Comparison of Pattern Extraction Capability between Self-Organizing Maps and Principal Component Analysis

Iseri, Yoshihiko
Graduate Student, Department of Urban and Environmental Engineering, Graduate School of Engineering, Kyushu University

Matsuura, Tomonori
Faculty of Science/Graduate School of Science and Engineering, University of Toyama : Professor

Iizuka, Satoshi
The Storm, Flood and Landslide Research Department, National Research Institute for Earth Science and Disaster Prevention : Senior Researcher

Nishiyama, Koji
Assistant Professor, Department of Urban and Environmental Engineering, Faculty of Engineering, Kyushu University

他

出版情報：九州大学工学紀要. 69 (2), pp.37-47, 2009-06-22. 九州大学大学院工学研究院
バージョン：
権利関係：

Memoirs of the Faculty of Engineering, Kyushu University, Vol. 69, No.2, June 2009

Comparison of Pattern Extraction Capability between Self-Organizing Maps and Principal Component Analysis

by

Yoshihiko ISERI[*], Tomonori MATSUURA[**] , Satoshi IIZUKA[***], Koji NISHIYAMA[†] and Kenji JINNO[††]

(Received May 1, 2009)

Abstract

Principal Component Analysis (PCA) has been extensively used for multivariate data analysis including climate data analysis. However, recent studies suggest the possible use of Self-Organizing Maps (SOM) for climate data analysis. In order to adequately utilize the PCA and SOM, it is crucial to clarify the difference of both methods. This study compared the pattern extraction capability of SOM and PCA in order to clarify the difference of both methods. Comparisons of the methods were performed by conducting two kinds of simulations. The first stimulation confirmed SOM can extract non-orthogonal patterns while the patterns extracted by PCA are mutually orthogonal. The second simulation investigated how the patterns extracted by each method would change when new inputs were added. When new inputs were added, PCA did not extract the pattern in the original inputs and additional inputs. On the other hand, SOM extracted the pattern in original and additional inputs.

Keywords: SOM, PCA, Simulation, Pattern extraction, Comparison

[*] Graduate Student, Department of Urban and Environmental Engineering
[**] Professor, Faculty of Science/Graduate School of Science and Engineering, University of Toyama
[***] Senior Researcher, The Storm, Flood and Landslide Research Department, National Research Institute for Earth Science and Disaster Prevention
[†] Assistant Professor, Department of Urban and Environmental Engineering
[††] Professor, Department of Urban and Environmental Engineering

1. Introduction

Principal Component Analysis (PCA) has been utilized in many research fields [1]. The PCA (and rotated PCA) has been also extensively applied in meteorology and it has revealed important feature of climate fields such as teleconnection patterns [2], climate variability [3] and so on. However, deficits of PCA have been also recognized. Controversial problem in applying PCA for climate fields is whether the spatial patterns extracted by PCA is real physical mode or not. For example, Ito [4] considered physical reality of the Arctic Oscillation (AO), which is defined as the first PC of monthly sea level pressure in the northern hemisphere, from many aspects and concluded the AO is not real mode but apparent mode. The situation in which a principal component (Arctic Oscillation) exists between real physical modes is presented in the schematic diagram of his paper [4]. Dommenget and Latif [5] mentioned that PCs of the dominant mode are often superposition of many different modes and the center of action derived by PCA can be different from the centers of action of real physical modes. Considering these pitfalls, they recommended to analyze climate data with different statistical methods, not only by PCA. The recently proposed method of the Self-Organizing Maps (SOM) is likely to become a complementary or alternative tool of the PCA. Reush et al. [6] compared the pattern extraction capability of SOM and PCA using synthetic example, and showed SOM is capable of extracting pattern in data without superposition of input data. In fact, Recent studies using SOM [7)-11)] have given evidences that the SOM will become a indispensable tool for meteorological analysis.

As mentioned above, the PCA (and rotated PCA) and SOM are nowadays applied for meteorological analysis, and both methods have similarities such that they can reduce data through extracting patterns of given data. However, these methods are actually different and conduct pattern extraction in different way. In order to utilize both methods according to the aim of analysis, it is crucial to grasp the difference of both methods.

This study compared the pattern extraction capability of SOM and PCA in order to clarify the difference of both methods. The comparisons of both methods were conducted by performing two kinds of simulations. The first simulation compared pattern extraction capability for non-orthogonal patterns using simulated data which is similar to the example investigated by Dommenget and Latif [5]. The second simulation investigated how the extracted pattern would change when new inputs were added.

