Epidemiologic Analyses of Spatial Clustering of Bovine Ephemeral Fever Outbreaks II. Principal Component Analysis

Takashi OGAWA, Kazuki ISHIBASHI[1)], Kazuhiko IMAMURA[2)], Sei KURASHIGE[3)], and Tadahiro INOUE

National Institute of Animal Health, 3–1–1 Kannondai, Tsukuba, Ibaraki 305, [1)]Chuo Livestock Hygiene Service Center, 2–1–3 Isouda, Hakata, Fukuoka 816, [2)]Department of Livestock Industry, Fukuoka Prefecture, 7–7 Higashi-Koen, Hakata, Fukuoka 812, and [3)]Ryochiku Livestock Hygiene Service Center, 424 Zendoji, Iida, Kurume, Fukuoka 839–11, Japan

(Received 11 March 1993/Accepted 22 October 1993)

ABSTRACT. The principal component analysis (PCA) was applied to analyze a correlation matrix of three variables on epidemic data of bovine ephemeral fever (BEF) outbreaks. These original data were summarized from the official outbreak report of Fukuoka Prefecture. The first and the second principal components of the PCA were interpreted as the infectious potency due to BEF virus and the prevention against BEF virus infection, respectively. The BEF outbreak areas were able to be classified epidemically into 4 groups by using the two principal components. The valuable epidemiological insights can be reasonably obtained from an application of the PCA. The results provided an important information for a further BEF vaccination campaign in the western part of Japan.—KEY WORDS: bovine ephemeral fever, epidemiology, principal component analysis, spatial clustering.

J. Vet. Med. Sci. 56(2): 223–226, 1994

The incidence patterns of farms affected with bovine ephemeral fever (BEF) in Fukuoka Prefecture in October of 1988 [5] have revealed a spatial clustering explained well by the negative binomial distribution [6]. The result also suggested that the clustering might be influenced by some multivariate factors (risk factors).

On the other hand, the principal component analysis (PCA) has been used as a method of multivariate analysis [1, 2, 4, 7]. The original variables (risk factors) of data are transformed into new, uncorrelated variables in the PCA and these transformed variables are called principal components. The principal components are results of reducing the number of the original variables without losing much of the information [3]. Also, the PCA has been applied for selecting a heterogeneous group from a population by a systematical judgment [7].

In the present paper we have attempted to explain epidemiologically the spatial clustering of the BEF outbreaks with a systematical judgment by using an application of the PCA and the results will be discussed in detail.

MATERIALS AND METHODS

Preparation of epidemic data for PCA: The PCA was carried out by using the same data which were used previously for an analyzing of the negative binomial distribution [6]. The original data of BEF outbreaks in October of 1988 were summarized from the official outbreak report of the Department of Livestock Industry in Fukuoka Prefecture (unpublished data). The data, especially on 22 municipal areas of the BEF outbreaks in Fukuoka Prefecture, were processed and summarized to the epidemic data of 11 municipals as follows: (1) the incidence rate of cattle per 10,000 head; (2) the incidence rate of farm (%); and (3) the vaccination rate of cattle (%).

The two incidence rates and vaccination rate of the BEF were calculated by the method described in the previous work [5] as follows:

$$\text{Incidence rate of cattle} = \frac{\text{Incidence of cattle}}{\text{Cattle population at risk}} \times 10{,}000$$

$$\text{Incidence rate of farm} = \frac{\text{Incidence of farm}}{\text{Farm population at risk}} \times 100$$

$$\text{Vaccination rate of cattle} = \frac{\text{Number of cattle vaccinated}}{\text{Number of dairy and beef cattle}} \times 100$$

The data regarding to the vaccination were also obtained from the annual report of the Department of Livestock Industry in Fukuoka Prefecture (unpublished data). The vaccinations were carried out during the period of May-July of 1988 on a voluntary basis under the direction of Fukuoka Prefecture.

The data of cattle and farm population at risk in October and during the period of May-July of 1988 were referred from the annual statistics of livestock dated February 1, 1988, which was published by the Statistics and Information Department, Ministry of Agricultures, Forestry and Fisheries (MAFF) of Japan [10].

Calculation and Mathematics of PCA [1, 2, 4, 7]: The PCA extracts new axes of variables (principal components) from original variables and summarizes the original information into new characteristics.

