Basic Method of Dynamic Modeling of the Global Change

Wasaburo Unno
Institute for Science and Technology, Kinki University
Higashi-Osaka, Osaka, 577 Japan

(Received December 28, 1992)

Abstract

The method of dynamic modeling is explored and applied to the global change. The repetitive use of principal component analyses (PCA) in multidimensional phase space describing the dynamical state of the global environment is the main tactics of the method. For various quantities monitored at sites distributed over the entire globe, amplitudes of variation in shorter time scales as well as the relevant time averages are taken as additional variables in order for revealing the nonlinear couplings among different time scales. Spatial and temporal derivatives of all those variables are then taken also as additional variables for describing the evolution of the dynamical state of the system.

The first PCA is intended to obtain the description of the system in terms of the principal components, and the second PCA is to describe the dynamical properties of the system. A method of maximum-entropy adjustment of the data is also explored. Thus, large amount of fragmentary data scattered in wide range of accuracy can be used in the PCA constructively, and the applicability of the method of dynamic modeling is greatly extended in the study of the global change as well as in other branches of the natural history.

Key words: Global change, Dynamical system, Principal component analysis, Natural history

1 Introduction

Population, Energy, and Environment are generally agreed as constituting three major problems of the coming century. They are strongly coupled with each other, but especially the third one, the environment, carries the results of all the natural and human activities. Therefore, the world-wide cooperation based on the deep understandings of the problem is urgently called for.

The most important direct quantity representing the global environment condition may be the annual average of the average global surface temperature. But, all the quantities that affect it and that are affected by it constitute the nonlinear dynamical system. Since the global environment is basically conditioned by liquid water and oxidizing atmosphere that have been continuously reformed by Life, special care must be taken for its complex characteristics as a multi-dimensional nonlinear dynamical system varying in wide ranges of time scales. An important aspect of the environmental problem is, therefore, the interaction of modes of different time scales.

The biggest concern of the present day, the global warming, seems to be mostly concentrated on the time scale of the order of 10 to 100 years. But, the daily variation and the seasonal variation are basic in meteorology, and the turbulent diffusivities that influence the global warming should depend on these shorter time variations. Also, there are long term variations such as deep ocean circulation having the time scale of the order of 1000 years. In the latter case, salt water cooled down to about 4°C in polar circumferential regions creeps to the bottom of ocean capturing CO_2 from the atmosphere and finally wells up in different parts of ocean, thus concerning the cooling of the earth in two different mechanisms. Therefore, not only those quantities that describe the present state of the global warming but also those quantities that represent variations of shorter and longer time scales must be included in the analysis of the global warming.

Our next concern is how to represent the dynamical system. Let us consider a huge dynamical system governed by a large set of partial differential equation. In general, we don't know how many modes of motion are excited and what are exciting mechanisms in the system. But, we can in principle measure various quantities, $Q_i(s_j, t_k)$, in various sites, s_j, at various time, t_k. Thus, we can construct a multi-dimensional phase space composed of the quantities Q_i's, and the state in each site at each time is represented by a point in that space. This description, however, is a static representation, which should be extended to a dynamic representation for chaotic dynamical systems. In other words, the dimension of the phase space has to be enlarged to include the dimensions representing temporal and spatial derivatives that can be generated from $Q_i(s_j, t_k)$.

The distribution of points in the multi-dimensional space of the dynamic representation thus constructed may be best analyzed by means of the principal component analysis (PCA). The dynamic state of the system, however cannot be fully accounted for by a single trial of PCA. The repetition of PCA, first for determining the static structure and then for hunting the mechanism, seems to be indispensable. The construction of this repetitive PCA scheme will be discussed later as the central task of the present paper.

The PCA, however, is very much restricted ordinarily by the lack of complete data to be possibly plotted in the multi-dimensional phase space. Especially, in the global warming study, the data are abundant but fragmentary, and random in accuracy and often lacking for particular sites and occasions. The PCA for incomplete data has been recently proposed by Unno and Yuasa (1992). The data are supplemented by the adjustment in the maximum entropy technique proper for PCA and even fragmentary data can be used constructively for defining the state of the system. The method will be summarized briefly in a subsequent section of this paper.