2. Methods

2.1 Self-Organizing Maps (SOM)

The SOM algorithm has been suggested by Kohonen [12), 13)] and it has been applied in wide ranges of studies such as engineering, medicine and agriculture. Recent studies in meteorology has also utilized SOM [7)-11)].

The SOM extracts patterns in high dimensional data and projects the extracted patterns on regularly arranged 2-dimensional grids. The 2-dimensional plane where the extracted patterns are arranged is called SOM-map and each of the grids on the SOM-map is denoted as node (or neuron). Each node has one reference vector, which shows extracted pattern and has the same dimension as input vector. Thereby if the SOM-map is composed of 100 nodes (e.g. 10×10 grids), 100 of reference vectors are arranged on the map. Reference vectors which are closely located on the SOM-map show similar patterns while the reference vectors separately located show different patterns. Each of the input vectors will be compared with all reference vectors and assigned into the

node where the most similar reference vector belongs to. In summary, the SOM algorithm extracts patterns from high dimensional data and arranges the extracted patterns on 2-dimensional plane (SOM-map). The SOM also classifies input data into the 2-dimensional plane by comparing the input vectors with the reference vectors.

The details of the basic version of the SOM algorithm [13] is given as follows. As the first step of the algorithm, map size (i.e. the number of nodes on the SOM-map) is determined and then give initial values of the reference vectors on the SOM-map. Random numbers can be used as the initial value of reference vectors although the first and second principal components obtained from input data can be used as initial value of reference vectors. As the next step, one of the input vectors is selected and the winner node c is identified by computing the distance between the selected input vector and all reference vectors. Although any metric can be used in order to compute the distance, Euclid distance is commonly utilized. The winner node c satisfies following equation for any i (i=1, 2, …, n; n is the number of nodes on the SOM-map).

$$\|x - m_c\| \le \|x - m_i\| \tag{1}$$

$x \in R^k$ is the selected input vector and k is the dimension of input vector. $m_i \in R^k$ is the reference vector in the i th node. After the winner node c is identified, reference vector m_i will be updated by the following function:

$$m_i(t+1) = m_i(t) + h_{ci}(t)\left[x(t) - m_i(t)\right], \quad (i = 1,2,\ldots,n) \tag{2}$$

Where t means regression step, and $h_{ci}(t)$ is a neighborhood function. In this study, the Gaussian is employed as neighborhood function:

$$h_{ci}(t) = \alpha(t) \cdot \exp\left(-\frac{\|r_i - r_c\|^2}{2\sigma^2(t)}\right) \tag{3}$$

Where $r_i \in R^2$ and $r_c \in R^2$ are position vectors of node i and c. Thereby, $\|r_i - r_c\|$ indicates the distance between the node i and the winner node c. $\alpha(t)$ is some monotonic decreasing function with t. $\sigma(t)$ is also some monotonic decreasing function and defines the width of the kernel. In this study, $\alpha(t)$ and $\sigma(t)$ are given as follows:

$$\begin{cases} \alpha(t) = \alpha(0)\dfrac{T-t}{T} \\ \sigma(t+1) = 1 + \left(\sigma(t) - 1\right)\dfrac{T-t}{T} \end{cases} \tag{4}$$

Where $\alpha(0)$ is the initial value of α, and T is the maximum of the regression step.

2.2 Principal Component Analysis (PCA)

Principal Component Analysis (PCA) has been widely used as multivariate technique, which reduces the multivariate data into a set of uncorrelated new variables [1]. In meteorology, data used for analysis is often described as spatial data or combination of many variables, and therefore the multivariate technique of PCA has been extensively utilized. Hannachi *et al.* [14] has provided a review on application of PCA in atmospheric science, mentioning PCA in atmospheric science is used for prediction purposes, smoothing purposes and the identification of teleconnection patterns.

Rotation of PCs [15), 16] is sometimes conducted in order to avoid pitfalls of PCA. Thereby comparison of SOM, PCA and rotated PCA would be informative. However, As this study rather focuses on the investigation of the properties of SOM, in order to simplify the comparisons and

clarify the characteristics of SOM, SOM and unrotated PCA are compared in this study.

