

# Analysis of Maritime Accidents using Fuzzy Clustering - Kinds, Causes and Number of Ships

Kazuyuki Tanaka  
Japan Coast Guard  
2-1-3, Kasumigaseki, Chiyoda-ku, Tokyo, Japan  
kazuyuki-tanaka@kaiho.mlit.go.jp

Mika Sato-Ilic  
University of Tsukuba  
1-1-1, Tennodai, Tsukuba-shi, Ibaraki, Japan  
mika@sk.tsukuba.ac.jp

**Abstract**— In this paper, we report the result of an analysis of maritime accidents using fuzzy clustering. The maritime accident data is observed as 3-way data (kinds, causes and number of ships every year). We used multivariate analysis, fuzzy cluster loading, and kernel fuzzy cluster loading to find the feature and trend of this data. As a result of the clustering, we obtained two clusters. In order to find the properties of the obtained clusters, we used fuzzy cluster loading and kernel fuzzy cluster loading.

## I. INTRODUCTION

Over the last 10 years, an average of 2,500 ships a year have met maritime accidents in Japan. Losses caused by maritime accidents, average 250 billion yen and approximately 180 lives a year in Japan. In addition to these facts, the marine environment has been polluted by oil spills and maritime traffic has come to a complete standstill due to maritime accidents. Once a marine accident occurs, human life and property are lost, so it is very important to prevent it.

The Japan Coast Guard (JCG) carries out activities to prevent maritime accidents and keep the sea safe and enjoyable. To ensure safe navigation, JCG enforces various laws and regulations, establishes aids to navigation system and prepares nautical charts. Also the JCG maintains a system that ensures prompt search and rescue operations to provide against maritime accidents. In addition to these things, JCG takes a survey of maritime accidents and analyzes the data.

In this study, we applied quantitative analysis using multivariate analysis for maritime accident data and the feature and similarity of a maritime accident are extracted. Moreover, 3-way data with the passage of time is treated

as data, the fuzzy cluster loading and the kernel fuzzy cluster loading to 3-way data are calculated, and with-time change of the relevance of accident kind and cause is presumed.

## II. DATA

We use maritime accident data based on the JCG's original data. This data has the structure of 3-way data where the 2-way data of the kinds and causes of maritime accident recorded for every year. Period of data is from 1975 to 2002. The data structure is shown in Fig. 1, and the number of maritime accidents is shown in Fig. 2.

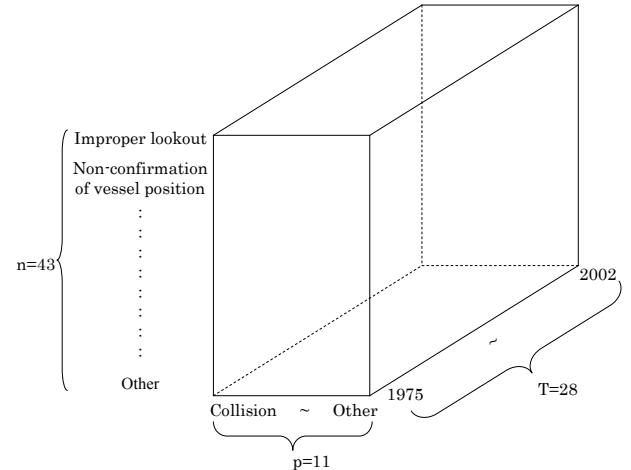


Fig. 1 Data Structure

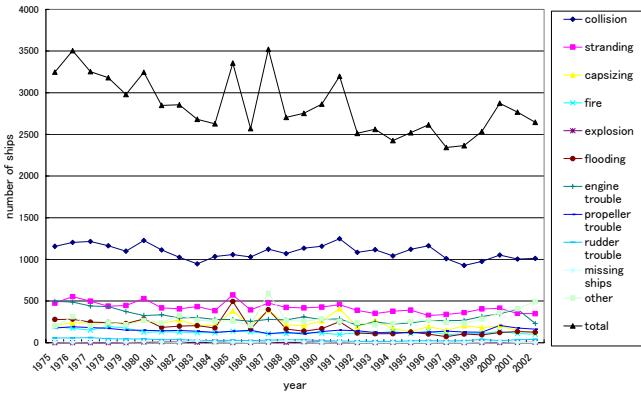


Fig. 2 Number of Maritime Accidents

### III. METHOD OF ANALYSIS

#### A. Fuzzy Cluster Loading

In order to obtain interpretation of fuzzy clustering result, we used the following model of fuzzy cluster loading [5] :