2 Construction of Data Base

Let $Q_i^{(s,t)}$ and $w_i^{(s,t)}$ denote the value of a monitored quantity, Q_i and the associated weight of the observation made at site s and at time t. From that, we can construct $Q_i^{(s_j,t_k)}$ and $w_i^{(s_j,t_k)}$ by interpolation, where s_j and t_k represent one of the mesh points of site s_j and time t_k. We will work exclusively on these values hereafter.

The weight $w_i^{(s_j,t_k)}$ is taken to be as large as $\exp[-\Delta^2]$ for the measurement of which the estimated probable error is Δ in units of the root-mean-square dispersion of Q_i. In particular, $w_i^{(s_j,t_k)}$ can be taken to be unity for high accuracy measurements and zero if no measurements are made. These $Q_i^{(s_j,t_k)}$ and $w_i^{(s_j,t_k)}$ (i=1...I, (j,k)=1...N) are the primary source data from which the multi-dimensional data set can be constructed. Subscripts j and k will be omitted from (s_j, t_k) hereafter, unless confusions may not happen.

We are now interested in the dynamics of time scale of the order of 10 years or more. Therefore, all the values of $Q_i^{(s,t)}$ are understood as the running averages over one year. However, since we are also interested in the coupling between the mode of the time scale in question and modes of daily and seasonal variations, the annual averages of the amplitudes of the daily and seasonal variations, denoted as $DQ_i^{(s,t)}$ and $SQ_i^{(s,t)}$, should be included in the data base. Thus, the quantity Q_i is extended to the vector, $\{Q_i\} = (Q_i, DQ_i, SQ_i)$.

Now, we turn to include the spatial (longitudinal and latitudinal) derivatives, ∇_λ and ∇_ϕ, and possibly the Laplacian, ∇^2, that are calculated from the original $Q_i^{(s,t)}$ data with the help of the interpolation formula in appropriate spatial intervals. The corresponding weights can be estimated accordingly. The inclusion of these spatial derivatives to the data base enables us to represent the dynamical system defined implicitly by a set of partial differential equations. Even if our primary interest is concerned with the global warming, sustaining mechanisms can have spatial structures that are substantially smaller than the earth radius.

3 Time Derivatives and Time Lags

Lastly, we consider the inclusion of temporal derivatives to the data base. In the analysis of the chaotic dynamical system , we usually employ the m-dimensional vector space composed of $\{X(t), X(t+\tau), ..., X(t+(m-1)\tau)$ belonging to time t (Takens, 1981). Here, τ denotes a representative time interval for the chaos. Equivalently, we can take time derivatives from the 0-th to the (m-1)-th. But, in treating the global warming, it would be more practical to employ only the first time derivative or at most up to the second time derivatives of many quantities, $\{Q_i\}$, by taking a number of quantities (dimension of embedding space) exceeding (2n+1) times the Hausdorf dimension, n, of the system ($3I > 2n+1$, or $6I > 2n+1$), (Takens, 1981). If necessary, the time derivatives of the spatial derivatives can be used as well. The correlations of these time derivatives, $\dot{Q}_i^{(s,t)}$ etc., with the principal components in the PCA describing the system should reveal the dynamic mechanisms driving the variation of the system, provided that the latter principal components could describe the system completely.

There is, however, another aspect that may escape from the above description. That is, the time lag problem. For instance, Bartusiak found a good correlation between the sunspot number and the stratospheric temperature in the north polar region when the west wind prevails in the equatorial stratoshere, but a good anti-correlation in the east wind period (see e.g. Nemoto,1992). In that case, the time lag of a half solar cycle to the north polar stratoshere temperature data for the period of east tropical stratospheric wind will bring a universally good correlation with the sunspot data. This example demonstrates the importance of a triple correlation which is usually disregarded in the PCA. Therefore, if such a triple correlation is known beforehand, the introduction of an appropriate time lag will be effective in deriving a significant principal component which would be otherwise overlooked.

4 Multi-Dimensional Representation

Various observational quantities, Q_i, amplitudes of their daily and seasonal variations, DQ_i and SQ_i, and spatial derivatives, ∇_λ and ∇_ϕ, and the Laplacian, ∇^2, of all of those quantities are denoted hereafter simply as Q_i. All of these quantities $Q_i^{(s,t)}$ are understood as the annual average values given for every site s_j and for every time t_k, and the corresponding weights of determination are denoted by $w_i^{(s,t)}$. Also, we have the time derivatives, $\dot{Q}_i(s,t)$, and the corresponding weights of determination, $w(\dot{Q}_i)^{(s,t)}$. These data form the content of the second data base which is the source data base for the PCA.