The following describes the mathematical procedure to compute PCs [1]. Assume ($m \times n$) matrix D consists of the data to be analyzed. Firstly, mean of each column is subtracted so that the average of each column becomes 0. This matrix of 0 mean column will be called as D during the remain of this chapter. Then, variance-covariance matrix S (or correlation matrix R) is computed from the matrix D. The variance-covariance matrix S is compute by:

$$S = \frac{1}{m-1} D' D \tag{5}$$

The variance-covariance matrix S is a symmetric ($n \times n$) matrix. Then, eigenvalue λ and eigenvector e of ($n \times n$) matrix S (or R) are computed by solving the following equation :

$$Se = \lambda e \tag{6}$$

It is usual to give constraint of $\|e_k\| = 1 (k = 1,2,\cdots,n)$ when EOF analysis is conducted. The ratio of the variance explained by the j th PC to total variance is called as proportion. The proportion of the j th PC is computed as following:

$$proportion\ \ of\ the\ j\ th\ PC = \frac{\lambda_j}{\sum_{l=1}^{n} \lambda_l} \times 100\,(\%) \tag{7}$$

Where λ_j is the j th largest eigen value. In this study, PCs are computed using variance-covariance matrix.

3. Simulation data

3.1. Data description

This study compared the difference of SOM and PCA in order to fully utilize both methods for climate data analysis. Although there are many purposes in climate data analysis, one of the important objectives is the identification of individual variation modes in climate. Extraction of individual variation modes in climate enables the investigation of the relationships of each climate mode and deepen our knowledge on climate system. The simulation 1 aims to compare the SOM and PCA on the extraction capability of individual variation modes. In the simulation 1, the climate variations are presented by *vector a*±, *b*± and *c*±. That is, three of variation modes *a*, *b* and *c* are assumed and each mode has positive and negative phase.

In addition, detection of climate change is also very important objective of climate data analysis. Although how the evidence of climate change will appear in climate data is unknown, the climate change may be detected as the appearance of climate pattern which did not exist in the past. The simulation 2 investigates the performance of PCA and SOM on the detection of new pattern which did not appear in the past data. The detail of the simulation data 1 and 2 is explained in the following sections 3.2 and 3.3.

3.2. Simulation data 1

The simulation data 1 is composed of three kinds of variation modes. Two of the variation modes are orthogonal to each other while one variation mode is not orthogonal to the other two variation modes. We firstly prepared 6×3 matrix whose rows present each input and columns represent the dimensions of each input. Each input is composed of positive or negative phase of 3 kinds of variation modes which are presented by *vector a*±, *b*± and *c*±. These vectors are arranged from the 1st to the 6th rows of the 6×3 matrix. The angle between *vector a*+ and *b*+ (*a*- and *b*-) is

$$\underset{900\times 3}{\zeta} = \begin{pmatrix} 1.000 & 0.000 & 0.000 \\ -1.000 & 0.000 & 0.000 \\ 0.000 & 1.000 & 0.000 \\ 0.000 & -1.000 & 0.000 \\ 0.500 & 0.707 & 0.500 \\ -0.500 & -0.707 & -0.500 \\ 1.000 & 0.000 & 0.000 \\ -1.000 & 0.000 & 0.000 \\ 0.000 & 1.000 & 0.000 \\ \vdots & \vdots & \vdots \\ 0.500 & 0.707 & 0.500 \\ -0.500 & -0.707 & -0.500 \end{pmatrix} \begin{matrix} \leftarrow \textit{vector a+} \\ \leftarrow \textit{vector a-} \\ \leftarrow \textit{vector b+} \\ \leftarrow \textit{vector b-} \\ \leftarrow \textit{vector c+} \\ \leftarrow \textit{vector c-} \\ \\ \\ \\ \\ \\ \\ \end{matrix} \tag{8}$$

90º, *vector a*+ and *c*+ (*a*- and *c*-) is 60 º, *vector b*+ and *c*+ (*b*- and *c*-) is 45 º. Norm of each vector is standardized into 1 so that each vector (variation mode) has the same amplitude. The simulation data 1 is generated by concatenating 150 sets of the 6×3 matrix into one matrix. This yielded the 900×3 matrix ζ, which is composed of the *vector a*±, *b*± and *c*±. The PCA and SOM were respectively applied for this 900×3 matrix ζ (shown in equation (8)).

3.3 Simulation data 2

18 of data matrices were utilized in the simulation 2. Firstly, we prepared 4×3 matrix whose rows present each input and columns show dimensions of each input. The second and fourth row show variation mode of *vector A*+ (0.707, 0.707, 0) and *vector A*- (-0.707, -0.707, 0), respectively. The first and third row consist of random numbers taken from uniform distribution ranging from [-1 1]. By concatenating 250 sets of this matrix into one matrix, 1000×3 matrix was obtained. Next, 4×3 matrix $M_{4\times 3}$ was prepared. The second and fourth row of this matrix are respectively variation mode of *vector B*+ (0.707, 0, 0.707) and *vector B*- (-0.707, 0, -0.707). The first and third row are composed of random numbers taken from the uniform random distribution ranging from [-1 1]. Then, *k* sets (*k*=5, 10, 15, 20, 25, 35, 45, 55, 65, 75, 100, 125, 150, 175, 200, 225, 250) of $M_{4\times 3}$ were respectively added to the 1000×3 matrix. This provided data matrices which have rows of 1020, 1040, 1060, 1080, 1100, 1140, 1180, 1220, 1260, 1300, 1400, 1500, 1600, 1700, 1800, 1900, 2000 and each of the matrices has 3 columns. Eventually, 18 of data matrices ψ_n (n=1,2,⋯,18) were obtained. The PCA and SOM were respectively applied for each of the 18 data matrices (shown ins equation (9)).

