheric concentrations of arsenic in Kawasaki, a) observed data with missing values, b) complemented data using as manganese as predictor element.

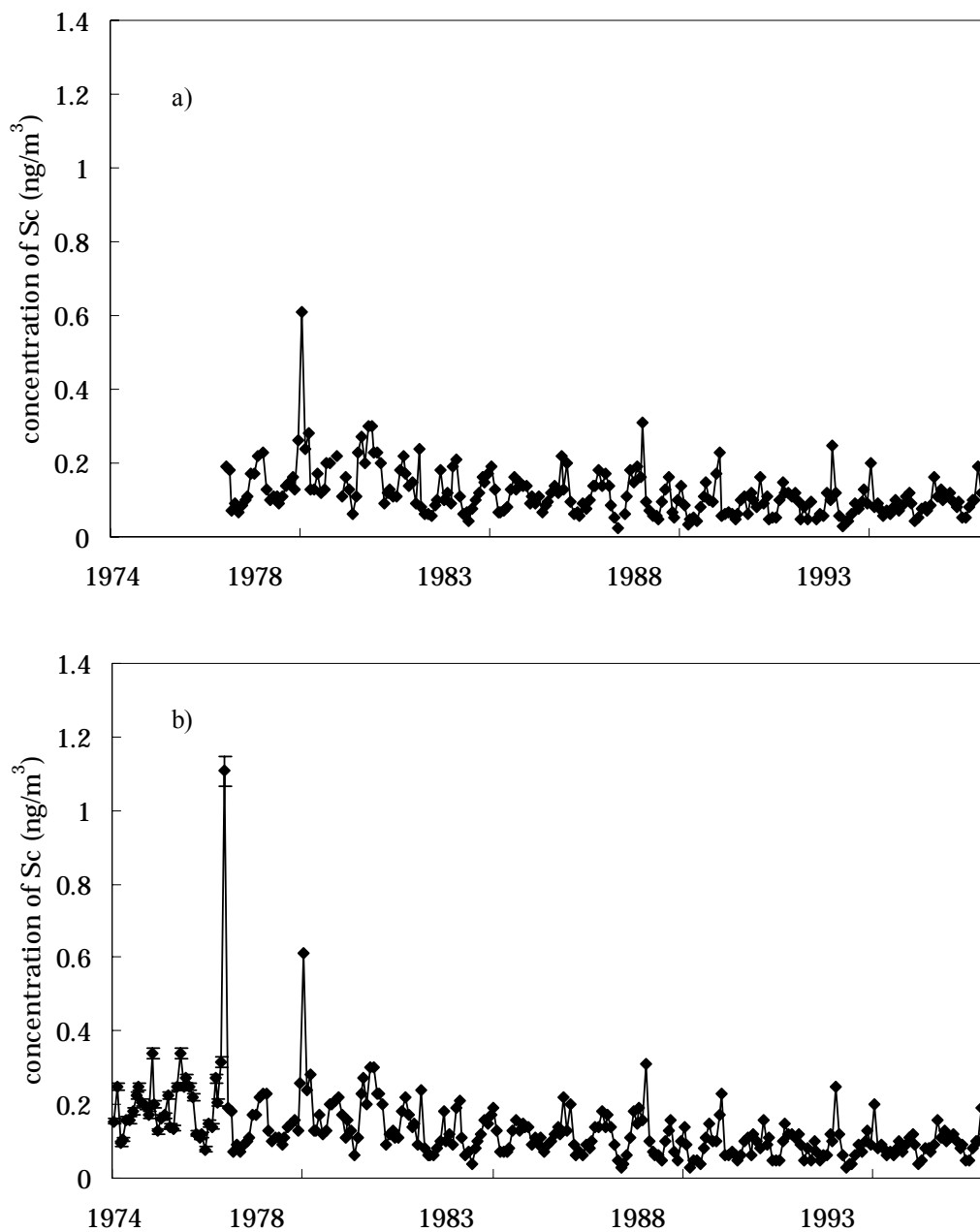


Fig.9 Variation of atmospheric concentrations of scandium in Kawasaki, a) observed data with missing values, b) complemented data using as aluminum as predictor element.