Before doing PCA, however, those Q_i and $\dot{Q}_i$ values should better be normalized to give q_i and $\dot{q}_i$ as follows,

$$q_i = [Q_i - < Q_i >]/\sigma_i, \quad (i = 1, 2, ..., m) \quad (1)$$

where $< Q_i >$ denotes the average over the whole (s_j, t_k) data and σ_i the rms dispersion,

$$< Q_i > = (\sum_{s,t} w_i^{(s,t)})^{-1} \sum_{s,t} w_i^{(s,t)} Q_i^{(s,t)} \quad (2)$$

and

$$\sigma_i^2 = (\sum_{s,t} w_i^{(s,t)})^{-1} \sum_{s,t} w_i^{(s,t)} [Q_i^{(s,t)} - < Q_i >]^2. \quad (3)$$

The phase space which is built by q_i-coordinates forms the embedding space in which dynamic characteristics are implicitly represented by observational quantities. The first PCA for obtaining principal components representing the system should be worked out in this space.

In the second PCA intended for finding mechanisms, the embedding space for the explicit dynamic representation should be composed of $\dot{q}_i$ in addition to the significant principal components p_l already obtained in the first PCA. Since no new significant principal components should appear, the result of the second PCA will give the expression of $\dot{q}_i$ in terms of a linear combination of p_l's, if the system was described completely by the principal components p_l.

5 Data Supplement

The usual PCA treats the case that the set $(q_1, q_2, ..., q_m)$ are measured with accuracy given by the weight $w_i^{(s,t)}$ for every site and time (s_j, t_k). This is not the case in the global warming in which $w_i^{(s,t)}$ is attributed to each $q_i^{(s,t)}$. We have formulated the maximum entropy method of the data supplement (Unno and Yuasa,1991) to make numerous imperfect data to be accessible to the PCA formalism.

To each pair of data $[w_i^{(s,t)}, q_i^{(s,t)}]$, we now propose to supplement a virtual pair $[v_i^{(s,t)}, x_i^{(s,t)}]$ such that

$$v_i^{(s,t)} = 1 - w_i^{(s,t)}, \qquad (4)$$

and the whole set of data including the virtual added data attains the maximum probability distribution under the following statistical constraints,

$$\sum_{s,t} v_i^{(s,t)} x_i^{(s,t)} = 0,$$

and

$$\sum_{s,t} v_i^{(s,t)} [x_i^{(s,t)}]^2 = \sum_{s,t} v_i^{(s,t)}, \qquad (5)$$

for all i. Thus, the set of measured q_i- values sprits into 2m subsets for each (s_j, t_k)-measurement in which the i-th variable attains either $q_i^{(s,t)}$ or $x_i(s,t)$. The weight for each subset is given by $\Pi_{i=1}^{m}[w_i, v_i]^{(s,t)}$ in which $[w_i,\ v_i]$ represents either w_i or v_i according as $q_i^{(s,t)}$ or $x_i^{(s,t)}$ is adopted for q_i in the subset. The variational determination of $x_i^{(s,t)}$ will be given later. We now proceed to the first PCA generalized for imperfect data. For the second PCA, similar supplementation of data should be performed for the time derivatives as well.

6 Generalized PCA

The principal components are the axes of the ellipsoidal distribution of datum points in the embedding space. The principal p_l of a point $(q_1, q_2, ..., q_m)$, for the perfect data $(w_i = 1)$, is expressed by the equation

$$p_l = \sum_{i=1}^{m} \mu_{li} q_i, \qquad (6)$$

in which the direction cosines μ_{li} are given by the condition that the sum of the principal components squared $\sum_{s,t}[p_l^{(s,t)}]^2$ be extremum for the variation of μ_{li} under the constraint that $\sum_{i=1}^{m} \mu_{li}^2 = 1$. In the present general case, 2^m subsets corresponding to two values (q_i, x_i) in every coordinate belong to every one of (s,t)-sets. Therefore, the above expression of p_l is generalized to be

$$p_l = \sum_{i=1}^{m} \mu_{li}(w_i q_i + v_i x_i). \qquad (7)$$

In the expression of p_l^2, however, terms having weight factors like $w_i v_i$ (same subscripts) should be omitted. We obtain,

$$p_l^2 = \sum_{i=1}^{m} [\mu_{li}^2 (w_i q_i^2 + v_i x_i^2)$$
$$+ 2 \sum_{j>k} \mu_{li}\mu_{lj}(w_i q_i + v_i x_i)(w_j q_j + v_j x_j)]. \qquad (8)$$