$$\underset{1000\times 3}{\psi_1} = \begin{pmatrix} -0.838 & -0.013 & 0.097 \\ 0.707 & 0.707 & 0.000 \\ -0.397 & -0.175 & 0.728 \\ -0.707 & -0.707 & 0.000 \\ -0.640 & 0.961 & -0.087 \\ 0.707 & 0.707 & 0.000 \\ \vdots & \vdots & \vdots \\ -0.707 & -0.000 & -0.707 \end{pmatrix} \begin{matrix} \leftarrow \textit{vector A+} \\ \leftarrow \textit{vector A-} \\ \\ \\ \\ \\ \\ \\ \end{matrix}, \quad \underset{1020\times 3}{\psi_2} = \begin{pmatrix} \underset{1000\times 3}{\psi_1} \\ \hline 0.610 & 0.071 & -0.894 \\ 0.707 & 0.000 & 0.707 \\ 0.400 & -0.614 & -0.330 \\ -0.707 & 0.000 & -0.707 \\ -0.999 & -0.093 & -0.311 \\ \vdots & \vdots & \vdots \\ -0.533 & -0.635 & -0.460 \\ -0.707 & 0.000 & -0.707 \end{pmatrix} \begin{matrix} \\ \leftarrow \textit{vector B+} \\ , \leftarrow \textit{vector B-} \\ \\ \\ \\ \\ \\ \end{matrix}$$

$$\cdots\cdots, \quad \underset{2000\times 3}{\psi_{18}} = \left(\begin{array}{ccc} \multicolumn{3}{c}{\underset{1000\times 3}{\psi_1}} \\ \hline 0.610 & 0.071 & -0.894 \\ 0.707 & 0.000 & 0.707 \\ 0.400 & -0.614 & -0.330 \\ -0.707 & 0.000 & -0.707 \\ -0.999 & -0.093 & -0.311 \\ \vdots & \vdots & \vdots \\ 0.946 & -0.554 & 0.626 \\ -0.707 & 0.000 & -0.707 \end{array}\right) \tag{9}$$

4. Result of comparison

4.1 Result for simulation data 1

Simulation 1 applied PCA and SOM for 900×3 matrix ζ. **Table 1** gives PCs and proportion of each PC obtained by applying PCA for the matrix ζ. As seen in **Table 1**, none of the variation modes (i.e. *vector a*±, *b*± and *c*±) is presented as PC. It is also seen that PC 1 and PC 2 account for 62.2% and 33.3% of total variance respectively, indicating that 95.5% of total variance is explained by the first two PCs.

Figure 1 shows bar chart of reference vectors obtained by applying 3×4 SOM for data matrix ζ. The node number on the SOM-map is given from the left to right in descending order (i.e. node 1 for the top left corner, node 4 for the top right corner, node 9 for the bottom left corner and node 12 for the bottom right corner). The *vector a*± were clearly shown as the reference vector of node 3 (*vector a*+) and 10 (*vector a*-) respectively, while the *vector b*± were presented in node 4 (*vector b*+) and 9 *(vector b*-). The *vector c*±, which were not orthogonal to *vector a*± and *b*±, were also presented in the node 1 (*vector c*+) and node 12 (*vector c*-).

