atmosphere

Article

An ISOMAP Analysis of Sea Surface Temperature for the Classification and Detection of El Niño & La Niña Events

John Chien-Han Tseng

Central Weather Bureau, Taipei 100, Taiwan; jchtsenghome@gmail.com

Abstract: Isometric feature mapping (ISOMAP) is a nonlinear dimensionality reduction method used for extracting features from spatiotemporal data. The traditional principal component analysis (PCA), a linear dimensionality reduction method, measures the distance between two data points based on the Euclidean distance (line segment), which cannot reflect the actual distance between the data points in a nonlinear space. By contrast, the ISOMAP measures the distance between two data points based on the geodesic distance, which more closely reflects the actual distance by the view of tracing along the local linearity in the original nonlinear structure. Thus, ISOMAP-reconstructed data points can reflect the features of real structures and can be classified more accurately than traditional PCA-reconstructed data points. Moreover, these ISOMAP-reconstructed data points can be used for cluster analysis by emphasizing the differences among the points more than those by the traditional PCA. In this study, sea surface temperature (SST) data points reconstructed using the traditional PCA and ISOMAP were compared. The classification based on these reconstructed SST points was tested using the Niño 3.4 index, which labels El Niño, La Niña, or normal events. The mean differences from the ISOMAP data points were larger than those from the traditional PCA data points. The ISOMAP not only helped differentiate the points in two different events but also provided better difference measurement of the points belonging to the same class (e.g., 82/83 and 97/98 El Niño events). On examining the evolution of the leading three temporal eigen components of the SST PCA, or especially the SST ISOMAP, we found that the trajectories were similar to the Lorenz 63 model on a phase space figure. This implies that NWP perturbations can be traced using the ISOMAP to measure growing unstable behaviors. Spatial eigenmodes (empirical orthogonal function) between the traditional PCA and ISOMAP were also determined and compared herein.

Keywords: El Niño; EOF; ISOMAP; La Niña; Niño 3.4; PCA; SST

check for updates

Citation: Tseng, J.C.-H. An ISOMAP Analysis of Sea Surface Temperature for the Classification and Detection of El Niño & La Niña Events. *Atmosphere* **2022**, *13*, 919. https://doi.org/10.3390/atmos13060919

Academic Editors: Bo-Wen Shen, Roger A. Pielke Sr. and Xubin Zeng

Received: 27 April 2022
Accepted: 31 May 2022
Published: 6 June 2022

1. Introduction

Classification or clustering is performed to understand differences among events (e.g., numerical data, colors, and object shapes or figures). The success of the classification method depends on the effective clarification for describing, measuring, and recognizing these differences. For example, to analyze the high dimensional data takes time and computing cost to extract information. Traditionally, the analyzed high dimensional data are dimensional reduction to low dimension 2D or 3D points and the difference can be measured simply by the distance in the low-dimensional space. Certainly, a meaningful low-dimensional space that can appropriately extract features from data must be identified.

Principal component analysis (PCA) is a traditional linear dimensionality reduction method. Leading PCA components extract the main variances of the original data, which are called explained variances. Fewer the leading PCA components and more the explained variances, the better the PCA results. For example, if the original data dimensions are 10,000, then three leading PCA components can be used to explain 80% original variances and the PCA results are very good. By contrast, the PCA results are worse when an excess of 100 leading PCA components are used to explain only 50% original variances. Therefore, many PCA modifications and other alternative techniques [1–3] attempted to address this imperfect interpretation in real data analysis. For example, the modifications including: rotation PCA [4], probabilistic PCA [5–7], Bayesian PCA [5,6], and kernel PCA [8], are available. Other alternative techniques are the independent component analysis (ICA) and independent subspace analysis (ISA), which found the hidden possible factors behind of the physical phenomena based on source signals rather than prominent variances as PCA. The framework of ICA or ISA was built on non-Gaussian distributions and the assumption of composed linearly of source signals [9–11]. The ICA or ISA looked for components or subspaces which were the most statistically independent as possible under the view of non-Gaussian probability distributions. Better classification results can be obtained after retrieving low-dimensional PCA or ICA components. Low-dimensional data points from the PCA or ICA can be classified more easily than the original high-dimensional data points.