Note that the weight of the q_i^2 term is not w_i^2 but w_i and that the $q_i x_i$ terms do not exist. This expression of p_l^2 holds for each (j,k) set which is composed of 2^m subsets.

The eigenvector $((\mu_{li}))(i=1,2,...,m)$ is determined from the variation of the following variation function,

$$S = S_0 - (\lambda_l/2)[\sum_{i=1}^{m} \mu_{li}^2 - 1], \qquad (9)$$

and

$$S_0 = (1/2N) \sum_{(s,t)=1}^{N} [p_l^{(s,t)}]^2 =$$

$$\sum_{i=1}^{m} [(1/2)\mu_{li}^2 + \sum_{i<j} \mu_{li}\mu_{lj} r_{ij}], \qquad (10)$$

where N is the total number of observing points (site s and time t) and r_{ij} denotes the correlation coefficient given by

$$r_{ij} = \sum_{(s,t)=1}^{N} (w_i^{(s,t)} q_i^{(s,t)} + v_i^{(s,t)} x_i^{(s,t)})$$

$$\times (w_j^{(s,t)} q_j^{(s,t)} + v_j^{(s,t)} x_j^{(s,t)}), \qquad (11)$$

an expression which is reducible to the correlations $< q_i q_j >$ if the distributions of x and q are statistically the same. Thus, the generalized PCA equations are reduced to the same form as the usual PCA equations,

$$\sum_{j=1}^{m} [(1-\lambda_l)\delta_{ij} + r_{ij}(1-\delta_{ij})\mu_{lj}] = 0, \ (i=1,2,...,m). \qquad (12)$$

The eigenvalues λ_l (l=1,2,...,m) and the corresponding eigenvectors (μ_{li}) are then obtained.

Now, it is easy to show that the rms dispersion of $p_l^{(s,t)}$ is equal to λ_l on account of equations (12),

$$\lambda_l = (1/N) \sum_{(s,t)=1}^{N} [p_l^{(s,t)}]^2. \qquad (13)$$

The probability of finding a set of observations having principal components $(p_1, p_2, ..., p_m)$ in a volume element $dp_1\, dp_2 ... dp_m$ obeys the ellipsoidal distribution

$$W(\mathbf{p})d\mathbf{p} =$$

$$[(2\pi)^m \prod_{l=1}^{m} \lambda_l]^{-1/2} \ \exp[-\sum_{l=1}^{m}(p_l^2/2\lambda_l)]d\mathbf{p}, \qquad (14)$$

ensuring the average of p_l^2 to be λ_l.

7 The Most Probable Data

We are now at the position to determine $x_i^{(s,t)}$ to provide the most probable data with the condition that $W(\mathbf{p})$ be maximum. There are two constraints for $x_i^{(s,t)}$ such that

$$\sum_{(s,t)=1}^{N} (w_i^{(s,t)} q_i^{(s,t)} + v_i^{(s,t)} x_i^{(s,t)}) = 0, \qquad (15)$$

and

$$\sum_{(s,t)=1}^{N} (w_i^{(s,t)} q_i^{(s,t)^2} + v_i^{(s,t)} x_i^{(s,t)^2}) = 1. \qquad (16)$$

However, these constraints can be fulfilled prior to the PCA by the zero point readjustment and by renormalization (by changing σ_i) after the following variational determination of $x_i^{(s,t)}$ has been done.

The variation function for determining $x_i^{(s,t)}$ is given by

$$W = \sum_{(s,t)=1}^{N} W^{(s,t)} = \sum_{(s,t)=1}^{N} [-(1/2)\sum_{l=1}^{m}(1/\lambda_l)p_l^{(s,t)^2}], \qquad (17)$$

where $p_l^{(s,t)^2}$ has been given in equation (8).