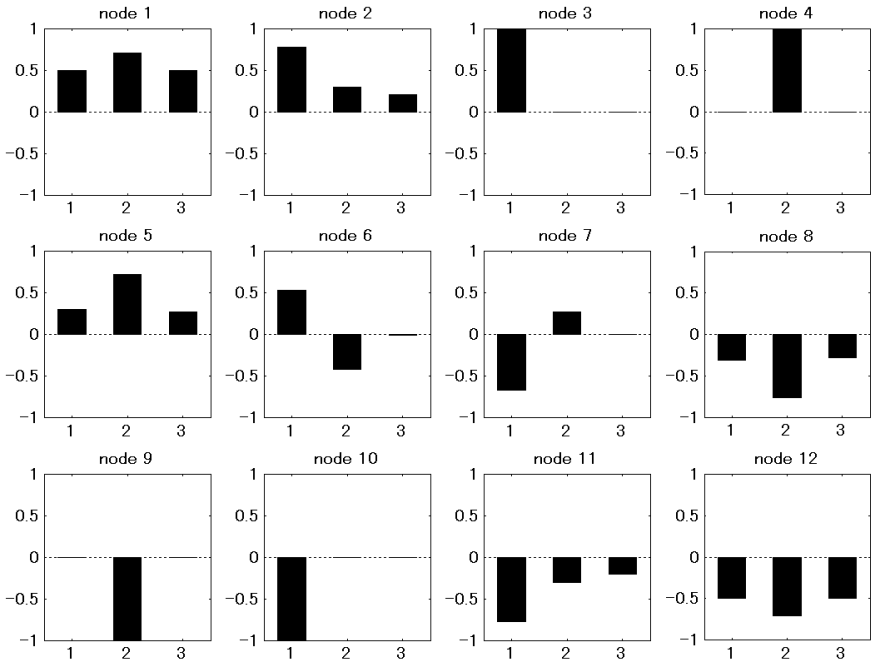

Fig. 1 Bar chart of reference vectors obtained by applying 3×4 SOM for data matrix ζ. The horizontal axis shows each dimension of reference vector in each node. Vertical axis shows the value of each dimension.

Table 1 Proportions and PCs obtained by applying PCA for ζ. Proportions are shown by percentage. Notice that each PC is composed of 3 dimensions.

	proportion (%)	**dimension 1**	**dimension 2**	**dimension 3**
PC 1	62.2	0.558	0.789	0.259
PC 2	33.3	0.816	-0.577	0.000
PC 3	4.5	-0.149	-0.211	0.966

4.2 Result for simulation data 2

Simulation 2 applied SOM and PCA for 18 of data matrices ψ_n (n=1,2,⋯,18) in order to investigate how the extracted patterns would change when new inputs were added. **Figure 2** shows the transition of PC 1, 2 and 3 against the increasing of new inputs. The horizontal axis shows the number of total inputs used for the computation of PCs and the vertical axis shows the value of each dimension. The vector A± (i.e. ± (0.707, 0.707, 0)) were perfectly presented as PC 1 when the number of inputs was 1000. However, the increasing of new inputs distorted the vector A± presented in the PC 1, and when the total inputs reached 2000, the PC 1 was completely different from the vector A±. The *vector A*± were neither presented in PC 2 nor PC 3. In addition, *vector B*± (i.e. ± (0.707, 0, 0.707)) were not presented in any of the PCs even if total inputs reached 2000. The transition of proportions of each PC is shown in **Fig. 3**. **Figure 3** shows that the addition of new inputs decreased the proportion of PC 1 while increased the proportion of PC 2. However, the changes of the proportion of PC 1 and PC 2 almost stopped when the total inputs reached approximately 1600.

Figure 4 shows the transition of reference vectors which are the most similar to each of the variation modes (i.e. *vector A*± and *B*±). The most similar reference vector was determined as the

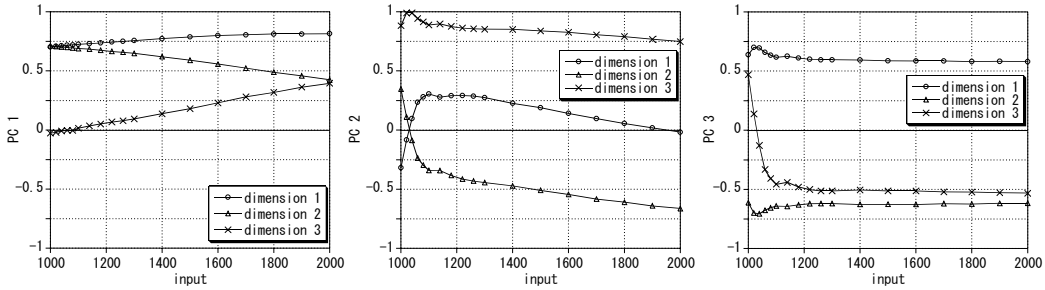

Fig. 2 Transition of PC 1 (left), PC 2 (center) and PC 3 (right) against the increasing of inputs.

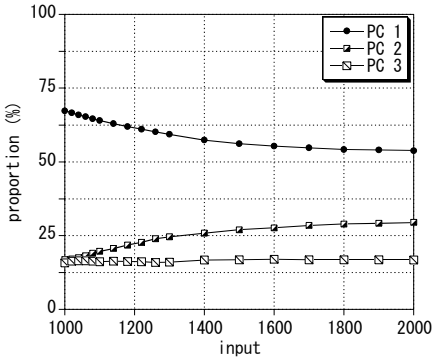

Fig. 3 Transition of proportion of each PC against the increasing of inputs. The proportion is shown by percentage.

