

Title	東アジアの酸性・酸化性物質の動態解明へのクラスタリング手法の適用
Author(s)	小山内, 尚
Citation	
Issue Date	2000-03
Type	Thesis or Dissertation
Text version	author
URL	<a href="http://hdl.handle.net/10119/657">http://hdl.handle.net/10119/657</a>
Rights	
Description	Supervisor:ホー ツー バオ, 知識科学研究科, 修士

# U sing C lustering T echn iques to I nvestigate A cidic and O xidative C om ponents in E ast A sia

H isashi O sanai

School of K now ledge Science  
Japan Advanced Institute of Science and Technology  
M arch 2000

K eyw ords: C lustering, K D D , regim e, N I E S , T he E ast A sian .

The E ast A sian area is one of the largest regions for the anthropogenic em ission of  $\text{NO}_x$  and  $\text{SO}_2$  [Rodhe 89]. In fact, because of population grow th and the rapid developm ent of industrial activity, it has been projected that this area w ill becom e the largest source of  $\text{NO}_x$  and  $\text{SO}_2$  in the world in the 21<sup>st</sup> century [G alloway 89]. It is im portant to analyze the present status of air pollution over the E ast A sian region, to evaluate the am ount of the anthropogenic em ission, to accurately predict the future situation by the use of com puter m odels, and to apply these results to w ork out counterm easures.

The book "Studies on the behavior of A cidic and O xidative C om ponents in E ast A sia" which investigated the pollution state of the present atm osphere is offered by N ational Institute for Environm ental Studies (N I E S).

Accumulated huge data currently, exceeds human's information processing capability. Therefore, there is a need of establishment of computer methods to exploit effectively these data. Recently, data mining methods have been expected to be efficient methods to extract knowledge and unknown information from large databases. It is well known that ozone of the troposphere background is an important compound that influences atmospheric oxidation phenomenon. Also, ozone is a very important component from the viewpoint of the global warming as a greenhouse gas.

The increase of tropospheric ozone in the Northern Hemisphere is reported recently in Europe and America, and the study of this increase attracts much attention from researchers. A report for these phenomena in East Asia region is indispensable. It is known that only  $\text{NO}_2$  becomes the outlier object of ozone in the troposphere, and there is an indication that the increase of the anthropogenic emission of  $\text{NO}_x$  also includes  $\text{NO}$  which has brought the increase in such ozone concentration. Thus, it is important to see the increase of ozone and  $\text{NO}_x$  from various angles, to govern chemistry of the tropospheric atmosphere, and to know the three-dimensional concentration distribution. It is an indispensable factor when considering the change of the tropospheric atmosphere. On the one hand,  $\text{SO}_2$  has played an important role as an outlier object of sulfuric acid of cause material of acid-rain.

It is said that a lot of  $\text{SO}_2$  flows in Japan from the part of East Asia area containing China and the Republic of Korea

The purpose of this study is to find new knowledge about behavior of acidic and oxidative component in East Asia using clustering techniques. Especially, we want to detect regimes which have deep influences on climate model and prediction.

This research consists of the following tasks which are required in main-steps of KDD (preparation of data, mining, analysis and understanding of results).

First, as the preparation of data, we do the removal of missing values and transformation of the different NIES databases into one unified database which is convenient for processing. These tasks are done by using Perl and PostgreSQL. Secondly, the data mining process is done on the unified database by using PAM clustering method and "natural" K discovery method. The programs for these algorithms were implemented by C language on UNIX workstation. Finally, the tree-dimensional structure of clusters discovered by the above methods are expressed visually.

Eventually, the following analysis results were obtained:

- (1) The "natural" number of clusters is four.
- (2) The pollution state of the cluster of the Northern part of the Chugoku district is severer than the contamination of other clusters.
- (3) In the cluster of the Southern part of Kyushu district, there is a significant correlation between temperature and sulfur dioxide.

The following conclusions can be drawn from the above-mentioned analysis results:

Each cluster discovered in (1) has no characteristics of the regime defined in this study, therefore it is not a regime. However, these clusters provide us insights to obtain results (2) and (3).

From the result (2), we can see clearly the relation between a pollution state and the transportation distance of an acid-rain contaminant.

From the result (3), we extracted a relationship between temperature and sulfur dioxide, that is the key point of the global cooling system. It can be considered that this is a precious relation that supports the global cooling system observed from actual data.

We hope that these results, especially (2) and (3), can be considered as a significant contribution to the field of environment studies.