



PERGAMON



Atmospheric Environment 35 (2001) 3155–3166

**ATMOSPHERIC
ENVIRONMENT**

www.elsevier.com/locate/atmosenv

Delineation of air-quality basins utilizing multivariate statistical methods in Taiwan

Tai-Yi Yu*, Len-Fu W. Chang

Graduate Institute of Environmental Engineering, National Taiwan University, 71, Chou-Shan, Taipei, Taiwan

Received 11 May 2000; received in revised form 29 October 2000; accepted 7 November 2000

Abstract

This study analyzed time-series data of air pollutants, O_3 and PM_{10} , to determine the division of air-quality basins in Taiwan by employing multivariate statistical methods, Varimax rotational method and cluster analysis. The databases of air pollutants, daily maximum 1-h O_3 and daily mean PM_{10} concentrations, were obtained from the ROC Environmental Protection Administration (ROC EPA) for the period from 1 July 1993 to 30 June 1998. The Varimax rotational method allowed us to delineate five homogenous PM_{10} subregions that cumulatively accounted for 85.6% of the total variance. The time-series analysis of rotated component scores associated with the PM_{10} subregions revealed that all divided subregions presented a strong seasonal cycle. Four of five subregions had higher component scores and PM_{10} concentrations from November to January. One subregion experienced higher values from March to May. The use of Varimax approach and cluster analysis on the O_3 and PM_{10} confirmed that O_3 was more demonstrative of the air-quality basins in Taiwan. Both the Varimax rotational method and the cluster analysis have specific advantages for the division of air-quality basins. This study also proposes a delineation of five air-quality basins having homogenous O_3 features as an alternative assignment of atmospheric carrying capacity control regions. © 2001 Elsevier Science Ltd. All rights reserved.

Keywords: Air-quality basins; Varimax rotational method; Cluster analysis; O_3 ; PM_{10}

1. Introduction

Air-quality basins are normally determined by meteorological and geographic conditions as well as political boundaries. Weather patterns, mountain topography, locations of emission sources, and differences in land use also play major roles in the formation and accumulation of secondary air pollutants. O_3 and portions of PM_{10} are secondary air pollutants, in which major source regions do not significantly influence its spatial behavior in the rural boundary layer (Vukovich and Fishman, 1986). Rather, the path and frequency of anticyclones dominate because of the requisite meteorological conditions (i.e., warm temperature, clear skies and stagnant conditions).

Scientists have recognized O_3 as a regional (Logan, 1989) and even global phenomenon (Liu et al., 1987). Thus, O_3 and PM_{10} could help to identify separated air-quality basins through multivariate statistical analysis of measured data.

The ROC EPA Air Pollution Control Act (ROC EPA, 1999) introduced a program on 20 January 1999 to prevent significant deterioration of air quality, maximum allowable incremental concentrations of air pollutants, and auditing rules for new or updated stationary sources, for the atmospheric carrying capacity control regions. The ROC EPA has the authority to assign atmospheric carrying capacity control regions according to weather and geographic conditions. The delineation of air-quality basins, which exhibit homogenous O_3 or PM_{10} concentrations, is an alternative method for forming atmospheric carrying capacity control regions. Many studies have adopted unrotated principal component analysis (UPCA) to examine the spatial patterns of meteorological

* Corresponding author. Tel.: + 886-2-8787-9037; fax: + 886-2-2362-6243.

E-mail address: d2507005@ms31.hinet.net (T.-Y. Yu).

measurements (Kidson, 1975; Wallace and Gutzler, 1981; Walsh and Mostek, 1981; Flocchini et al., 1981) and source contribution of air pollution measurements (Gaarenstroom et al., 1977; Gatz, 1978; Alpert and Hopke, 1980, 1981; Pitchford et al., 1981). However, Buell (1975) indicated the shape of the boundary area to which the correlation coefficients belong often helps to determine the topography of the unrotated principal components. Therefore, the UPCA is adopted mainly for data reduction and can be somewhat misleading when applied for spatial delineation. Eder (1989) and Eder et al. (1993) applied the Varimax rotational technique to analyze the SO_4^{2-} concentrations in precipitation and daily 1-h maximum O_3 concentrations over non-urban areas of the eastern US (the station numbers were 40 and 77) to determine the spatial features of air pollution measurements. This application allowed the former investigators to delineate seven and six homogenous subregions, accounting for 74.2 and 64.02% of the total variance, respectively. Ashbaugh et al. (1984) utilized the Varimax rotational technique to obtain the spatial patterns of 40 inter-site correlations of sulfur concentrations in the western United States. The first five components were identified with 57.3% of the total variance in the data. The first and second components were attributed to copper smelter emissions and episodic incursion of sulfur. Juang et al. (1996) also utilized the Varimax rotational method to analyze the combined data set of five air pollutants, SO_2 , PM_{10} , CO , O_3 and NO_x , to delineate seven air basins with the monitoring data of 1994. Yu and Chang (1999) demonstrated the spatial and temporal features of O_3 concentrations and component scores over Taiwan via the UPCA and Varimax rotational approaches. Most importantly, the Varimax rotational technique supplies a statistical, objective, and better physical interpretation that enables PM_{10} and O_3 to be divided into homogenous subregions.

The cluster analysis is another effective multivariate statistical method that splits a data set into several similar groups. Kalkstein et al. (1987) grouped similar synoptic weather patterns with three varying cluster procedures, while other scientists (Fernau and Samson, 1990a,b) also applied cluster analysis to define specific periods with similar meteorological data and precipitation chemistry. Before performing the division of air-quality basins, selecting clustering algorithms, specifying the measured distance and choosing adequate grouping numbers are the major features of the objective classification of cluster analysis.

This investigation continues the work of Yu and Chang (1999) by utilizing the Varimax rotational method and the cluster analysis to determine the divided subregions that exhibit homogenous O_3 or PM_{10} concentrations. Moreover, it also ascertains the better criteria pollutant and the better multivariate approach for the delineation of air-quality basins. This study also presents

a novel method to separate air-quality basins through two multivariate statistical methods of the spatial and temporal variability of O_3 and PM_{10} concentrations found in Taiwan. Results in this study will hopefully stimulate the implementation of an air pollution control program on atmospheric carrying capacity.

2. Method

The air-quality-monitoring network continuously monitors the air quality via its 72 stations spread over Taiwan. The major air pollutants that are monitored include SO_2 , NO , NO_2 , CO , O_3 , PM_{10} and NMHC. Fig. 1 illustrates the locations of the monitoring stations and the political boundaries in Taiwan. The historical data of daily maximum 1-h values of O_3 and daily mean values of PM_{10} concentrations are compiled into two distinct databases, lasting from 1 July 1993 to 30 June 1998. These data sets can be viewed mathematically as a time series of high-dimensional vectors, each of which has 63 O_3 components and 71 PM_{10} components (i.e. the number of total observational stations with measured O_3 and PM_{10} data).

