Acta Med. Nagasaki 46 : 33－37

Study on the Relationship of Indoor Air Quality Indexes by Correspondence Analysis and Homogeneity Analysis

Baojun PAN[1)], Feng XU[1)], Zhiren WU[2)], Kiyoshi AOYAGI[4)], Hiroshi SAITO[3)], Taiichiro TAKEMOTO[4)]

1) Fujian Centers for Disease Control, Fuzhow, China
2) Okusa Clinic, Nagasaki Yurino Hospital, Nagasaki
3) Department of Preventive Medicine and Health Promotion, Nagasaki University School of Medicine
4) Department of Public Health, Nagasaki University School of Medicine

We investigated the relationships among air quality indexes, such as the total number of bacteria "BAC", concentration of carbon dioxide "CO2", relative humidity "HUM", concentration of negative ion "NI" and opening time "TIM" in the environment of indoor public place. The data were divided into 3 levels based on its percentile and analyzed by correspondence analysis (ANACOR) and homogeneity analysis (HOMALS) in SPSS 8.0/10.0. Using these methods, we analyzed among 2 variables and their 3 levels, there were the relationships between BAC and CO2, CO2 and TIM, and BAC and TIM. When analyzing by HOMALS, among 3 variables and their 3 levels, there were some similar relationships: BAC, CO2 and TIM; BAC, CO2 and NI. The increase of CO2 and BAC were changed with TIM. The higher CO2 tended to increase BAC and to decrease NI. These results indicated under general conditions, CO2 can show the status of the air BAC. Increasing the ventilation and the relative humidity suitably may raise NI concentration and improve the air quality.

Although the analytical results of the air quality by ANACOR and HOMALS are similar to that of chi-square test, ANACOR and HOMALS can be clearly shown in the plots with 2 or 3 axes. These two methods (ANACOR and HOMALS) may be useful to show the relationships among indoor air quality indexes more sensitively and directly.

ACTA MEDICA NAGASAKIENSIA 46 : 33－37, 2001

Key Words: indoor public places, air quality indexes, correspondence analysis (ANACOR), homogeneity analysis (HOMALS)

Address Correspondence: Zhiren Wu, M.D.
Okusa Clinic, Nagasaki Yurino Hospital No.302 Maeda Bldg. Namesi 1-29-3, Nagasaki 852-8061 Japan
TEL: +81-95-856-5798 FAX: +81-95-856-6663
E-mail: hitoshi-go@h8.dion.jp

Introduction

For analyzing the qualitative or quantitative data of indoor air, t-test, ANOVA, correlation, regression and Chi-square test etc. have been used. The authors introduce two new methods for processing the data: correspondence analysis (ANACOR) and homogeneity analysis (HOMALS) in Statistical Package for the Social Science (SPSS) 8.0/10.0 for Windows. These analyses can show the relationships among variables visually in the plots with two or three dimensions. To study the relationships among air quality indexes, such as the total number of bacteria "BAC", concentration of carbon dioxide "CO2", relative humidity "HUM", concentration of negative ion "NI" and opening time "TIM", we evaluated the environment of indoor place (Karaoke, dancing halls etc, with central air conditioning) using ANACOR and HOMALS.

Data and Methods

Statistical methods

ANACOR analyzes the correspondence relationship between row and column variables (2 dimensions). Five normalization methods can be used: CANONICAL, PRINCIPAL, RPRINCIPAL, CPRINCIPAL, value -2 to +2. The running output contents are: the tables (TABLES) (Table 1), the row profiles, the column profiles (PROFILE), inertia, the singular value (SINGULAR)/ Eigen values (EIGEN) of each dimensions, proportion explained, cumulative proportion, row scores, column scores (SCORES), contribution of row points of each dimension (CONTRIBUTIONS) and 5 plots (TRROWS, TRCOLUMNS. ROWS, COLUMNS and JOINT). The most important plot is JOINT (Row and Column Scores Chart). The distances (Euclidean distances) between categories of the row variable and between

Table 1. The contingency tables and its Chi-square values for 5 air indexes (variables)

Air index (Var.)	LB	MB	HB	x^2 test
LC	17	28	9	x^2=27.035
MC	35	79	26	
HC	8	24	32	P=0.000
	LC	MC	HC	
T0	31	38	11	x^2=28.253
T1	15	52	21	
T2	8	50	32	P=0.000
	LB	MB	HB	
T0	29	36	15	x^2=13.436
T1	18	48	22	
T2	13	47	30	P=0.009
	LN	MN	HN	
LH	27	4	1	x^2=13.078
MH	19	8	2	
HH	1	2	2	P=0.011