Tenebaum et al. [12] proposed iso metric feature mapping (ISOMAP) to solve the classification problem and obtain well-distributed low-dimensional data points. They pointed out that the traditional PCA considers the data linearly; for example, time evolution is resolved by the linear evolution of the original data arrangement. The geopotential height can be imagined to evolve in a month by daily data. The 30-times data being considered is constrained by the linear time variation. The Geopotential height does not evolve linearly in 1 month. However, despite this, the covariance matrix of the Geopotential height is counted linearly in the PCA. When these linear considerations are used to determine nonlinear variation, the data points cannot be discriminated and are sticky together by the view of low dimensional principal axes. Classifying concentrated or sticky together data points is difficult, and it leads to classification failure. The linear PCA separated points cannot represent the actual distances between the data points, leading to imperfect and false classification. In the ISOMAP, the original nonlinear relations in the data are built by establishing the nearest neighbors. The ISOMAP maintains the linearity of small domains but reflects nonlinear variations in the larger domain. In other words, this is called manifold consideration or manifold learning. That means the small local domain of the manifold is the homeomorphism as the Euclidean space. Dimensionality reduction through the ISOMAP reflects the real and nonlinear variations between the data points; in brief, data points can be pushed away more than that with the traditional linear PCA. Thus, classification can be effectively performed after ISOMAP dimensionality reduction.

In this study, we used sea surface temperature (SST) data to perform traditional El Niño classification. Classification results obtained through the traditional PCA and ISOMAP were compared. We presented the leading eigen components constructed using the PCA/ISOMAP as the space coordinate values over time, which depicted trajectories. On examining the data point trajectories after the ISOMAP, we found that the evolution exhibited the Lorenz 63 model on a phase space figure. The SST data points switched to different El Niño, normal, and La Niña events and spiral trajectories were never repeated. The ISOMAP SST point trajectories indicated that El Niño cycles like the chaotic behaviors as the phase space plot showing. Based on this indication, we believe that ISOMAP analysis can be a diagnostic tool for tracing the circulation evolution and evaluating ensemble perturbations, ensemble forecast spread, or even the numerical weather prediction (NWP) score between forecasts and analysis.

Some studies have reported that there are several types of El Niño/La Niña events (e.g., warm pool, dateline, Central Pacific, Eastern Pacific, and El Niño Modoki [pseudo El Niño] types) [13–15] and have differentiated these events by the geographical positions in which the SST anomaly centers/patterns are located. They combined PCA, correlation, regression-PCA methods, and different indices (Niño 1 + 2, Niño 3, Niño 4, and Niño 3.4) for an improved recognition of these different types of El Niño/La Niña events. These different events have varied connections with different weather/climate patterns. The improved recognition allows a more accurate forecast and greater resilience of their corresponding extreme weather events. However, the aforementioned methods are not straightforward and their results do not allow simple data visualization. We found the SST points constructed using the ISOMAP to reflect the differences in the types of El Niño/La Niña events in a simpler manner because these differences could be measured directly on the basis of the distance between the constructed SST points.

The remainder of this paper is organized as follows. In Section 2, we explain the definition of Niño 3.4 index for different SST events, the SST data, and the concept of ISOMAP. In Section 3, we show PCA/ISOMAP-reconstructed points and classification results. In the final section, we summarise major findings and the future applications regarding ISOMAP.

2. Data and Methods

The SST data used herein were obtained from version 5 of the NOAA NCDC ERSST (Extended Reconstructed global Sea Surface Temperature) data set, based on COADS data, collected from January 1980 to December 2021. To distinguish the El Niño, normal, or La Niña events are based on the Niño 3.4 (170° W–120° W, 5° S–5° N) index from NOAA's Climate Prediction Center. El Niño events were defined in Niño 3.4 region when the moving 3-month average SST anomaly exceeded 0.5 °C for at least 5 months. By contrast, La Niña events/anti-El Niño events were defined when the average SST anomaly was lower than 0.5 °C for 5 months. Determining the relationship between this index and the Pacific Ocean domain (120° E–60° W, 30° S–30° N) SST pattern would be interesting. Most studies on El Niño have focused on the large-scale circulation patterns and not on the Niño 3.4 index area.

The concept of ISOMAP is shown in Figure 1. The real data points are located in the warp surface, as shown in the arc curve in Figure 1. During PCA calculation, we assume that the variation (temporal or spatial) is linear and the relation between data points is like a short straight line. The PCA relation can be considered a type of Euclidean distance. However, the real distance between point c and point d is greater than the Euclidean distance. Thus, PCA always fails to present the real situation. To use a linear tool such as linear algebra eigen solutions, we must rearrange the relation, the distance, according to the longer straight line in Figure 1. The distance between point c and point d on the geodesic line truly reflects the distance in the real warp surface, the curve, in Figure 1. In brief, we consider the neighbor's situation and we do not take the 'shortcut' between the given data points.