$$u_{ik} = \sum_{a=1}^p x_{ia} z_{ak} + \varepsilon_{ik}, \quad i = 1, \dots, n, \quad k = 1, \dots, K \quad (1)$$

where,

$u_{ik}$  the obtained fuzzy clustering result as the degree of belongingness of an object  $i$  to a cluster  $k$

$$\left( u_{ik} \in [0,1], \sum_{k=1}^K u_{ik} = 1, \sum_{i=1}^n u_{ik} > 0 \right)$$

$x_{ia}$  the value of an object  $i$  to a variable  $a$

$\varepsilon_{ik}$  an error

$z_{ak}$  the fuzzy degree which represents the amount of loading of a cluster  $k$  to a variable  $a$

$n$  number of objects

$p$  number of variables

$K$  number of clusters

We call these  $z_{ak}$  fuzzy cluster loadings. This parameter will show how each cluster can be explained by each of the variables.  $K$  is given in advance.

The model (1) is rewritten as

$$\mathbf{1} = U_k X \mathbf{z}_k + \mathbf{e}_k \quad (2)$$

using

$$U_k = \begin{pmatrix} u_{1k}^{-1} & 0 & \dots & \dots & 0 \\ 0 & u_{2k}^{-1} & 0 & \dots & \vdots \\ \vdots & \vdots & \vdots & \vdots & 0 \\ 0 & \dots & \dots & \dots & u_{nk}^{-1} \end{pmatrix},$$

$$X = \begin{pmatrix} x_{11} & x_{12} & \dots & x_{1p} \\ x_{21} & x_{22} & \dots & x_{2p} \\ \vdots & \vdots & \dots & \vdots \\ x_{n1} & x_{n2} & \dots & x_{np} \end{pmatrix},$$

$$\mathbf{1} = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}, \quad \mathbf{z}_k = \begin{pmatrix} z_{1k} \\ z_{2k} \\ \vdots \\ z_{pk} \end{pmatrix}, \quad \mathbf{e}_k = \begin{pmatrix} e_{1k} \\ e_{2k} \\ \vdots \\ e_{nk} \end{pmatrix}, \quad e_{ik} \equiv u_{ik}^{-1} \varepsilon_{ik}$$

To minimize  $\mathbf{e}_k^T \mathbf{e}_k$ , the estimate of  $\mathbf{z}_k$  is obtained as

$$\mathbf{z}_k = (X^T U_k^2 X)^{-1} X^T U_k \mathbf{1}. \quad (3)$$

From

$$\mathbf{e}_k^T \mathbf{e}_k = (U_k \mathbf{e}_k)^T (U_k \mathbf{e}_k) = \mathbf{e}_k^T U_k^2 \mathbf{e}_k, \quad \mathbf{e}_k = (\varepsilon_{1k}, \dots, \varepsilon_{nk})^T,$$

we estimate  $\mathbf{z}_k$  of (1) to minimize weighted regression error. On the other hand, to minimize  $\mathbf{e}_k^T \mathbf{e}_k$  of (1), the estimate of  $\mathbf{z}_k$  is obtained as

$$\mathbf{z}_k = (X^T X)^{-1} X^T U_k^{-1} \mathbf{1}.$$

This is different from  $\mathbf{z}_k$  of (3). So we find  $\mathbf{z}_k$  of (3) is different from the estimate obtained from weighted least square regression. The details of the difference with related methods are discussed in [6].