Table 2. The main commands used in ANACOR and HOMALS for 5 variables

```
GET FILE = 'C:\SPSS\SAV\Air.sav'.
VALUE LABELS BAC 1 'LB' 2 'MB' 3 'HB'
             /CO2 1 'LC' 2 'MC' 3 'HC'
             /HUM 1 'LH' 2 'MH' 3 'HH'
             /NI  1 'LN' 2 'MN' 3 'HN'
             /TIM 1 'T0' 2 'T1' 3 'T2'.
ANACOR TABLE = CO2(1,3) BY BAC(1,3)
       /DIMENSION = 2 /NORMALIZATION = CANONICAL
       /PRINT = TABLE SCORES PROFILE CONTRIBUTIONS
       /PLOT = JOINT.
HOMALS VARIABLES = BAC(3) CO2(3) TIM(3)
       /DIMENSION = 2 /PRINT = FREQ EIGEN DISCRIM QUANT
       /PLOT = QUANT.
```

categories of the column variable are approximations of Chi-square distances (Table 2).

HOMALS analyzes the homogeneity relationship among row and column variables (2 or 3 dimensions), plotting charts with 2 or 3 dimensions. The most important chart is QUANT (Category Quantification) (Table 2).

Data source and processing:

Data on the air index was collected in Fuzhou, China (1989-1998, 258 points), and a database "Air.sav" was made, using SPSS.

Air index classification:

We measured the total number of bacteria (BAC), concentration of carbon dioxide (CO2), relative humidity (HUM), concentration of negative ion (NI) and opening time (TIM). Based on the percentile (P) of air index value (<P_{25}, P_{25}-P_{75} and >=P_{75}), the values were divided into 3 levels (low (L), middle (M) or high (H)). BAC divided into low bacteria (LB), middle bacteria (MB) and high bacteria (HB), CO2 into low carbon dioxide (LC), middle carbon oxide (MC) and high carbon oxide (HC), HUM into low humidity (LH), middle humidity (MH) and high humidity (HH), NI into low negative ion (LN), middle negative ion (MN) and high negative ion (HN), TIM into T0 (before opening), T1 (1 hours later) and T2 (2 hours later) (Table 1).

The main commands used:

When using ANACOR, DIMENSION=2; HOMALS, DIMENSION=2 or 3 (Table 2). This example uses CANONICAL (default) for normalization.

The main running results

The results by ANACOR:

The main running results are shown in Table 3-5, and the most important outputs are Fig 1-4 (The JOINT plot is introduced only in this paper). The values of the coordinate points (LC, MC, HC, and LB, MB, and HB) in Fig 1 (the same as in Fig 2-4) can find from Table 4-5. For example:

LC (-0.408, 0.419);
LB (-0.484, 0.372);
HB (0.942, 0.062), and so on.

Table 3. 2 dimensions, Inertia and Proportion Explained

Dimension	Singular Value	Inertia	Proportion Explained	Cumulative proportion
1	0.31916	0.10187	0.972	0.972
2	0.05403	0.00292	0.028	1.000
3	—	0.00145	1.000	1.000

Table 4. Row Scores and Contribution of Row points to each dimension

Carbon dioxide (%)	Marginal Profile	Row Scores		Contribution	
		Dim 1	Dim 2	Dim 1	Dim 2
LC (<0.05)	0.209	-0.408	0.419	0.109	0.682
MC (0.05~0.0711)	0.543	-0.291	-0.177	0.144	0.314
HC (≥0.0712)	0.248	0.980	0.033	0.747	0.005
Total	1.000	—	—	1.000	1.000

Table 5. Column Scores and Contribution of Column points to each dimension

Total Number Bacteria	Marginal Profile	Column Scores		Contribution	
		Dim 1	Dim 2	Dim 1	Dim 2
LB (<7)	0.233	-0.484	0.372	0.171	0.597
MB (7~17.99)	0.508	-0.260	-0.202	0.107	0.385
HB (≥18)	0.260	0.942	0.062	0.722	0.019
Total	1.000	—	—	1.000	1.000

Table 3-5 and Fig 1-4 showed that the higher carbon dioxide value in the environment was associated with higher total number of bacteria. The increases of carbon dioxide and the total number of bacteria in the air were associated with the opening time of indoor to public places. The higher humidity environment tended to be higher negative ion concentration.