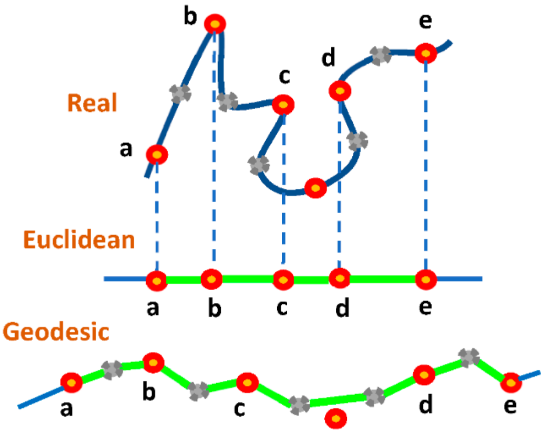


Figure 1. The concept of ISOMAP. The given data points are marked by a–e. Their neighbors are gray star points. The green segments are the measured distances of these data points.

Tenebaum et al. [12] proposed to plot the nearest neighbor graph that is used to reflect the distance on the warp surface. This means the distance between point c and point d is not calculated using the Euclidean distance, the shortcut straight distance method, but by including the other three points. Under geodesic framework, after considering the neighbors, the shortest distance is decided again and one red circle point is bypass to the distance counting. The geodesic distance is calculated on the basis of this neighbor graph, and then, this distance relation is used to form the covariance matrix that can be solved by the traditional PCA or multi-dimensional scaling (MDS) method [16]. The ISOMAP (Algorithm 1) is shown below [17–19].

Algorithm 1: ISOMAP

```
Step 1. construct the matrix of squared pairwise similarities $\mathbf{D}$, $\mathbf{D}_{i\,j} = \left|\left|\boldsymbol{y}_i - \boldsymbol{y}_j\right|\right|^2$, the distance matrix, measured on temporal dimension. The $i, j$ are temporal indexes.
Step 2. build the weighted graph based on the $\mathbf{D}$ according to how many neighbors of each point
Step 3. estimate the geodesic distances $\mathbf{D}_G$ by finding the shortest paths on the weighted graph (Dijkstra's algorithm [20])
Step 4. define $\mathbf{B} = -1/2\,\mathbf{J}\,\mathbf{D}_G\,\mathbf{J}^{\mathbf{T}}$, where $\mathbf{J} = \mathbf{I} - 1/N$, I is the identity matrix, and $N$ is the number of data points
Step 5. solve eigen problem $\mathbf{BP} = \mathbf{P}\Sigma$
Step 6. computing the leading principal vectors by $\mathbf{X} = \mathbf{P}\Sigma^{1/2}$
```

To measure the ISOMAP low dimensionality sufficient criterium is the residual variance [12], which is defined as follows:

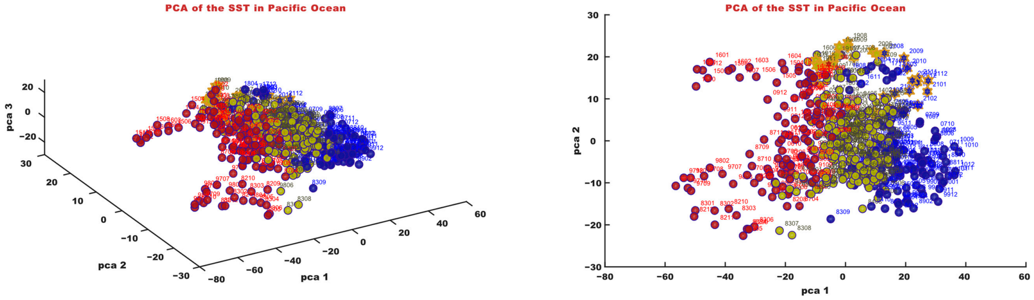
$$1 - R^2(\mathbf{D}_M\,,\ \mathbf{D}_G).$$

$\mathbf{D}_G$ is the geodesic distance matrix used in steps 3 and 4 in Algorithm 1 for solving the eigen problem, $\mathbf{D}_M$ is the geodesic distance matrix reconstructed from the ISOMAP, the low-dimensional eigen (principal) components (leading modes of step 6 in Algorithm 1), and R is the coefficient of correlation. The larger the correlation between $\mathbf{D}_M$ and $\mathbf{D}_G$ is, the lower the residual variance is, and low-dimensional components can be used to approximate the original high-dimensional structure [12,18]. In this paper, we constructed the distance matrix based on the temporal dimension. We assumed that the time variation of the SST is not linear. The detailed arrangements of the matrices used are given in the Appendix A.