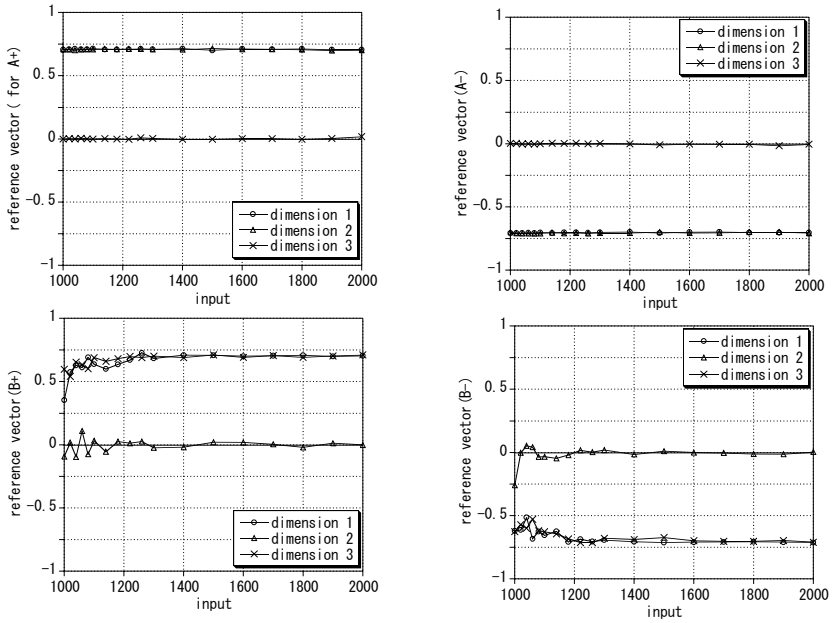

Fig. 4 Transition of the most similar reference vectors for *vector A*+ (upper left), *A*- (upper right), *B*+ (lower left) and *B*- (lower right) against increasing of inputs.

reference vector which showed the smallest root mean square error for corresponding variation vector. As seen in **Fig. 4**, *vector A*± were clearly represented in reference vectors even though new inputs were added.

5. Discussion

5.1 Discussion for simulation 1

The result confirmed that the patterns extracted by PCA can be different from the individual modes of original variations [15]. There are at least two reasons for this consequence. The first reason is that if there are many variation modes in the data, PCA identifies somewhere between the variation modes as PC 1, because PCA detects the direction of maximum variance as PC 1. **Figure 5** shows PC 1 and v*ector a*+, *b*+ and *c*+. As seen in **Fig. 5**, the direction of maximum variance, which is presented by the PC 1, exists between the 3 vectors. The second reason is that the patterns extracted by PCA need to be orthogonal to each other. Because of this property, if the individual variation modes are not orthogonal to each other, at least one of the variation modes can not be presented as PC. In this simulation, the PC 1 was quite different from all of the original variation modes (see **Fig. 5**) and the angle between the PC 1 and *vector a*+, *b*+, *c*+ were 56.08º, 37.91º and 14.91º, respectively. This means that all vectors were not orthogonal to the PC 1, and thereby the remaining two PCs (i.e. PC 2 and PC 3), which need to be orthogonal to PC 1, were also different from all of the original variation modes. On the other hand, the SOM, which is free from the orthogonal constraint, succeeded in extracting all individual variation modes separately. This capability of SOM on the separation of individual variation patterns is consistent with the discussion given in earlier studies [6), 11)].

The result also clearly showed difference of each method in how the relationships of extracted patterns are presented. Although PCA guarantees the extracted patterns are orthogonal to each other, SOM arranges the obtained patterns on the 2-dimensional map, allowing to compare the similarities (or dissimilarities) of the extracted patterns by their relative locations on the map. Moreover, as

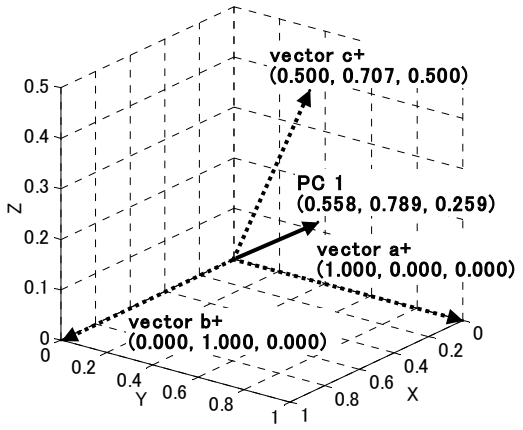

Fig. 5 3-dimension plot of PC 1 and *vector a*+, *b*+ and *c*+. PC 1 is shown by solid arrow. *Vector a*+, *b*+ and *c*+ are shown by dotted arrow.

pointed out in several studies [6), 9), 11)], various patterns such as transition pattern from a positive phase to negative phase can be presented on the map.