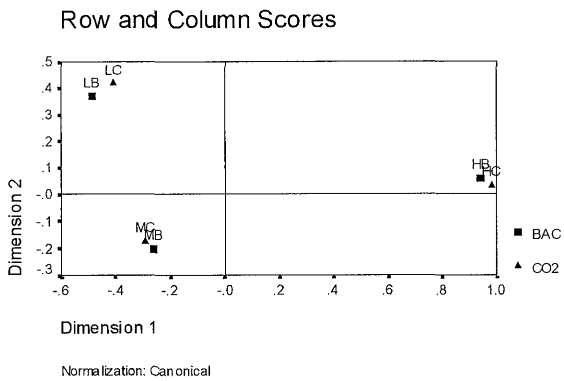

Figure 1. A correspondence relationship (combined plot) between BAC and CO2 for row and column points (JOINT chart)

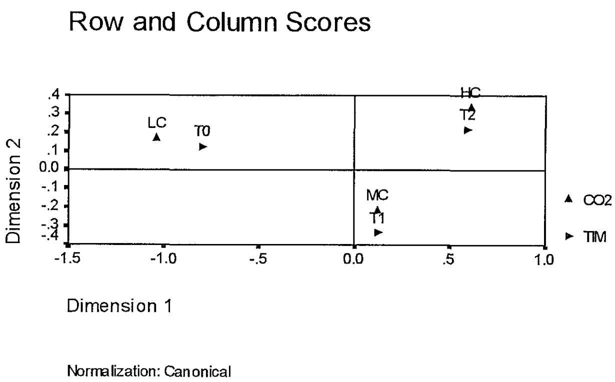

Figure 2. A correspondence relationship (combined plot) between CO2 and TIM for row and column points (JOINT chart)

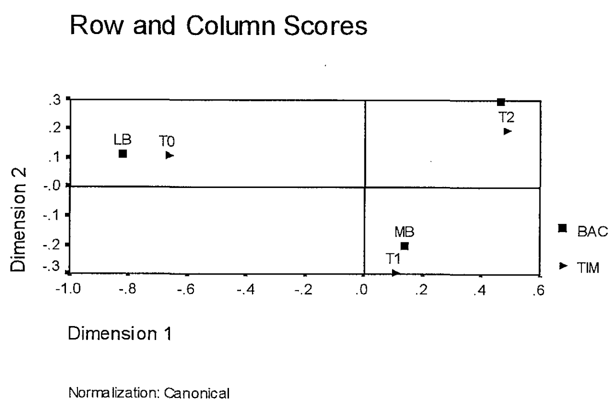

Figure 3. A correspondence relationship (combined plot) between BAC and TIM for row and column points (JOINT chart)

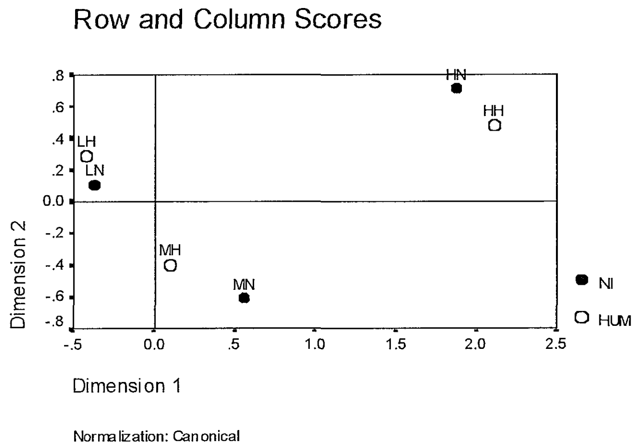

Figure 4. A correspondence relationship (combined plot) between NI and HUM for row and column points (JOINT chart)

The results by HOMALS:

The main running results are shown in Table 6-8. The following results can be printed out: Marginal Frequencies (FREQ, Table 6), the Eigen value of each dimension (EIGEN), Discrimination measures per variable per dimension, Category Quantifications of each Dimensions (Table 2, Table 7-8). The most important charts are showen in Fig 5-7 (The QUANT plot is introduced only in this paper, plotting Category Quantifications).