Taking the variation $\delta x_i^{(s,t)}$, we obtain

$$\frac{\delta W}{\delta x_i^{(s,t)}} = -v_i^{(s,t)} \sum_{l=1}^{m} \frac{1}{\lambda_l} [\mu_{li}^2 x_i^{(s,t)}$$

$$+ \sum_{j \neq i} \mu_{li}\mu_{lj}(w_j^{(s,t)} q_j^{(s,t)} + v_j^{(s,t)} x_j^{(s,t)})] = 0. \qquad (18)$$

This is a set of simultaneous linear algebraic equations $(i=1,2,...,m)$, decoupled with different (s,t) sets. The solution $x_i^{(s,t)}$ is easily obtained in linear combination of $q_j^{(s,t)}(j=1,2,...,m)$.

For illustration, for the case of $m=2$, we obtain the eigenvalues,

$$\lambda_1 = 1 + r, \quad \text{and} \quad \lambda_2 = 1 - r \quad (\text{for } r > 0), \qquad (19)$$

omitting superscript $^{(s,t)}$ in the variables and subscript $_{12}$ in r_{12}. The corresponding eigenvectors (principal components) are

$$\begin{pmatrix} \mu_{11} \\ \mu_{12} \end{pmatrix} = \begin{pmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} \mu_{21} \\ \mu_{22} \end{pmatrix} = \begin{pmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix}. \qquad (20)$$

Then, equations (18) are reduced to

$$x_1 = rv_2x_2 = rw_2x_2 = rw_2q_2$$

and

$$-rv_1x_1 + x_2 = rw_1q_1, \quad (21)$$

and the solution is given by

$$x_1 = \frac{r(rv_2w_1q_1 + w_2q2)}{1 - r^2v_1v_2}$$

and

$$x_2 = \frac{r(w_1q_1 + rv_1w_2q_2)}{1 - r^2v_1v_2}. \quad (22)$$

For small r, x_1 depends more on q_2, and x_2 on q_1.

Whenever $q_i^{(s,t)}$ is poorly observed ($w_i^{(s,t)} \simeq 0$), the $x_i^{(s,t)}$ value can be employed in place of $q_i^{(s,t)}$ for representing the state (s,t).

8 The Second PCA

Let us suppose that the observation is made for so many quantities and so many sites and times as the dynamic behavior of the system can be described completely by the principal components. Then, the second PCA, which is the mechanism hunting procedure, can be made, using the time derivatives of the observed quantities as the additional dimensions representing the system. Since the dynamic characteristics have been already described by the principal components in the first PCA, there should appear no additional principal components having significant eigenvalues in the second PCA, and the correlations of the time derivatives with the principal components hitherto obtained should reveal the mechanisms driving the dynamics of the system. In other words, the differential equations thus obtained could describe the mechanisms of the global warming, if the key quantities are sufficiently observed.

9 Concluding Remark

The earth environment forms a chaotic dynamical system in various time scales. Multi-dimensional representation is appropriate to describe the system, and the way of constructing such representation is discussed in detail. Then, the PCA is introduced to analyze the system. To reveal the cause of the global warming, the second PCA procedure is recommendable. The above scenario will work, if the data are complete. For the global change, data are numerous but mostly fragmentary. In that case, the adjusted data should be adopted for supplementing the incomplete data. The method (Unno and Yuasa, 1992) of determining the adjusted data is to maximize the probability given by equation (14) by adding an adjusted value of weight $(1 - w)$ for the observed value of weight w. In this way, all the fragmentary data that have been useless can be taken into the PCA constructively without violating the accuracy of the analysis. Thus, the utility of the PCA is also very much extended. Nowadays, data can be obtained rather easily and personal computers to analyze the data are also available. There will be a new science, if there is will. The present paper is intended to provide the method.

References

[1] Nemoto,J.,1992, *What happened at the end of this Century?* (in japanese), Chikuma Syobo, p.160.

[2] Takens,F., 1981, in D.A.Rand and L.-S.Young(eds.), *Lecture Note in Math.*, Springer-Verlag, Berlin, p.336.

[3] Unno,W. and Yuasa,M., 1992, *Astrophys.and Space Sci.*,**189**, 271.