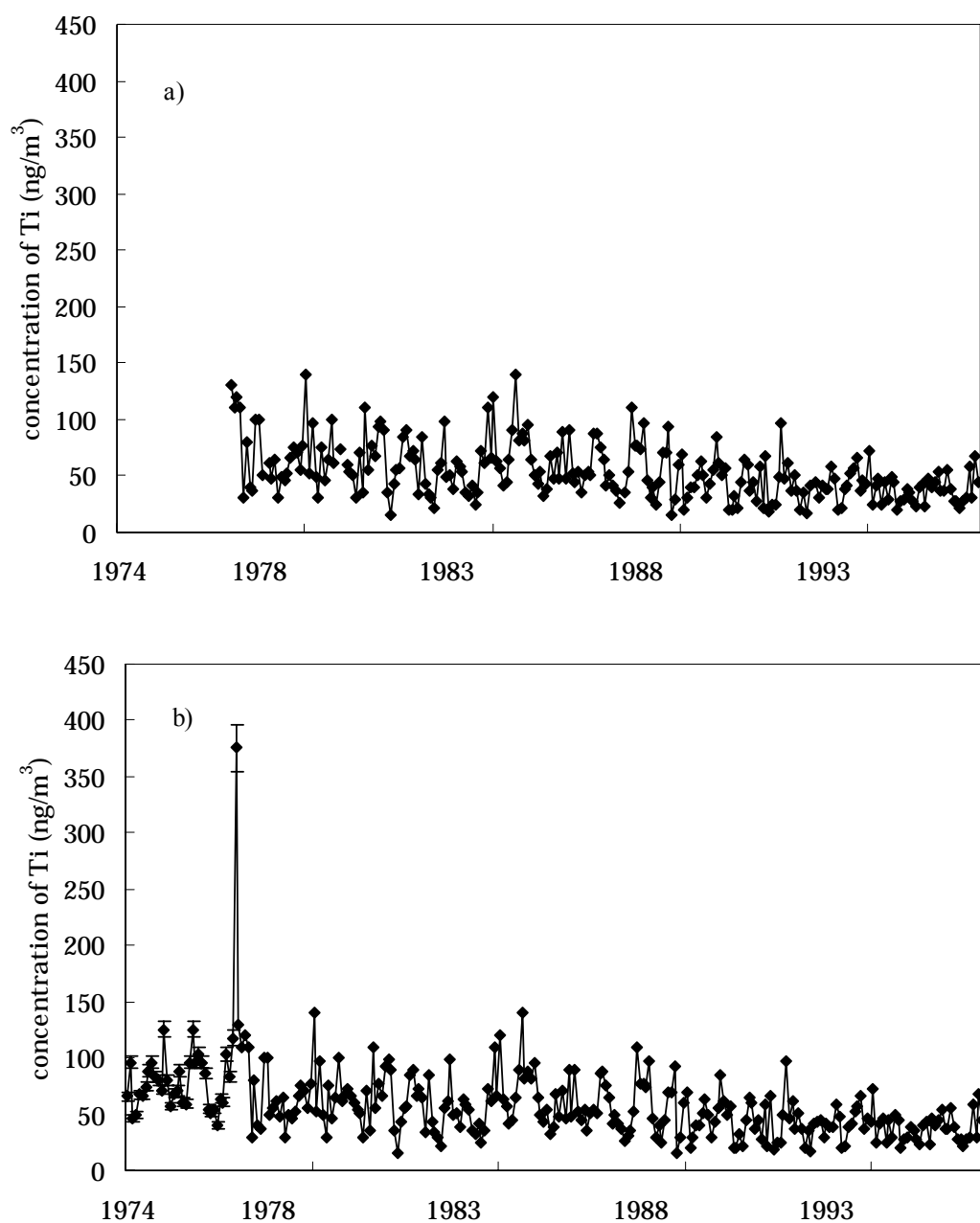


Fig.10 Variation of atmospheric concentrations of titanium in Kawasaki, a) observed data with missing values, b) complemented data using as aluminum as predictor element.