The classifier used in this study is the smooth support vector machine (SSVM) that replaces the plus function in the non-smooth SVM by a smooth function [21]. All test results presented in the next paragraphs were obtained from 20-times 5-fold cross-validation. This means that 80% data were randomly selected as the training set and the remaining 20% data formed the testing set in each validation. Then, the average training errors/testing errors of this 20-times validation were calculated.

3. Classification

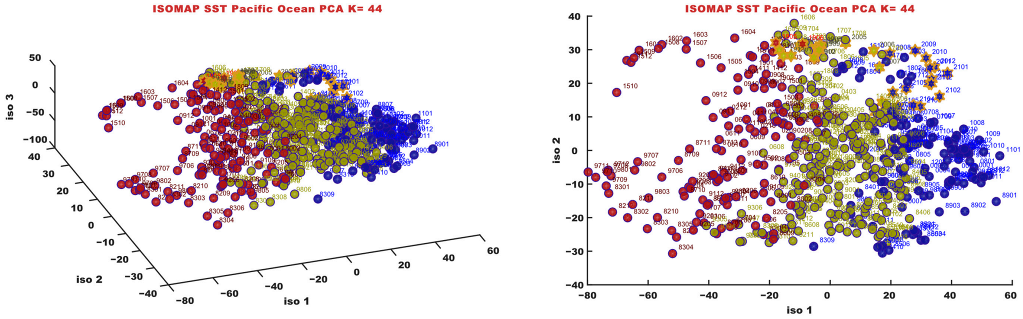
Three leading temporal eigenvectors from the PCA and ISOMAP were used for presenting points reconstructed after dimensionality reduction (step 6 in the Algorithm 1). The leading 20 eigenvectors were used for testing the classification results. Figure 2a presents the 3D structure of the PCA-reconstructed points, and Figure 2b shows the 2D structure of the PCA-reconstructed points. The El Niño events were labeled red, normal events were marked yellow, and La Niña events were labeled blue. In Figure 2, most El Niño, normal, and La Niña events were already well separated. This means that meteorologists are correctly using the Niño 3.4 index to define the El Niño event. Inevitably, some points were stuck together, which probably led to the classifier failure. Under this situation, if the sufficient distance between the PCA data points can be obtained, the better is for the classification. As mentioned in the previous section, the solution can be the ISOMAP that efficiently separates the points by measuring the geodesic distance, thus reflecting the actual space distance variation between the data points.



(**a**) (**b**)

Figure 2. (**a**) 3D structure composed of the three leading PCA components. (**b**) 2D structure composed of the two leading PCA components.

During the application of the ISOMAP, the number of neighbors should be determined first, which was done by evaluating the residual variance. In this study, the number of neighbors for 41 years of SST data was 44. Similar to the PCA calculation, the leading three eigen components of the ISOMAP are presented in Figure 3a. The ISOMAP points were indeed more separated than the PCA points. Some events were significantly different from others even if they belonged to the same class (El Niño: 82/83, 97/98, 15/16; La Niña: 84/85, 88/89, 98/99). The 3D structure of the ISOMAP-reconstructed data points exhibited more space variations, with the data points being well separated, and allowed grouping of the data points into different clusters. Figure 3b is the 2D structure of Figure 3a and shows well-separated data points compared with those in Figure 2b (PCA calculation).



(**a**) (**b**)

Figure 3. (**a**) 3D structure composed of the three leading ISOMAP components. (**b**) 2D structure composed of the two leading ISOMAP components.

The data points for the recent 31 months were marked by a star sign outside the circle of points. If the center of Figures 2 and 3 represented the climatology mean, the recent 31-month trajectory was far from the center. This indicates that the recent SST data evolved differently from the previous SST data. Using the past data to describe the recent 31-month variation would be difficult. To highlight this situation, we faded the data points before the recent 31-month period but made those for the recent 31 months clear (Figure 4). We found the trajectory to be like number 8, the circular shape. On using animation to demonstrate these 41-year SST data points, we could find the trajectories in a circular motion, similar to the Lorenz 63 model [22] on a phase space figure. The SST points swung between the El Niño, normal, and La Niña events and would not repeat their passages. These trajectory behaviors corresponded to those reported in previous studies [23,24], which indicated that the change in ENSO events is a low-order chaotic system.