#### B. Kernel Fuzzy Cluster Loading

Kernel method originally developed in the context of support vector machines [2], the efficient advantage of which has been widely recognized in many areas. The essence of the kernel method is arbitrary mapping from lower dimension space to higher dimension space. Note that the mapping is an arbitrary mapping, so we do not need to find the mapping; this is called the kernel trick.

Suppose an arbitrary mapping  $\Phi$ :

$$\Phi : R^n \rightarrow F, \quad \mathbf{x}, \mathbf{y} \in R^n, \quad \dim(R^n) < \dim(F)$$

where  $F$  is a higher dimension space than  $R^n$ .

We assume

$$k(\mathbf{x}, \mathbf{y}) = \Phi(\mathbf{x})^T \Phi(\mathbf{y})$$

where  $k$  is the kernel function which is defined in  $R^n$  and  $\mathbf{x}, \mathbf{y} \in R^n$ .

Typical examples of the kernel function are as follows:

$$k(\mathbf{x}, \mathbf{y}) = (\mathbf{x} \cdot \mathbf{y})^d \quad (\text{Polynomial kernel of degree } d) \quad (4)$$

$$k(\mathbf{x}, \mathbf{y}) = \exp\left(-\frac{\|\mathbf{x} - \mathbf{y}\|}{2\sigma^2}\right) \quad (\text{Gaussian kernel}) \quad (5)$$

$$k(\mathbf{x}, \mathbf{y}) = \tanh(\alpha(\mathbf{x} \cdot \mathbf{y}) + \beta) \quad (\text{Sigmoid kernel}) \quad (6)$$

By the introduction of this kernel function, we can analyze the data in  $F$  without finding the mapping  $\Phi$  explicitly. Although we have not obtained the characteristic of the mathematical space of  $F$  yet, many studies that show the classification of data according to mapping from  $R^n$  to  $F$  are reported. [3], [7]

From (3), we can obtain the following:

$$\begin{aligned} z_k &= (X' U_k^{-2} X)^{-1} X' U_k \mathbf{1} = ((U_k X)' (U_k X))^{-1} (U_k X)' \mathbf{1} \\ &= (C_k' C_k)^{-1} C_k' \mathbf{1} \end{aligned} \quad (7)$$

where

$$\begin{aligned} C_k &= (c_{ia(k)}), \quad i = 1, \dots, n, \quad a = 1, \dots, p, \\ c_{ia(k)} &\equiv u_{ik}^{-1} x_{ia}, \quad i = 1, \dots, n, \quad a = 1, \dots, p. \end{aligned}$$

Using  $C_{a(k)}' = (c_{1a(k)}, \dots, c_{na(k)})$ , we can represent (7) as follows:

$$z_k = (C_{a(k)}' C_{b(k)})^{-1} (C_{a(k)}' \mathbf{1}), \quad a, b = 1, \dots, p, \quad (8)$$

where  $(C_{a(k)}' C_{b(k)})$  is the  $p \times p$  matrix that have  $C_{a(k)}' C_{b(k)}$  as  $(a, b)$ -th element, and  $(C_{a(k)}' \mathbf{1})$  is the  $p \times 1$  vector that have  $C_{a(k)}' \mathbf{1}$  as  $a$ -th element. That is,

$$\begin{aligned} C_k' C_k &= (C_{a(k)}' C_{b(k)}), \quad a, b = 1, \dots, p \\ C_k' \mathbf{1} &= (C_{a(k)}' \mathbf{1}), \quad a = 1, \dots, p. \end{aligned}$$

From (7) and (8), the fuzzy cluster loading in  $F$  is defined as follows:

$$\tilde{z}_k = (\Phi(C_{a(k)}) \Phi(C_{b(k)}))^{-1} (\Phi(C_{a(k)}) \Phi(\mathbf{1})) , \quad a, b = 1, \dots, p \quad (9)$$

where  $\tilde{z}_k$  shows the fuzzy cluster loading in  $F$ .