2.1. The Varimax rotational method

The Varimax rotational method, as developed by Kaiser (1958), increases the segregation between component loadings and more clearly defines a distinct clustering of intercorrelated data and ultimately makes spatial interpretation easier. This technique rotates the predetermined principal components while maintaining that the individual components remain orthogonal to each other. Horel (1981) indicated that the UPCA and the Varimax rotational approaches maximize the sum and the variance of squared correlation coefficients, respectively. The front analysis maximizes the first moment statistic (the sum) that causes the correlation coefficients to be poorly distributed (i.e., many loadings with similar magnitudes), resulting in spatial patterns that are difficult to interpret. Maximizing the second moment (the variance) causes the correlation coefficients to be widely distributed (i.e., many loadings close to zero) in an easier explanation for spatial patterns.

The analysis is initiated by extracting a square, symmetrical correlation matrix (63 columns \times 63 rows for O_3 and 71 columns \times 71 rows for PM_{10}), the original data matrix was 63 stations \times 1826 days for O_3 and 71 stations \times 1826 days for PM_{10}). The average values of daily maximum 1-h values of O_3 and daily mean values of PM_{10} concentrations at respective stations were substituted for the lost data. The normalized value is (1)

$$Z_{ik} = \frac{C_{ik} - \mu_i}{S_i}, \quad (1)$$

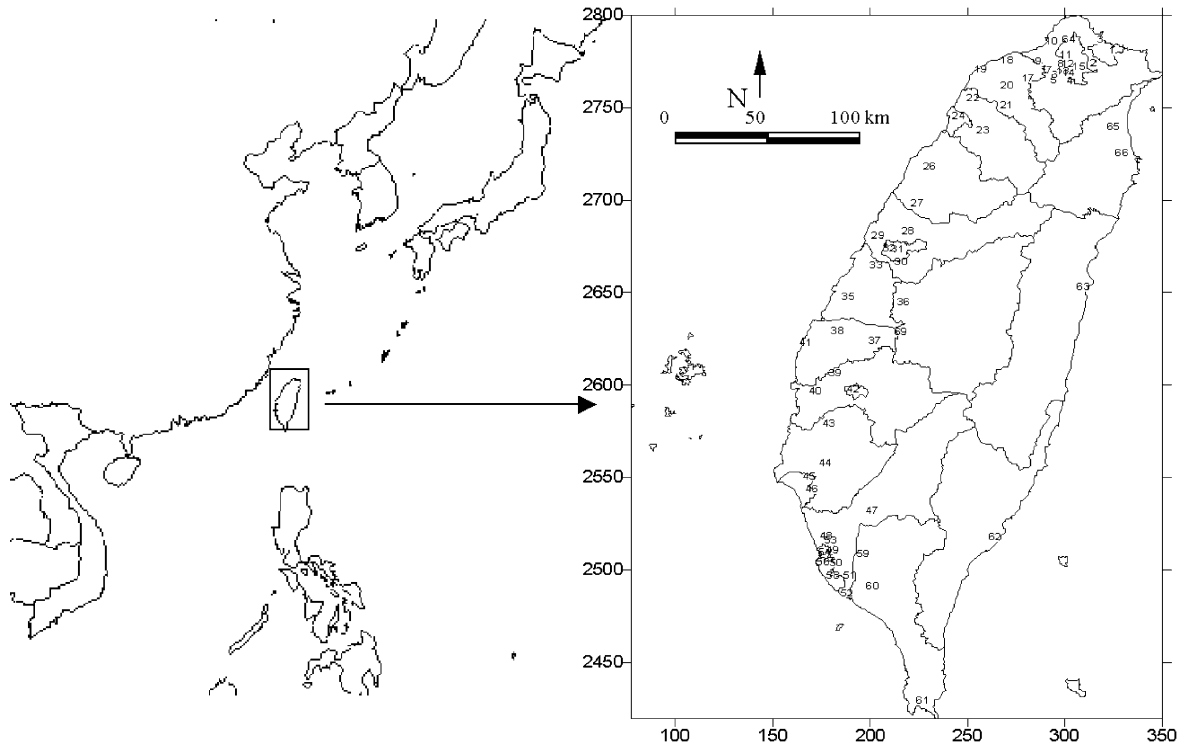


Fig. 1. Locations of monitoring stations and political boundaries over Taiwan (Taiwan, ROC).

where Z_{ik} denotes the normalized value of the k th observation on the i th stations, C_{ik} represents the k th O_3 or PM_{10} concentration of the i th station, μ_i is the mean value of the i th station, and S_i denotes the standard deviation of the i th station. A correlation matrix is significantly more appropriate than a covariance matrix for resolving spatial oscillations (Overland and Preisendorfer, 1982). Moreover, a correlation matrix presents the isopleths of component loadings, which can be regarded as the correlation coefficients between the rotational component and individual stations. In sum, the Varimax rotational method aims to obtain the maximum T value, which is represented as

$$T = n \sum_{j=1}^n \sum_{i=1}^n \left(\frac{A_{ij}}{h_i} \right)^4 - \sum_{j=1}^n \left(\sum_{i=1}^n \frac{A_{ij}^2}{h_i^2} \right)^2, \quad (2)$$

$$h_i^2 = \sum_j \frac{A_{ij}^2}{h_i^2},$$

$$Z_{ik} = \sum_{j=1}^n A_{ij} R_{jk},$$

where n denotes the number of stations, A_{ij} represents the loading of the i th station on the j th rotated principal component, h_i^2 is the communality of the i th station, and R_{jk} stands for the component score of the k th variable for

the j th rotated principal component. The standardized scores of the rotated components are symbolized by

$$P_{jk} = R_{jk} / \sqrt{\lambda_j}, \quad (3)$$

where λ_j denotes the eigenvalue or variance of the j th rotated component and P_{jk} represents the standardized scores of the k th variable on the j th rotated principal component.