Fig 5-7 showed that under general conditions, the increases of CO2 density and BAC were associated with opening time, and with the decrease of NI. CO2

Table 6. The ranking classification of 4 variables, frequency and its code

Air Index (Var.)	Classification	Freq.	Ranking Class	Labels
Carbon Dioxide	<0.050	54	1	LC
(CO2)	0.050-	140	2	MC
	0.071-	64	3	HC
Total Number	<7	60	1	LB
Bacteria (BAC)	7-	131	2	MB
	18-	67	3	HB
Negative Ion	<10	47	1	LN
(NI)	10-	14	2	MN
	35-	5	3	HN
Opening Time	Before open.	80	1	T0
(TIM)	After open. 1 hr	88	2	T1
	After open. 2 hrs	90	3	T2

Table 7. Discrimination measures for 3 variables and 3 dimensions

Air Index (Var.)	Dim. 1	Dim. 2	Dim. 3
Total Number Bacteria (BAC)	0.452	0.426	0.042
Carbon Dioxide (CO2)	0.578	0.530	0.090
Opening Time (TIM)	0.495	0.244	0.843

Table 8. Category quantifications of 2 dimensions for 3 variables

Air Index (Var.)	Categories	Dim. 1	Dim. 2
Carbon Dioxide	1	1.107	0.815
(CO2)	2	0.090	-0.673
	3	-1.130	0.785
Total Number	1	0.927	0.407
Bacteria (BAC)	2	0.076	-0.624
	3	-0.980	0.855
Opening Time	1	0.958	0.651
(TIM)	2	-0.079	-0.554
	3	-0.774	-0.037

in the air showed the status of BAC and the level of NI. Increasing the ventilation decreased the accumulation of CO2 and BAC in the air, raised the concentration of NI, and improved the indoor air quality.

The HOMALS method could analyze three or more variables (indexes) of the air quality. It may be useful

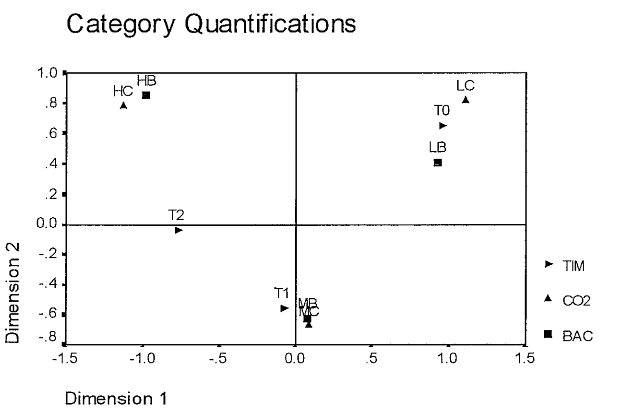

Figure 5. A homogeneity relationship (combined plot) among CO2, BAC and TIM for row and column points (QUANT chart)

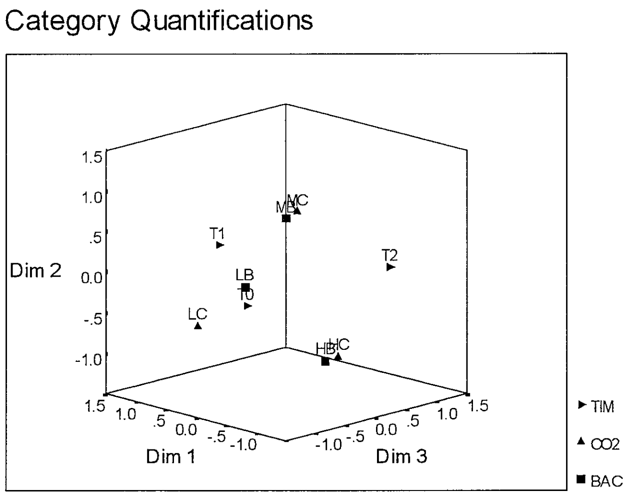

Figure 6 — 1. Homogeneity relationship (combined plot) among CO2, BAC and TIM for row and column points (3 Dim. Chart in SPSS 8.0)

to show the relationship between ranking categories and the air quality indexes by plotting with 2 or 3 axes. It may be a more sensitive and direct method than the analytical method of χ^2 test.