Using the kernel method, we can estimate  $\tilde{z}_k$  without

finding  $\Phi$  as follows:

$$\tilde{z}_k = (k(C_{a(k)}, C_{b(k)}))^{-1} (k(C_{a(k)}, \mathbf{1})), \quad a, b = 1, \dots, p \quad (10)$$

where  $k$  is kernel function.

## IV. ANALYSIS BASED ON 2-WAY DATA

### A. Clustering

The result of clustering using complete linkage method is shown in Fig. 3.

From this result we can see that we can divide this data between 1980 and 1981 except 1985 and 1987.

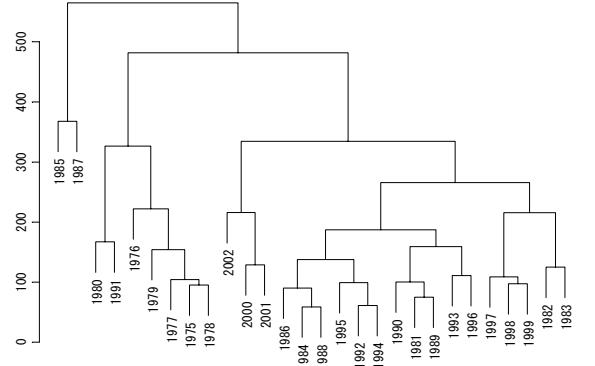


Fig. 3 Clustering Result

### B. Calculation of Fuzzy Cluster Loading

Using the pre-information of the former section, we applied the fuzzy clustering algorithm named FANNY.[1], [4].

The objective function is defined as follows:

$$J(U) = \sum_{k=1}^K \left( \sum_{i=1}^n \sum_{j=1}^n (u_{ik})^m (u_{jk})^m \delta_{ij} / \left( 2 \sum_{l=1}^n (u_{lk})^m \right) \right) \quad (10)$$

where,

$u_{ik}$  the obtained fuzzy clustering result as the degree of belongingness of an object  $i$  to a cluster  $k$

$$\left( u_{ik} \in [0,1], \quad \sum_{k=1}^K u_{ik} = 1, \quad \sum_{i=1}^n u_{ik} > 0 \right)$$

$n$  number of objects

$K$  number of clusters

$\delta_{ij}$  is defined as follows;

$$\delta_{ij} = \sum_{h=1}^p (x_{ih} - x_{jh})^2 \quad (i, j = 1, \dots, n, \quad h = 1, \dots, p)$$

where,

$p$  number of variables

$x_{ih}$  the value of an object  $i$  to a variable  $h$

$m$  given parameter which controls belongingness.

$$(1 < m < \infty)$$

The purpose of FANNY is to estimate  $U = (u_{ik})$  in order to minimize the objective function (10).

We put  $m = 2$  and  $K = 2$ . The result of fuzzy clustering is shown in table 1. In this table, each value shows degree of belongingness to each cluster.

From this result we can see that years 1975 to 1981 belong to Cluster 1 and years 1982 to 2002 belong to Cluster 2 except to 1985.

Table 1 Fuzzy Clustering Result ( $m = 2$ )

year	Cluster 1	Cluster 2
1975	0.55	0.45
:	:	:
1981	0.52	0.48
1982	0.48	0.52
:	:	:
1985	0.53	0.47
:	:	:
1987	0.43	0.57
:	:	:
2002	0.48	0.52

The result of fuzzy cluster loading is shown in Fig. 4.

From this result we can see that Cluster 1 is related to explosion and rudder trouble and Cluster 2 is related to stranding, explosion, rudder trouble and missing ships.