5.2 Discussion for simulation 2

The result indicates that the SOM has good capability of detecting new patterns which did not exist in the past data. **Figure 4** suggests that new pattern can be immediately reflected to reference vector by addition of new inputs. In this simulation, addition of just 20 inputs significantly improved the most similar reference vector to v*ector B*±, which is new pattern, and the addition of 40 inputs gave fairly good approximation of *vector B*±. Increasing of new inputs kept decreasing the RMSEs for *vector B*±. Eventually, addition of 500 new inputs gave reasonable approximation of *vector B*±, giving the most similar reference vector for *vector B*+ as (0.70707, 0.01958, 0.70922) with the RMSE of 0.01971. **Figure 4** also presents SOM succeeded in extracting *vector A*± as reference vectors, even if the number of new inputs reached 1000.

On the other hand, as shown in the **Fig. 2**, the *vector B*± were not presented as PC even though the addition of new inputs reached 1000. **Figure 2** also showed PCA failed in retaining the *vector A*± as PC, when new pattern were added. **Figure 6** shows the transition of PC 1 against the increasing of new inputs. As clearly seen in the **Fig. 6,** the direction of PC 1 gradually changed from *vector A*+ (original pattern) to *vector B*+ (new pattern) as the number of new inputs increased. The reason for this change is that the addition of new inputs decreased the proportion of *vector A*± while the proportion of *vector B*± increased, and this resulted in another direction of maximum variance which is related to the *vector B*±. Thereby the direction of PC 1 (direction of maximum variance) changed when new inputs were added. Eventually, addition of 1000 of new inputs captured neither the *vector A*± nor *vector B*±, giving the direction between the *vector A*+ and *B*+ as the direction of PC 1. In short, the SOM showed good capability of detecting new pattern while preserving original pattern. On the other hand, PCA showed difficulty in finding new pattern, and it even distorted the original patterns when new patterns were added.

Although this simulation indicated SOM is able to extract the original pattern and new pattern when new pattern is added, addition of too many new inputs which contains a number of new patterns maybe deteriorates the original patterns because SOM can not extract the patterns which exceed the number of reference vectors on the map. However, even in such a situation, enlargement of map size would still keep the original patterns and also extract new patterns as reference vectors. Notice that although the random numbers taken from uniform distribution were also added, the

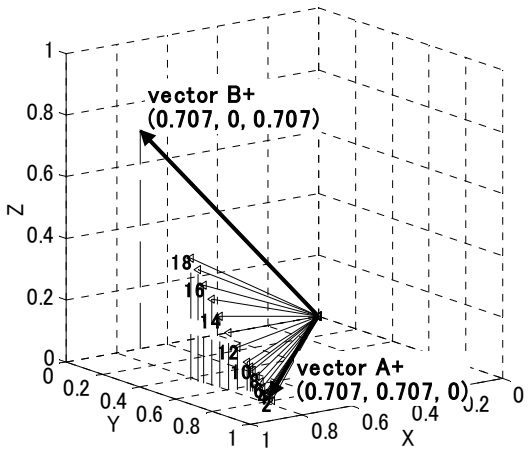

Fig. 6 3-dimensional plot of *vector A+*, *B+* and PC 1 of 18 data matrices. The thick solid arrows are vector A+ and B+. The thin solid arrows are PCs. The numbers labeled near by the thick solid arrows show data matrix number. (the data matrix of even numbers are labeled).

SOM succeeded in detecting pattern hidden in the noised data. This robustness of SOM for extracting patterns in the data with noise is consistent with the results of Reush *et al.* [6].

The capability of SOM for detecting new pattern with keeping pre-existing pattern suggests that the SOM would be useful for the investigation of climate change. In fact, the studies applying SOM for climate data [9), 11)] succeeded in detecting decadal scale of climate variability. In order to detect climate change signals, it is important to compare the past climate patterns with the recent climate patterns. While the simulation indicates the PCA is insensitive in extracting new patterns (one of the possible reasons for this is that the new pattern does not account for significant amount of total variance), SOM seems to have good capability of detecting new pattern with preserving pre-existing pattern. This capability of SOM allows to compare the past pattern with new pattern and to analyze what changes have happened from the past to the present. If the patterns associated with climate change is orthogonal to the dominant pattern in the past, the PCA would be able to extract new climate pattern as PCs, however, the patterns associated with climate change is not likely to be orthogonal to the past climate pattern, and it will take a long time until the new pattern occupies significant amount of total variance. Considering these results, the use of SOM for the extraction of the past and new pattern in climate data and the comparison of the extracted patterns is likely to provide useful information which can not be obtained by PCA.