In the division of air-quality basins, features of geographic and emission inventories are two dominant factors. Fig. 2 represents the spatial patterns of these two factors over Taiwan. Yu and Chang (1999) utilized the Varimax rotational technique to divide Taiwan into five homogenous O_3 subregions (Fig. 3) that account for 75.5% of the total variance. The O_3 subregions are termed Midwestern Taiwan (M-W-T), Tainan-Kaohsiung-Pingtung (T-K-P), Taipei-Keelung (T-K), Tao-Yuan (T-Y), and eastern Taiwan (E-T) in decreasing order of explained contribution. This investigation delineates five homogenous PM_{10} subregions that account for 85.6% of the total variation. The influence regimes of the first five PM_{10} rotated components, P1, P2, P3, P4 and P5, are depicted in Fig. 4. The isopleths of component loadings were determined with Kriging approach, a useful geostatistical gridding method. Five maps of the first

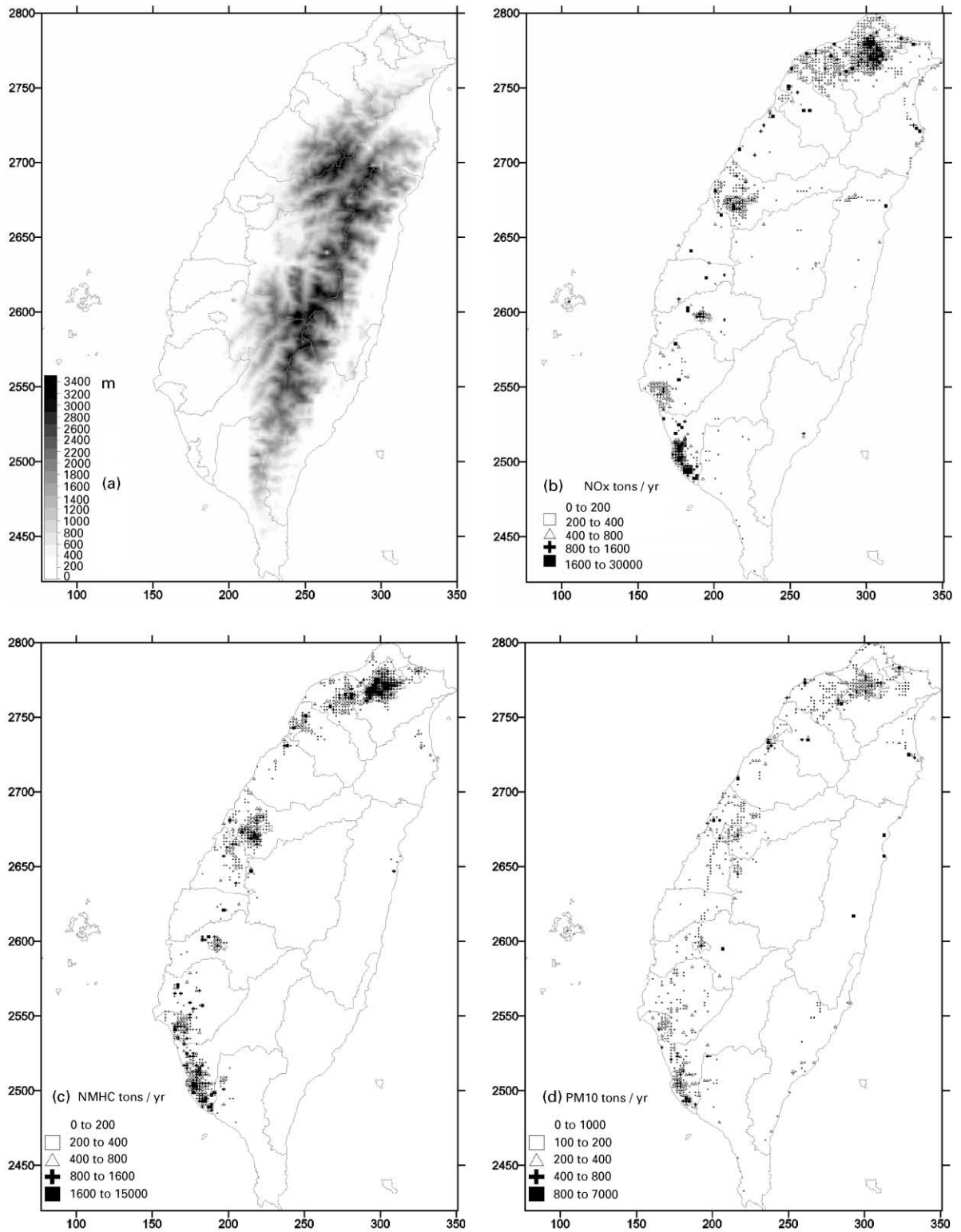


Fig. 2. (a) Topography, (b) NO_x emission inventory over Taiwan, (c) NMHC, and (d) PM₁₀ emission inventory over Taiwan (Source: ROC EPA, 1997).

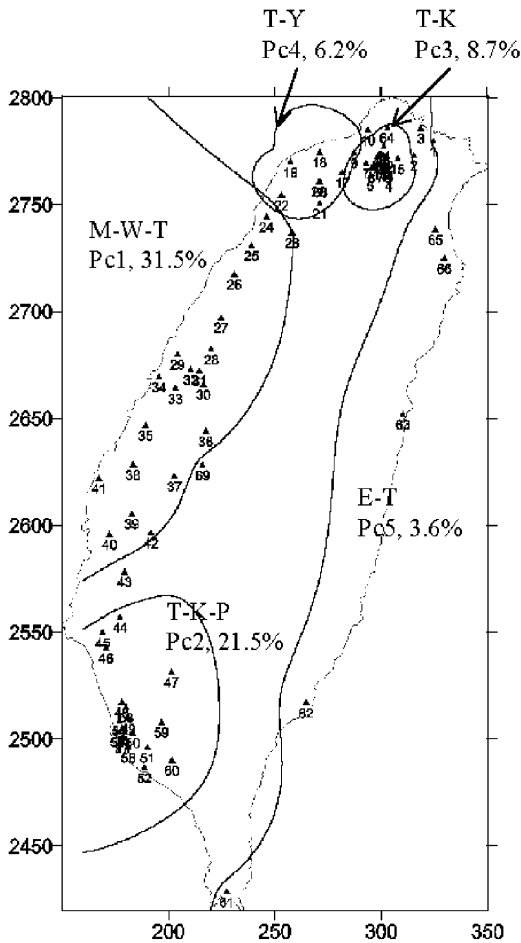


Fig. 3. Five homogenous O_3 subregions (Yu and Chang, 1999).

five components can be incorporated into one figure by plotting the suitable loading of each station on their respective rotated principal components. Each of the five rotated principal components identifies an influence regime that distinguishes a segregated area of study domain. Accounted variances of the five components are 39.2, 29.3, 7.7, 5.3 and 4.1%. The PM_{10} subregions, as delineated by the 0.5 or 0.6 component loading isopleths, exhibit insignificant overlapping.