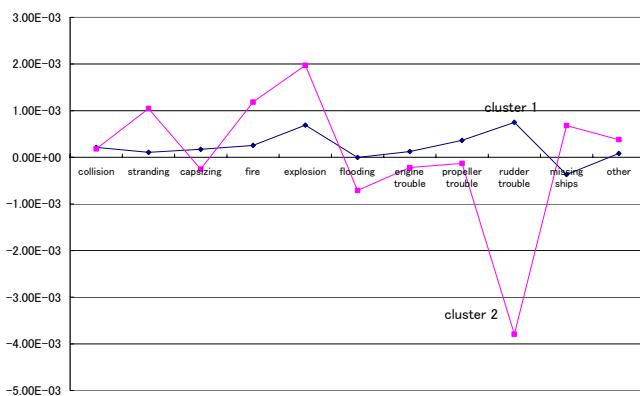


Fig. 4 Fuzzy Cluster Loading Result

### C. Calculation of Kernel Fuzzy Cluster Loading

We calculated kernel fuzzy cluster loading using same fuzzy clustering result of Table 1. We use the polynomial kernel of degree 2 as the kernel function. The result of kernel fuzzy cluster loading is shown in Fig. 5.

From this result we can see that both Cluster 1 and Cluster 2 are related to explosion, rudder trouble and missing ships but the relation of Cluster 2 is larger than that of Cluster 1.

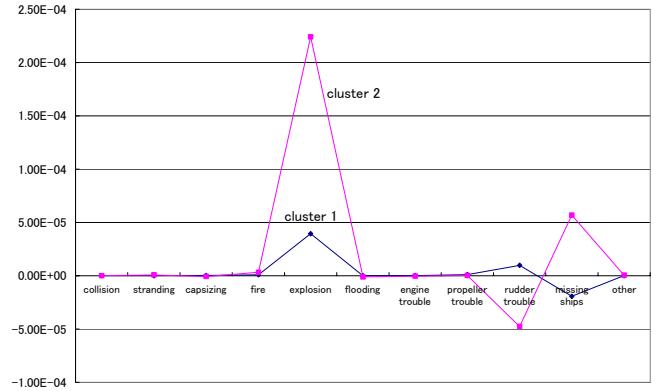


Fig. 5 Kernel Fuzzy Cluster Loading Result

From comparing these results we can consider that these clusters are characterized by three causes (explosion, rudder trouble and missing ships) for the reason of the kernel method characteristic that is noise elimination.

## V. ANALYSIS BASED ON 3-WAY DATA

### A. The Method of calculate Fuzzy Cluster Loading/Kernel Fuzzy Cluster Loading based on 3-way data

If fuzzy clustering is performed for every year with 3-way data, it does not restrict the agreement of each year's fuzzy cluster. Therefore, clustering result of each year cannot be compared. So we made a new super-matrix by arranging the 2-way data procession of each year perpendicularly, then we can get the degree of belongingness under the same fuzzy cluster. In other words, we can obtain the change of the degree of belongingness with time under an eternal cluster.

Using this degree of belongingness, we can obtain the fuzzy cluster loading/kernel fuzzy cluster loading under the cluster the same as the degree of belongingness.

So we can compare the degree of attribution and the fuzzy cluster loading through a time.

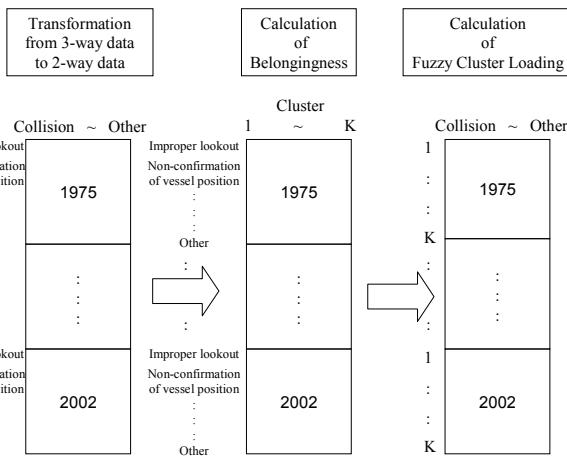


Fig. 6 Method of calculate Fuzzy Cluster Loading/Kernel Fuzzy Cluster Loading based on 3-way Data

### B. Calculation of Fuzzy Cluster Loading based on 3-way data

We calculated fuzzy cluster loading based on 3-way data using the fuzzy clustering result by using FANNY algorithm [4], where  $m = 2$  and  $K = 2$ . The result of fuzzy cluster loading using 3-way data is shown in Fig. 7.