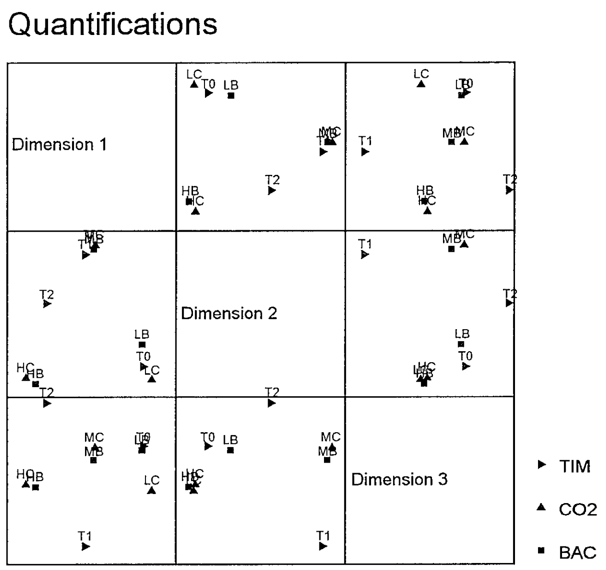

Figure 6 — 2. Homogeneity relationship (combined plot) among CO2, BAC and TIM for row and column points (3 Dim. Chart in SPSS 10.0)

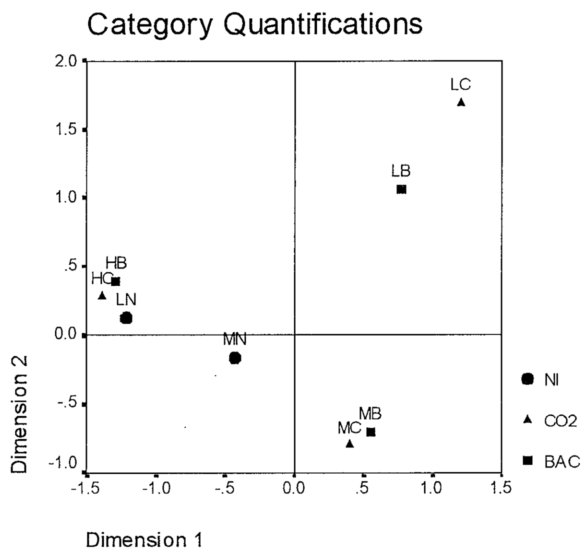

Figure 7. A homogeneity relationship within CO2, BAC and NI for the ranking classifications (2 Dim. chart)

Discussion

The ANACOR & HOMALS are more novel and imagery than classical methods. Three kinds of data may be collected in survey (quantitative, qualitative and ranked data). The methods, such as t-test, ANOVA and χ^2 test etc, have been used as classical statistical analysis. However, the ANACOR and HOMALS in SPSS 8.0/10.0 are used as new analytical methods recently. When the data are transformed into rank (for example in this paper, the air quality indexes were transformed into 3 ranks ("L", "M" and "H"), the ANACOR can analyze the relationships between different classifications for 2 variables, and can imaginably show the 2 dimensions charts additionally. Furthermore, the HOMALS can show 3 dimensions charts, and these charts are more easily understood and imaginable in SPSS 10.0 edition than in 8.0 edition.

There are more complex mathematical and statistical concepts involved in these two methods, for example, canonical correlation, principal component analysis, discriminatory analysis, eigenvalue, score, contribution and Euclidean distance etc. These concepts are worth studying by researchers. These two methods may be useful in order to obtain the charts of 2 or 3 dimensions more imaginably.

Conclusions

1. The relationships among 5 air quality indexes at indoor places (total number bacteria, carbon dioxide, humidity, negative ion and opening time) were analyzed by two new methods (ANACOR and HOMALS) in SPSS 8.0/10.0.
2. These two methods mentioned above could analyze the relationships between different classifications for 2 variables, and imaginably show 2 or 3 dimensions charts (HOMALS method). These charts were more novel, more easily understood and imaginable, but it is necessary to understand some complex statistical concepts involved in these analysis and plotting.

Acknowledgement

Mr. Chen Ming-Jun and Ms. Zheng Yan participated in this investigation.

References

1) SPSS Inc. SPSS Categories, Chicago: Marketing Department of SPSS Inc, B49-B62, 1990
2) Pan B, Que S, You M, et al: General Packages for Scientific Research (SPSS), Fuzhou: Fujian Science & Tech. Publish. House, 755-756, 1995
3) Pan B, Liu S, Chen Z, et al: Study on Birth Weight, Fetal Position and the Mother's Age by Homogeneity Analysis. Journal of Shanghai Medical University, 26(5): 385, 1999
4) Pan B, Que S, You M, et al: A Training Material on SPSS 8.0/9.0/10.0 for Windows, Jan.1999/2000