2.2. A time-series analysis of the rotated components

Time-series analysis is an effective means of characterizing the relationships between air pollutants' concentrations and the rotated principal components. This approach has been successfully applied in the aerometric data, including SO_4^{2-} concentrations in precipitation (Eder, 1989), SO_2 ambient air concentrations (Ashbaugh et al., 1984) and O_3 ambient air concentrations (Eder et al., 1993). The temporal variation of each subregion can

be assessed by examining the standardized scores of the rotated principal components: the standardized scores are weighted summed values whose magnitudes depend on the daily mean PM_{10} concentrations and the loadings of rotated components. All of the rotated component scores have a mean of zero and a standard deviation of one since the scores are standardized. The positive and negative scores correspond to higher and lower values than average PM_{10} concentrations.

The daily principal component scores accurately reflect the temporal variance of each subregion when plotted as a time series (Fig. 5). Such information also provides a valuable reference for scientists to select PM_{10} episodes where the highest PM_{10} concentrations are observed. According to our results, the optimum period for simulating potential reductions in PM_{10} concentrations resulting from various emission strategies is to be determined. For example, model designers could determine the optimum simulation period for the P1 subregion by examining the daily time series of component scores associated with the first rotated principal component. The occurrence rates of standardized principal component scores greater than 1, 2, 3, 4, 5, 6 and 7 are 37.7, 29.8, 23.0, 17.5, 12.4, 8.9 and 6.0% during 5 yr period. The time series indicates that 53 days in 1994 had standardized principal component scores greater than 5. No other year for this subregion recorded more days where the standardized component scores were greater than 5. In addition, 11 of the 53 days occurred within a month, from 1 April to 30 April, representing the most intensive PM_{10} episode experienced by the P1 subregion and, subsequently, the optimum simulation period.

The seasonal cycles were depicted by the time series of median values of component scores associated with the five subregions over a 5-yr period. (Fig. 5f) 30-day moving average values of the median score were chosen to enhance the characterization of the seasonal variation for each PM_{10} subregion. The standardized component scores of subregions P2, P3, P4 and P5 were positive values from October to January. All PM_{10} subregions had negative component scores in summer, from June to August. The P1 subregion exhibits positive component scores during spring, from March to May.

The statistics for each PM_{10} subregion were summarized as the box plots in Fig. 6. The box plots included the 5th, 10th, 25th, 50th, 75th, 90th and 95th percentile values. PM_{10} concentrations in the P2 subregion are considerably higher than in the other subregions across all percentiles from September to March. Correspondingly, the percentile of days that concentrations exceeded 100 (35.9%) and $125 \mu g m^{-3}$ (20.3%) are also the greatest in the entire domain (Table 1). The P1 subregion had the highest PM_{10} concentrations during summer, while P3 displayed second highest concentrations for the percentage of days exceeding 100 (30.8%) and $125 \mu g m^{-3}$ (11.4%). The P1, P2, P3 and P4 subregions had the

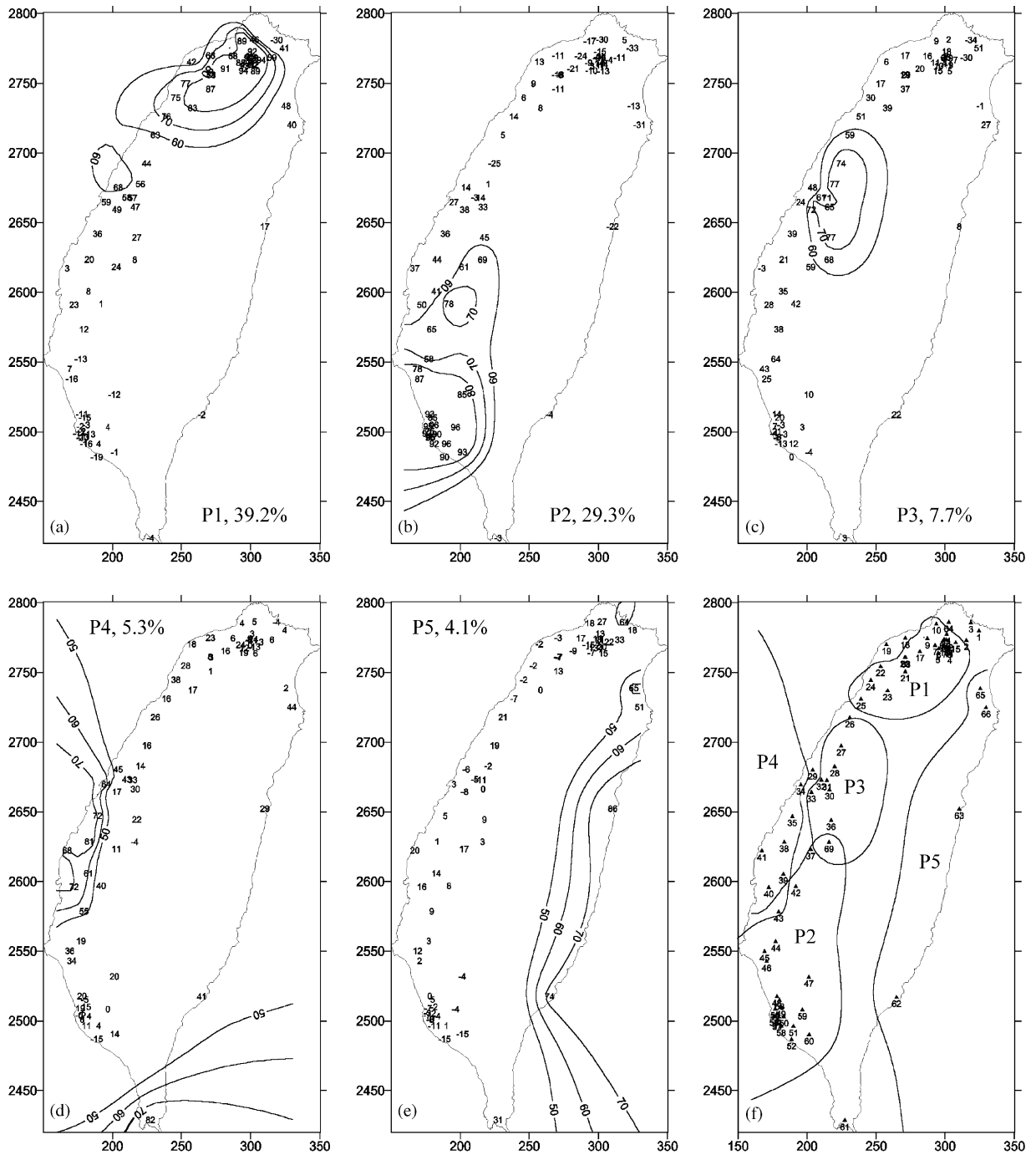


Fig. 4. (a)–(e) PM₁₀ component loadings ($\times 100$) associated with five rotated principal components. (f) Five homogenous PM₁₀ concentration regions delineated by the maximum component loadings.

highest PM₁₀ concentrations in April, December, November and November, respectively. Subregions P1, P2, P3 and P4 experienced the minimum PM₁₀ concentrations during February, June, June and June, respectively.