From this result we can see that both Cluster 1 and Cluster 2 are related with explosion, rudder trouble and missing ships but the other causes stay unchanged. And we can see that fuzzy cluster loadings stay unchanged before 1980 but these have sharp fluctuations after 1981.

### C. Calculation of Kernel Fuzzy Cluster Loading based on 3-way data

We calculated kernel fuzzy cluster loading based on 3-way data using the same fuzzy clustering result as above. We use the polynomial kernel of degree 2 as the kernel function. The result of kernel fuzzy cluster loading

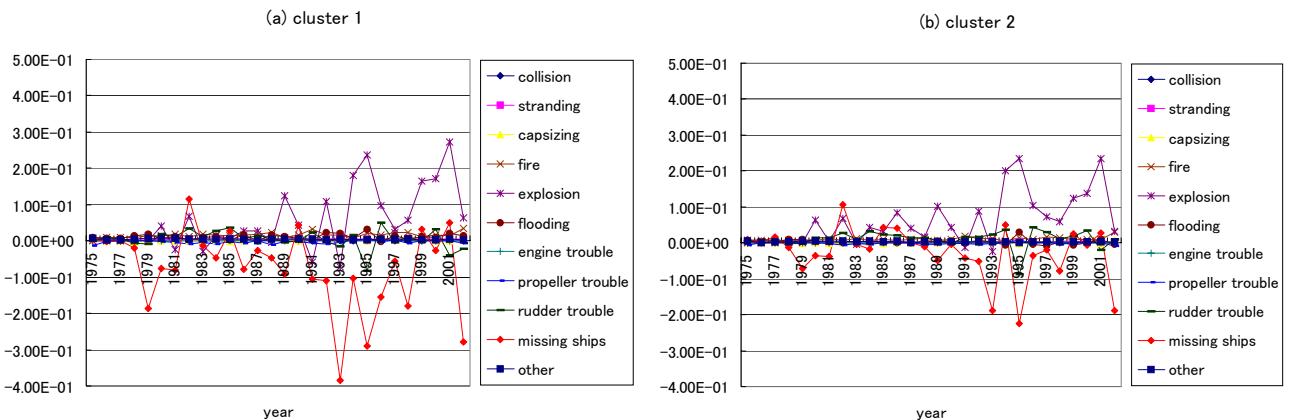


Fig. 7 Fuzzy Cluster Loading using 3-way Data Result

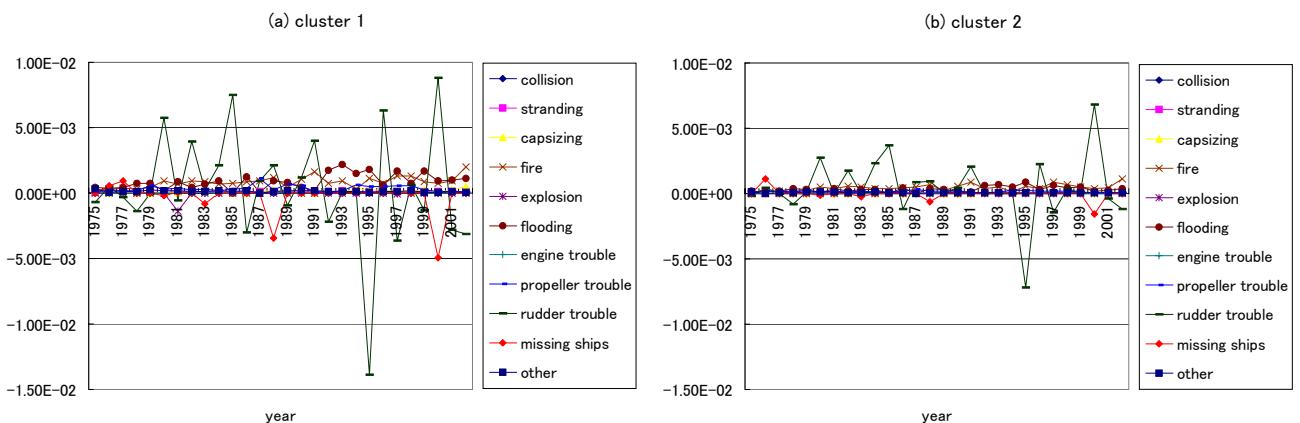


Fig. 8 Kernel Fuzzy Cluster Loading using 3-way Data Result

using 3-way data is shown in Fig. 8.