At first the PCA was calculated from the correlation matrix of three variables by using the unit-length standardization method so that each variable has a mean equal to zero and a variance equal to one. The PCA was calculated as follows:

Let R be the correlation matrix of the original variables and X be the column vector of the standardized variables:

$$R=\begin{pmatrix} 1 & r_{12} & r_{13} \\ r_{12} & 1 & r_{23} \\ r_{13} & r_{23} & 1 \end{pmatrix}, \quad X=\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}.$$

Let l be the column vector such that $l^T l = 1$, i.e.,

$$l = (l_1 \quad l_2 \quad l_3)^T,$$

where $l_1^2+l_2^2+l_3^2 = 1$. To obtain the linear combination $z = l_1x_1 + l_2x_2 + l_3x_3$ that gives the maximum variance, we must find the solution of the following equation:

$$|R - \lambda I| = 0,$$

where I is a unit matrix; λ is a Lagrange multiplier. This problem is equivalent to the problem to find eigenvalues of R. In the case of three variables, we obtain three eigenvalues $\lambda_1, \lambda_2, \lambda_3$ ($\lambda_1 > \lambda_2 > \lambda_3$) and the corresponding eigenvectors:

$$l_i = (l_{i1} \quad l_{i2} \quad l_{i3})^T, \qquad i=1,2,3.$$

Then, the ith principal component, z_i, is given by the linear combination of the ith eigenvector and the standardized variables:

$$z_i = l_{i1}x_1 + l_{i2}x_2 + l_{i3}x_3, \qquad i=1,2,3.$$

The ith eigenvalue divided by the sum of all the eigenvalues, $\lambda_i/(\lambda_1 + \lambda_2 + \lambda_3)$, represents the proportion of total variation explained by the ith principal component.

The calculation of the PCA was processed according to the Statistical Analysis System (SAS) [9] by using the HP 9000/720 work station in the Tsukuba Computer Center for Agriculture, Forestry and Fisheries Research of the MAFF.

RESULTS

Input data for PCA: The epidemic data (three variables × 11 municipal areas) are shown in Table 1. The rates distribution of the three variables tended to disperse respectively. The correlation matrix calculated from the three variables of the epidemic data is also shown in Table 2. The correlation coefficient between variables of the two incidence rates (variable name: CATTLE and FARM) was high while the others were low. The clustering of BEF outbreaks was not well explained with the vaccination control.

Results of PCA: The eigenvalues of the correlation matrix are shown in Table 3-a. The eigenvalue of the first principal component (PRIN1) reflecting the largest amount of the variation was 1.882 that was 62.75 percent of the total variation. The cumulative proportion of the first two eigenvalues was 0.9386 reflecting a major variation observed among the three variables. Table 3-b shows eigenvectors of correlation matrix. The incidence rate of cattle (CATTLE) contributed 0.6541 to the PRIN1 and 0.6905 to farm (FARM). On the other hand, the vaccination rate (VACCINE) contributed only 0.9341 to the PRIN2.

The results showed that the PRIN1 could be interpreted as the factor relating to the incidence rate, namely, infectious potency due to BEF virus. Furthermore, the

Table 1. Epidemic data with 3 variables of bovine ephemeral fever outbreaks in 11 municipal areas of Fukuoka Prefecture in October 1988 for principal component analysis

Municipal area	Incidence rate of BEF[a] in: cattle per 10,000	Incidence rate of BEF[a] in: farm (%)	Vaccination rate against BEF (%)
1) Fukuoka	89.18	16.07	14.80
2) Chikushi	52.02	10.0	2.89
3) Munakata	26.37	7.69	21.24
4) Kasuya	34.99	18.18	0.0
5) Itoshima	71.61	20.0	17.48
6) Kita-Kyusyu	15.48	4.0	9.80
7) Onga	73.26	12.50	31.14
8) Iizuka	6.73	5.56	4.10
9) Kurate	33.39	7.14	57.10
10) Kaho	25.40	4.30	1.02
11) Kurume	2.49	0.90	55.07
Mean	39.17	9.67	19.51

a) BEF: Bovine ephemeral fever.

Table 2. Correlation matrix among 3 variables for principal component analysis

Variable	CATTLE[a]	FARM[b]	VACCINE[c]
CATTLE	1.0	0.7906	−0.1016
FARM	*	1.0	−0.2985
VACCINE	*	*	1.0

a) CATTLE: Incidence rate of bovine ephemeral fever per 10,000 head of cattle.
b) FARM: Incidence rate of bovine ephemeral fever per cent (%) of farm.
c) VACCINE: Vaccination rate against bovine ephemeral fever per cent (%) of cattle.