2.3. Cluster analysis

The cluster analysis categorizes different groups according to similarities or distances. Euclidean distance is the chosen scale of similarity for normalized data of air

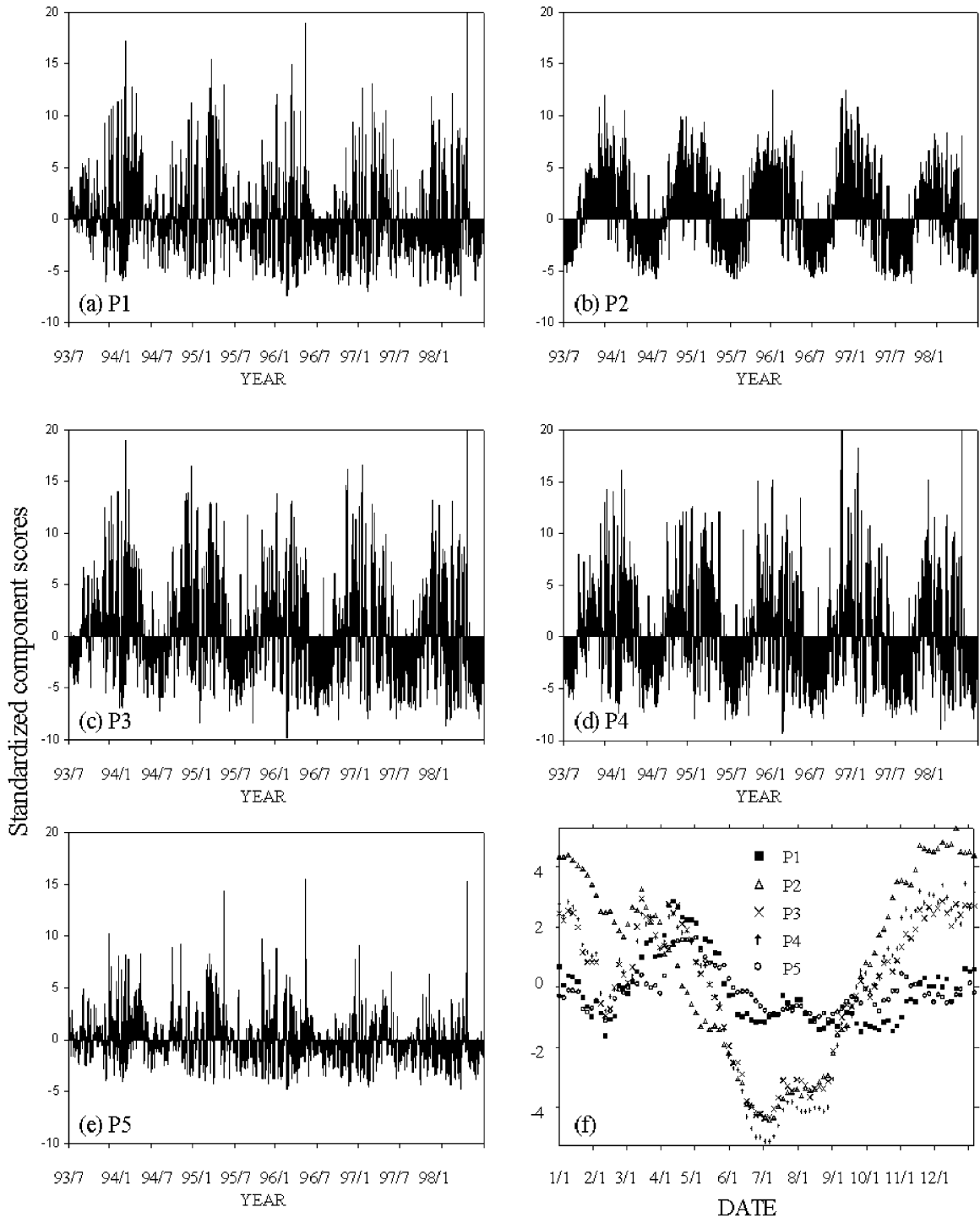


Fig. 5. (a)–(e) Daily time-series of the standardized principal component scores associated with the five homogenous PM_{10} subregions. (f) Seasonal time-series as defined by the 30-day moving values of median principal component scores for each of the homogenous PM_{10} subregions.