6. Conclusion

This study compared pattern extraction capability of SOM and PCA using simulated data. The result indicated the followings.

- The result confirmed that SOM can extract the patterns which are not orthogonal to each other while PCA extracts the patterns as set of orthogonal patterns. The capability of SOM for extracting non-orthogonal patterns provides more various patterns than PCA.
- As PCA is expected to achieve, PCA extracts the PC as the direction of maximum variance. However, as shown in the simulations (also discussed in literatures [5), 6), 11)]), the direction of maximum variance can be obtained as the mixing mode of individual variation modes.
- When new pattern was added, SOM succeeded in extracting both pre-existing pattern and new pattern as reference vectors while PCA extracted neither pre-existing pattern nor new pattern. Because the PCA identifies the direction of maximum variance, the addition of new pattern

resulted in the identification of pattern which exist in the pre-existing pattern and new pattern.

It should be noticed this study conducted the simulations under the very simplified situation (i.e. only a few patterns are included in the data) and compared just some aspects of both methods despite these methods have a number of ways to be utilized. Therefore, application of the methods for actual meteorological data and evaluation of other aspects of these methods such as classification capabilities should be conducted in future study.

Acknowledgements

We are grateful to Prof. Ryuichi Kawamura for constructive comments on this study.

References

1) I.T., Jolliffe: Principal component analysis, *Springer series in statistics*, (1986).
2) J. M., Wallace and D. S. Gutzler; Teleconnections in the geopotential height field during the northern hemisphere winter, *Monthly Weather Review.*, Vol.109, No.4, pp.784-812 (1981).
3) J. D., Horel,: A rotated principal component analysis of the interannual variability of the northern hemisphere 500mb height field, *Monthly Weather review*, Vol.109, No.10, pp.2080-2092 (1981).
4) H., Ito; Reconsideration of the true versus apparent arctic oscillation, *Journal of climate*, Vol.21, No.10, pp.2047-2062 (2008).
5) D., Dommenget, and M. Latif; A cautionary note on the interpretation of EOFs, *Journal of Climate*, Vol.15, No.2, pp.216-225 (2002).
6) D. B., Reush, R. B. Alley, et al.; Relative performance of self-organizing maps and principal component analysis in pattern extraction from synthetic climatological data, *Polar geography*, Vol.29, No.3, pp.188-212 (2005).
7) T., Cavazos; Using self-organizing maps to investigate extreme climate events: an application to wintertime precipitation in Balkans, *Journal of climate*, Vol.13, No.10, pp.1718-1732 (2000).
8) K., Nishiyama, S. Endo, et al.; Identification of typical synoptic patterns causing heavy rainfall in the rainy season in Japan by a self-organizing map, *Atmospheric Research*, Vol.83, No.2-4, pp.185-200 (2007).
9) J., Leloup, Z., Lachkar, et al.; Detecting decadal changes in ENSO using neural networks, *Climate dynamics*, Vol.28, No.2-3, pp.147-162 (2007).
10) J., Leloup, M. Lengaigne, et al.; Twentieth century ENSO characteristics in the IPCC database, *Climate dynamics*, Vol.30, No.2-3, pp.277-291 (2008).
11) T., Tozuka, J.-J. Luo, et al.; Tropical Indian ocean variability revealed by self-organizing maps, *Climate dynamics*, Vol.31, No.2-3, pp.333-343. (2008).
12) T. Kohonen; Self-organized formation of topologically correct feature maps, *Biological cybernetics*, Vol.43, No.1, pp.59-69 (1982).
13) T. Kohonen; Self-organizing maps, *Springer-Verlag*, Heidelberg (2001).
14) A., Hannachi, I.T., Jolliffe, et al.; Empirical orthogonal functions and related techniques in atmospheric science: A review, *International journal of climatology*, Vol.27, No.9, pp.1119-1152 (2007).
15) M. B., Richman: Rotation of principal components, *Journal of climatology*, Vol.6, No.3, pp.293- 335 (1986).
16) R. Kawamura: A rotated EOF analysis of global sea-surface temperature variability with interannual and interdecadal scales, *Journal of physical oceanography*, Vol.24, No.3, pp.707-715 (1994).