Table 3. Eigenvalues and eigenvectors of correlation matrix
a. Eigenvalues

	Eigenvalue	Difference	Proportion	Cumulative
PRIN1[a]	1.8825	0.9491	0.6275	0.6275
PRIN2[b]	0.9334	0.7492	0.3111	0.9386
PRIN3[c]	0.1841	*	0.0614	1.0
	3.0	*	1.0	*

b. Eigenvectors

Variable	PRIN1	PRIN2	PRIN3
CATTLE	0.6541	0.3451	−0.6731
FARM	0.6905	0.0910	0.7176
VACCINE	−0.3089	0.9342	0.1787

a) PRIN1: First principal component.
b) PRIN2: Second principal component.
c) PRIN3: Third principal component.

PRIN2 could be interpreted as the factor relating to the vaccination rate, namely, prevention against BEF virus infection.

The component scores among the 11 municipal areas which induced from the PRIN1 and PRIN2 are presented in Table 4. The 11 areas could be classified into 4 groups (I, II, III, and IV) by two dimensions of the axes of the PRIN1 and PRIN2.

The epidemiological characteristics of the 4 groups could be explained as follows:

(Group I); the three municipal areas scored PRIN1>0 and PRIN2>0 were located along the coastline and were affected considerably with the BEF irrespective of a constant vaccination control.

(Group II); the two areas scored PRIN1>0 and PRIN2<0 were located in the areas adjacent to the Group I and were affected certainly with the BEF because the vaccination rates were low.

(Group III); the four areas scored PRIN1<0 and PRIN2<0 were located in the inland except a part of Munakata and Kita-Kyusyu. These areas were less affected with the BEF irrespective of low vaccination rates.

(Group IV); the two areas scored PRIN1<0 and PRIN2>0 were also located in the inland. These areas revealed less the BEF and high vaccination rates.

Figure 1 shows the geographical location of the 4 groups in the map of Fukuoka Prefecture.

Table 4. Component scores of the epidemic data with 3 variables of bovine ephemeral fever outbreaks in 11 municipal areas of Fukuoka Prefecture in October 1988

Municipal area	PRIN1	PRIN2	Group
1) Fukuoka	1.92	0.48	I
5) Itoshima	1.91	0.45	I
7) Onga	0.91	0.98	I
4) Kasuya	1.14	−0.82	II
2) Chikushi	0.58	−0.60	II
3) Munakata	−0.54	−0.10	III
10) Kaho	−0.63	−1.09	III
8) Iizuka	−0.96	−1.15	III
6) Kita-Kyusyu	−1.02	−0.81	III
9) Kurate	−0.98	1.61	IV
11) Kurume	−2.34	1.06	IV

DISCUSSION

The spatial clustering of BEF outbreaks could be classified epidemically into 4 groups by using the two principal components of the PCA. The two components were independent (orthogonal) and could be interpreted as the BEF infection and prevention factors, respectively. Finally, the geographical locations of the 4 groups were shown as the epidemic zones onto the map of Fukuoka Prefecture (Fig. 1). For example, Fukuoka, Itoshima and Onga where the areas located along the coastline (group I) revealed high incidence of the BEF regardless of a constant vaccination rate (30 to 40%).

In the previous papers with the meteorological analysis [5, 8], it has been suggested that the low-level jet stream