From this result we can see that Cluster 1 is related with explosion, rudder trouble and missing ships, and Cluster 2 is related with rudder trouble. We can see that fuzzy cluster loadings stay unchanged before 1980 but these have sharp fluctuations after 1981 similar to fuzzy cluster loading using 3-way data.

## VI. CONCLUSION

In this paper, we made an analysis of maritime accidents using fuzzy clustering. We used cluster analysis, fuzzy cluster loading, and kernel fuzzy cluster loading to find the feature and trend of this data.

From the result of clustering, we can see that this data include 2 clusters, one is the term before 1980 and the other is the term after 1981 excluding 1985 and 1987. The reason is that we can consider these two years are outlier from Fig.2.

In order to find the properties of the obtained clusters, we used fuzzy cluster loading and kernel fuzzy cluster loading. From the result of fuzzy cluster loading, we can see that Cluster 1 is related with explosion and rudder trouble and Cluster 2 is related with stranding, explosion, rudder trouble and missing ships. And from the result of kernel fuzzy cluster loading, we can see that both Cluster 1 and Cluster 2 are related with explosion, rudder trouble and missing ships but the relation of Cluster 2 is larger than that of Cluster 1. From comparing these results we can consider that these clusters are characterized by three causes (explosion, rudder trouble and missing ships) for the reason of the kernel method characteristic that is noise elimination.

We have maritime accident data as 3-way data, so we tried to apply fuzzy cluster loading and kernel fuzzy cluster loading to this 3-way data.

From the result of fuzzy cluster loading using 3-way data, we can see that both Cluster 1 and Cluster 2 are related with explosion, rudder trouble and missing ships but the other causes stay unchanged. And we can see that fuzzy cluster loadings stay unchanged before 1980 but these have sharp fluctuations after 1981. From the result of kernel fuzzy cluster loading using 3-way data, we can see that Cluster 1 is related with explosion, rudder trouble and missing ships and Cluster 2 is related with rudder trouble. And we can see that fuzzy cluster loadings stay unchanged before 1980 but these have sharp fluctuations after 1981 the same as fuzzy cluster loading using 3-way data. It is thought that the reasons can be classified into

two clusters also from the amount of change of this fuzzy cluster loading and the kernel fuzzy cluster loading can be explained.

## REFERENCE

- [1] J.C. Bezdek, *Pattern Recognition with Fuzzy Objective Function Algorithms*, Plenum Press (1987)
- [2] N. Cristianini and J. Shawe - Taylor, *An Introduction to Support Vector Machines and Other Kernel-Based Learning Methods*, Cambridge University Press (2000)
- [3] R. Inokuchi and S. Miyamoto, LVQ Clustering and SOM using a Kernel Function, *20th Fuzzy System Symposium*, pp. 65-68, (2004)
- [4] L. Kaufman, P. J. Rousseeuw, *Finding Groups in Data*, Wiley (1990)
- [5] M. Sato-Ilic, On Kernel based Fuzzy Cluster Loadings with the Interpretation of the Fuzzy Clustering Result, *International Journal of Computational and Numerical Analysis and Applications*, Vol. 4, No. 3, pp. 265-278, (2003)
- [6] M. Sato-Ilic, On Fuzzy Clustering based Regression Models, *North American Fuzzy Information Society*, pp. 216-221, (2004)
- [7] B. Schölkopf, S. Mika, C. J. C. Burges, P. Knirsch, K. Müller, G. Rätsch, and A. J. Smola, Input Space Versus Feature Space in Kernel-Based Methods, *IEEE Transactions on Neural Networks*, Vol. 10, No. 5, (1999)