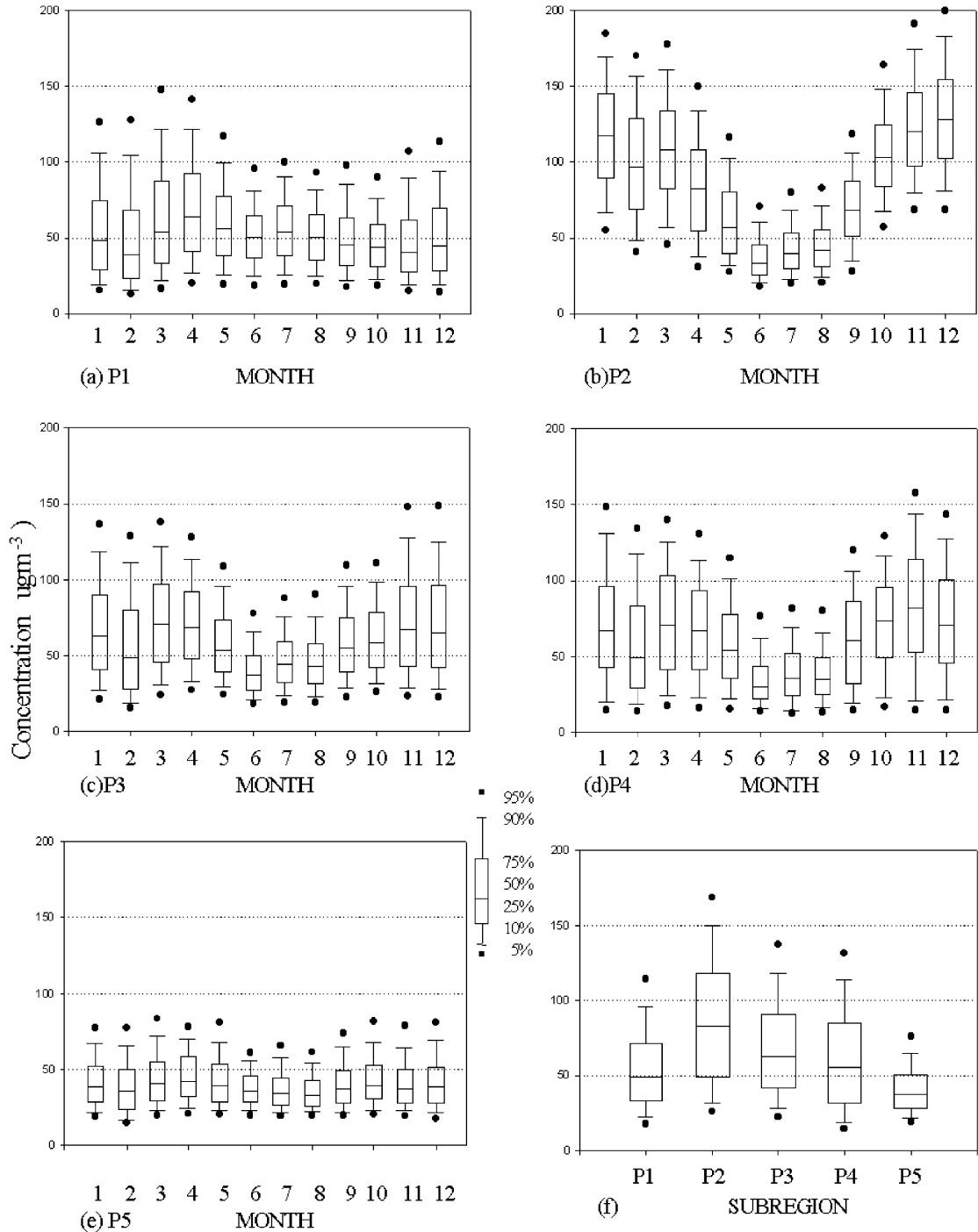


Fig. 6. (a)–(e) Boxplots of the monthly PM₁₀ concentrations associated with the five PM₁₀ subregions. (f) Boxplots of the PM₁₀ concentrations associated with the five subregions.

pollutants as well as the daily maximum 1-h O₃ values and the daily mean PM₁₀ concentrations. Ward’s minimum variance clustering method, which merges two clusters that result in the smallest increase in the sum-of-

squares, was the clustering approach employed herein. The plots of percentage change in total root-mean-square deviation (TRMSD) against the number of clusters provide the criteria for determining the adequate

number of clusters (Dorling et al., 1992). The TRMSD is recalculated as the sum of root-mean-square deviations of normalized values of air pollutants from its cluster average at every step. The percentage change in TRMSD values, with the number of O₃ and PM₁₀ clusters, is presented in Fig. 7. Dorling et al. (1992) indicated that adequate numbers of clusters could be located as sudden breaks or large percentage changes during the clustering process because the merging of clustered trajectories have significantly different wind directions and wind speeds. As the cluster numbers of O₃ and PM₁₀ are greater than five and four, we can gain limited explained variances through adding the number of clusters. Therefore, four and five clusters of O₃ and PM₁₀ were employed to analyze the spatial features of Taiwan's monitoring station.

The spatial evaluation of four and five clusters of O₃ and PM₁₀ could be an appropriate method to choose the better criteria pollutant for delineating air quality basins via the cluster analysis. The difference, as obtained

Table 1
Summary statistics of daily mean PM₁₀ ($\mu\text{g m}^{-3}$) concentrations associated with each PM₁₀ subregion

PM ₁₀ subregion	N	Mean	% > 100	% > 125
P1	49,740	55.3	5.85	2.31
P2	34,280	85.6	35.88	20.31
P3	11,807	69.3	30.80	11.42
P4	11,932	61.9	16.17	6.51
P5	8812	41.4	1.09	0.27
Domain	11,6571	65.3	17.91	8.80

from Fig. 8a and b, between four and five clusters of O₃ is that the later diagram delineates the northwestern part of Taiwan as two isolated groups. The division of O₃ subregions into five clusters via the Varimax rotational method has a similar result in Fig. 3, except that stations 23, 24 and the T-K O₃ subregion were classified as the same cluster and stations 44, 45, 46 and the O₃ M-W-T subregion were categorized as the same cluster. Fig. 8c and d illustrate that some stations (46, 3, 64, 34 and 1) within the same political boundary were arranged into distinct groups depending on whether they had four or five PM₁₀ clusters. The division of PM₁₀ subregions into five clusters via the Varimax rotational method has similar delineating results, except for the small overlap on the boundaries of the P1, P2, P3 and P4 subregions. The difference between four and five clusters of PM₁₀ is that the later graph divides the northwestern and middle part of Taiwan into two diverse groups.

O₃ is a more suitable pollutant than PM₁₀ for the division of air-quality basins according to both the Varimax method and cluster analysis. The Varimax rotational method provides the accounted variances of rotated components, correlation coefficients between monitoring stations and rotated components, and optimum simulation period for different subregions. The cluster analysis provides the quantified criteria for the selection of cluster numbers and the hierarchy of similarities. Thus, both methods have their specific advantages.

2.4. The division of air quality basins

Five air-quality basin divisions are proposed for Taiwan based on political boundaries as well as the

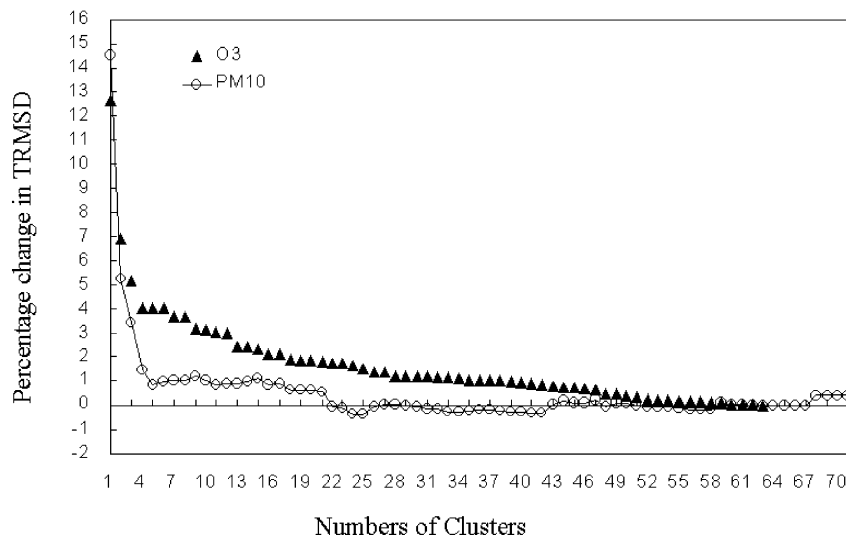


Fig. 7. The percentage change in TRMSD with the numbers of clusters.