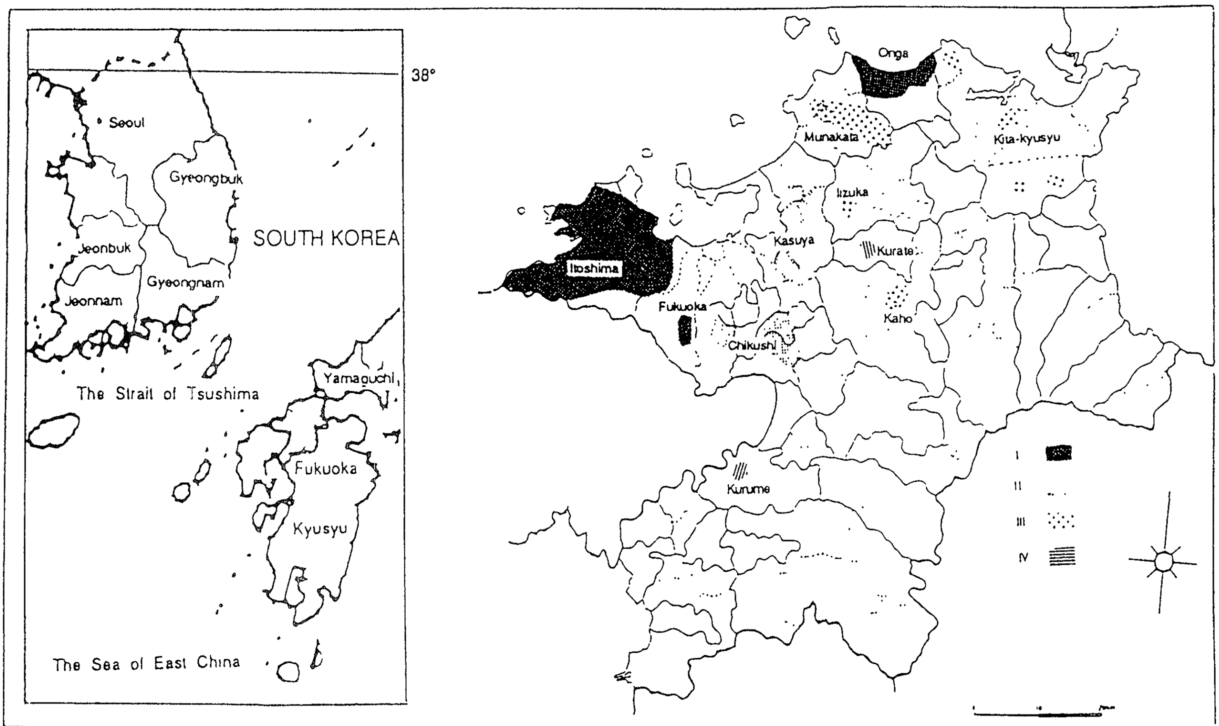

Fig. 1. Geographical location in Fukuoka Prefecture occurred bovine ephemeral fever in 1988. The outbreak areas could be epidemically classified into 4 groups by the principal component analysis.

from China or South Korea had an impact on the area located in the coastline along the Strait of Tsushima (northern part of Kyusyu Island of Japan). Next, it has been speculated epidemiologically that the BEF virus was carried across the Strait on vectors themselves carried on the jet stream to Fukuoka Prefecture [5, 8]. So it could be easily understood that the reason why the high incidence of the BEF appeared in these areas. In contrast, the inland areas of Fukuoka Prefecture had less outbreaks of BEF irrespective of a low vaccination rate because the areas were not affected directly with such stream or vectors.

The results of the PCA presented valuable epidemiological insights on the spatial clustering of BEF outbreaks. By judging the results of the PCA, it would allow to make a systematical judgment of information for a further BEF vaccination campaign in the western part of Japan.

ACKNOWLEDGEMENT. The authors would like to express gratitude to Dr. K. Yamamura, National Agro-Environment Institute of Tsukuba for his valuable advice in preparing the manuscript.

REFERENCES

1. Anderson, T. W. 1984. An Introduction to Multivariate Statistic Analysis, 2nd ed., Wily, New York.
2. Kendall, M. 1980. Multivariate Analysis, 2nd ed., Griffin, High Wycombe.
3. Lafi, S. Q. and Kaneene, J. B. 1992. An explanation of the use of principal-components analysis to detect and correct for multicollinearity. *Prev. Vet. Med.* 13: 261–275.
4. Mardia, K. V., Kent, J. T., and Bibby, J. M. 1979. Multivariate Analysis. Academic Press, London.
5. Ogawa, T. 1992. Epidemiologic investigation of bovine ephemeral fever outbreaks in Kyusyu Island in Japan during the fall of 1988. *Prev. Vet. Med.* 14: 69–76.
6. Ogawa, T., Ishibashi, K., Imamura, K., Kurashige, S., and Hatakeyama, H. 1993. Epidemiologic analyses of spatial clustering of bovine ephemeral fever outbreaks. I. Application of the negative binomial distribution. *J. Vet. Med. Sci.* 55: 177–179.
7. Okuno, C., Kume, H., Haga, T., and Yoshizawa, T. 1985. Methods of Multivariate Analyses, 2nd ed., Nikkagiren, Tokyo (in Japanese).
8. Shirakawa, H., Ishibashi, K., and Ogawa, T. 1993. A comparison of the epidemiology of bovine ephemeral fever in South Korea and south-western Japan. *Aust. Vet. J.* 70 (in press).
9. Statistical Analysis Systems Institute. 1988. SAS/STAT User's Guide. Release 6.03 Edition, Cary, NC.
10. Statistics of Livestock, Japan. 1989. Statistics and Information Department, Ministry of Agriculture, Forestry and Fisheries, Tokyo.