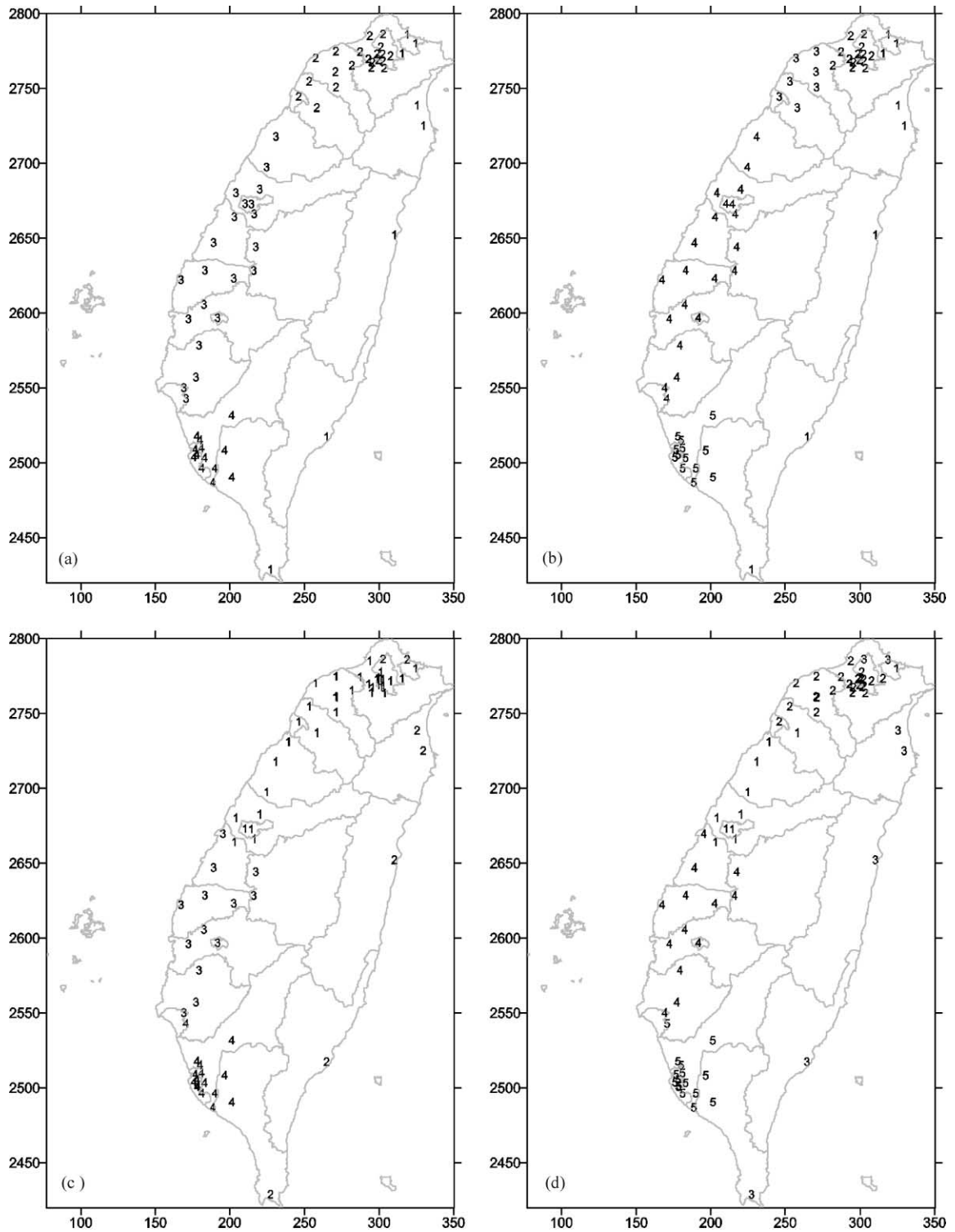


Fig. 8. (a) Four; (b) five clusters for O_3 ; (c) four; and (d) five clusters for PM_{10} , the same numbers marked on the position of monitoring sites present the same clusters.

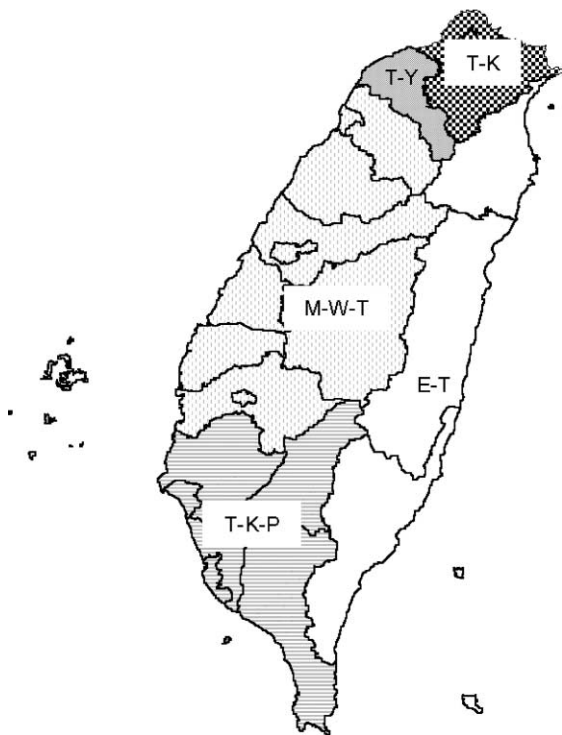


Fig. 9. Five proposed air-quality basins over Taiwan.

O₃ subregions separated by the Varimax rotational method and the cluster analysis (Fig. 9). These air-quality basins are termed as T-K, T-Y, M-W-T, T-K-P and E-T. Local governments within the same air-quality basin should devote their efforts to enhance air quality. Evaluating the relationships between the boundary lines of five air-quality basins, mountain topography and emission inventory of NO_x and NMHC, we gained the following results. First, geographic condition separated Taiwan into eastern and western parts. Secondly, spatial patterns of emission inventory divided western Taiwan into four distinct air-quality basins. On the basis of the delineation of measured ozone data, 10 counties located within M-W-T air-quality basins have the highest priorities to precede emission abatement strategies of air pollutants, with a special focus on O₃ problems.

3. Summary

This study examined the spatial and temporal variabilities of O₃ and PM₁₀ concentrations over Taiwan from July 1993 to June 1998 via the Varimax rotational method and cluster analysis. The Varimax approach identified five homogenous PM₁₀ subregions, accounting for 85.6% of the total variation and each sub-

region exhibited unique PM₁₀ concentrations. Analysis of the rotated principal component scores as well as the actual PM₁₀ concentrations revealed a strong seasonal cycle in the P1, P2, P3 and P4 subregions. One of five PM₁₀ subregions presented its higher component scores and PM₁₀ concentrations during spring. Other four PM₁₀ subregions showed their highest component scores and PM₁₀ concentrations from November to January.

The division of air-quality basins should be based on O₃ instead of PM₁₀ since: some stations within the same political area were attributed to varying PM₁₀ clusters and subregions, and a small overlap exists among different PM₁₀ subregions created by the Varimax approach. This study divided Taiwan into five air-quality basins based on the political boundary lines and the Varimax method and cluster analysis on O₃. These five air-quality basins could provide the EPA with an alternative assessment of the atmospheric carrying capacity control regions. The counties within the same air-quality basins are proposed to devote their efforts to improve ozone problems.

References

- Alpert, D.J., Hopke, P.K., 1980. A quantitative determination of sources in the Boston urban aerosol. *Atmospheric Environment* 14, 1137–1146.
- Alpert, D.J., Hopke, P.K., 1981. A determination of the sources of airborne particles collected during the regional air pollution study. *Atmospheric Environment* 15, 675–687.
- Ashbaugh, L.L., Myrup, L.O., Flocchini, R.G., 1984. A principal component analysis of sulfur concentrations in the western United States. *Atmospheric Environment* 18, 783–791.
- Buell, C.E., 1975. The topography of empirical orthogonal functions. Preprints of the Fourth Conference on Probability and Statistics in Atmospheric Sciences, pp. 188–193.
- Dorling, S.R., Davies, T.D., Pierce, C.E., 1992. Cluster analysis: a technique for estimating the synoptic meteorological controls on air and precipitation chemistry-method and applications. *Atmospheric Environment* 26A, 2575–2581.
- Eder, B.K., 1989. A principal component analysis of SO₄²⁻ precipitation concentrations over the eastern United States. *Atmospheric Environment* 23, 2739–2750.
- Eder, B.K., Davis, J.M., Bloomfield, P., 1993. A characterization of the spatiotemporal variability of non-urban ozone concentrations over the eastern United States. *Atmospheric Environment* 27A, 2645–2668.
- Fernau, M.E., Samson, P.J., 1990a. Use of cluster analysis to define periods of similar meteorology and precipitation chemistry in Eastern North American. Part I: Transport patterns. *Journal of Applied Meteorology* 29, 735–750.
- Fernau, M.E., Samson, P.J., 1990b. Use of cluster analysis to define periods of similar meteorology and precipitation chemistry in Eastern North American. Part II: Precipitation patterns and pollutant deposition. *Journal of Applied Meteorology* 29, 751–761.

- Flocchini, R.G., Cahill, T.A., Pitchford, M.L., Eldred, R.A., Feeney, P.J., Ashbaugh, L.L., 1981. Characterization of particles in the arid west. *Atmospheric Environment* 15, 2017–2030.
- Gaarenstroom, P.D., Perone, S.P., Moyers, J.L., 1977. Application of pattern recognition and factor analysis for characterization of atmospheric particulate composition in southwest desert atmosphere. *Environmental Science and Technology* 11, 795–800.
- Gatz, D.F., 1978. Identification of aerosol sources in the St. Louis area using factor analysis. *Journal of Applied Meteorology* 17, 600–608.
- Horel, J.D., 1981. A rotated principal component analysis of the interannual variability of the northern hemisphere 500 mb height field. *Monthly Weather Review* 109, 2080–2092.
- Juang, B.J., Liaw, C.M., Sheu, F.C., 1996. A study on the delineation of air basins at Taiwan. *The Proceedings of 13th Air Pollution Control Technologies Conference*, Taipei, Taiwan, pp. 967–974.
- Kaiser, H.F., 1958. The Varimax criterion for analytic rotation in factor analysis. *Psychometrika* 23, 187–201.
- Kalkstein, L.S., Tan, G., Skindlov, J.A., 1987. An evaluation of three clustering procedures for use in synoptic climatological classifications. *Journal of Climate and Applied Meteorology* 26, 717–730.
- Kidson, J.W., 1975. Eigenvector analysis of monthly mean surface data. *Monthly Weather Review* 103, 177–186.
- Liu, S.C., Trainer, M., Fehsenfeld, F.C., Parrish, D.D., Williams, E.J., Fahey, D.W., Hubler, G., Murphy, P.C., 1987. Ozone production in the rural troposphere and the implications for regional and global ozone distribution. *Journal of Geophysical Research* 92, 4191–4207.
- Logan, J.A., 1989. Ozone in rural areas of the United States. *Journal of Geophysical Research* 94, 8511–8532.
- Overland, J.E., Preisendorfer, R.W., 1982. A significance test for principal components applied to a cyclone climatology. *Monthly Weather Review* 119, 1–4.
- Pitchford, A., Pitchford, M., Malm, W., Flocchini, R., Cahill, T., Walther, E., 1981. Regional analysis of factor affecting visual air quality. *Atmospheric Environment* 15, 2043–2054.
- ROC Environmental Protection Administration, 1999. Air Pollution Control Act, <http://www.epa.gov.tw/english/laws>.
- ROC Environmental Protection Administration, 1997. Investigation of emission inventory and evaluation of emission abatement strategies on air pollutants.
- Vukovich, F.M., Fishman, J., 1986. The climatology of summertime O₃ and SO₂ (1977–1981). *Atmospheric Environment* 20, 2423–2433.
- Wallace, C.E., Gutzler, D.G., 1981. Teleconnections in the geopotential height field during the northern hemisphere winter. *Monthly Weather Review* 109, 784–812.
- Walsh, J.E., Mostek, A., 1981. A quantitative analysis of meteorological anomaly over the United States, 1900–1977. *Monthly Weather Review* 108, 615–639.
- Yu, T.Y., Chang, L.F.W., 1999. Division of air basins utilizing principal component analysis. *Journal of the Chinese Institute of Environmental Engineering* 9, 209–218.