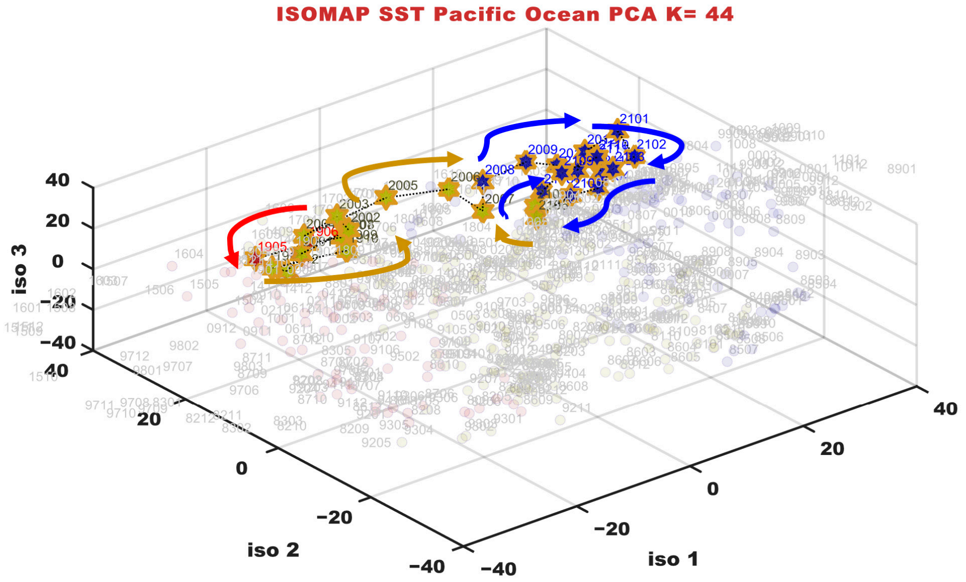


Figure 4. 3D structure composed of the three leading ISOMAP components but only has the recent 31-month data points highlighted. The curves marked by arrow s depict the time evolution (May 2019 to December 2021) of these points. The trajectory shows the end of last El Niño (red arrows), circling in normal events (yellow arrows), dramatically changing to be La Niña (blue arrows) with a bigger circle route then back to two-month normal events, and to be La Niña again.

One spatial eigenmode and its corresponding temporal eigenmode acted as a pair in the PCA or ISOMAP calculation. The original data matrix could simply be multiplied with the temporal mode to obtain the spatial mode and vice versa. After the examination of the temporal modes, the leading three spatial eigenmodes, the empirical orthogonal function (EOF), were presented in Figures 5 and 6 (from the PCA and ISOMAP respectively). The number in the top legend of all panels of Figures 5 and 6 is the ratio of the explained variance. With the three leading eigenmodes, the explained variances of the ISOMAP were not as good as those of the PCA. The first eigenmodes of both methods were similar, but the second and third eigenmodes were slightly different. These small differences in the three leading eigenmodes led to obvious differences in the distribution of the corresponding temporal points. In fact, no obvious differences were observed between the PCA and ISOMAP in the first 20 leading spatial eigenmodes (not shown). Slight rearrangement by using the ISOMAP can help separate the data points, allowing their grouping into different clusters.

The residual variances mentioned in the previous section reflect the similarity between the original covariance and low-dimensional covariance. Residual variances of the ISOMAP with the different nearest neighbor numbers are presented in Figure 7. Residual variances became smaller as the neighbor number increased. However, the lower residual variance values did not mean that the nearest neighbor number in ISOMAP was considerably better. It depends on what kind of purpose the ISOMAP needs to achieve. If we want to do classification, we cannot choose the largest nearest neighbor number. Basically, when we connected all the points together in the ISOMAP, the matrix used to solve eigenmodes was similar to the covariance matrix in the PCA. The method connected all points is actually the MDS. When the distance is measured in Euclidean space, the MDS is identical to the PCA [18]. The PCA-reconstructed points in Figure 2 were not distributed well for classification. This is also the reason for not considering all neighbors to count the ISOMAP because the temporal data points are all connected to each other, and thus, the advantage of the ISOMAP is lost. On the other hand, the explained variances from the ISOMAP were not good, but their residual variances were sufficiently lower to reconstruct low-dimensional structures approximate to the high-dimensional structures. The advantage of ISOMAP gave the reasonable distance estimations between the disconnected points through the Dijkstra's algorithm, which decided the shortest path through the neighbors and visited all the points. In brief, the disconnected two points were connected by their neighbors, while preferably no single point is isolated. It also pointed out that try to avoid using too small nearest neighbor number to prevent some points from being isolated.

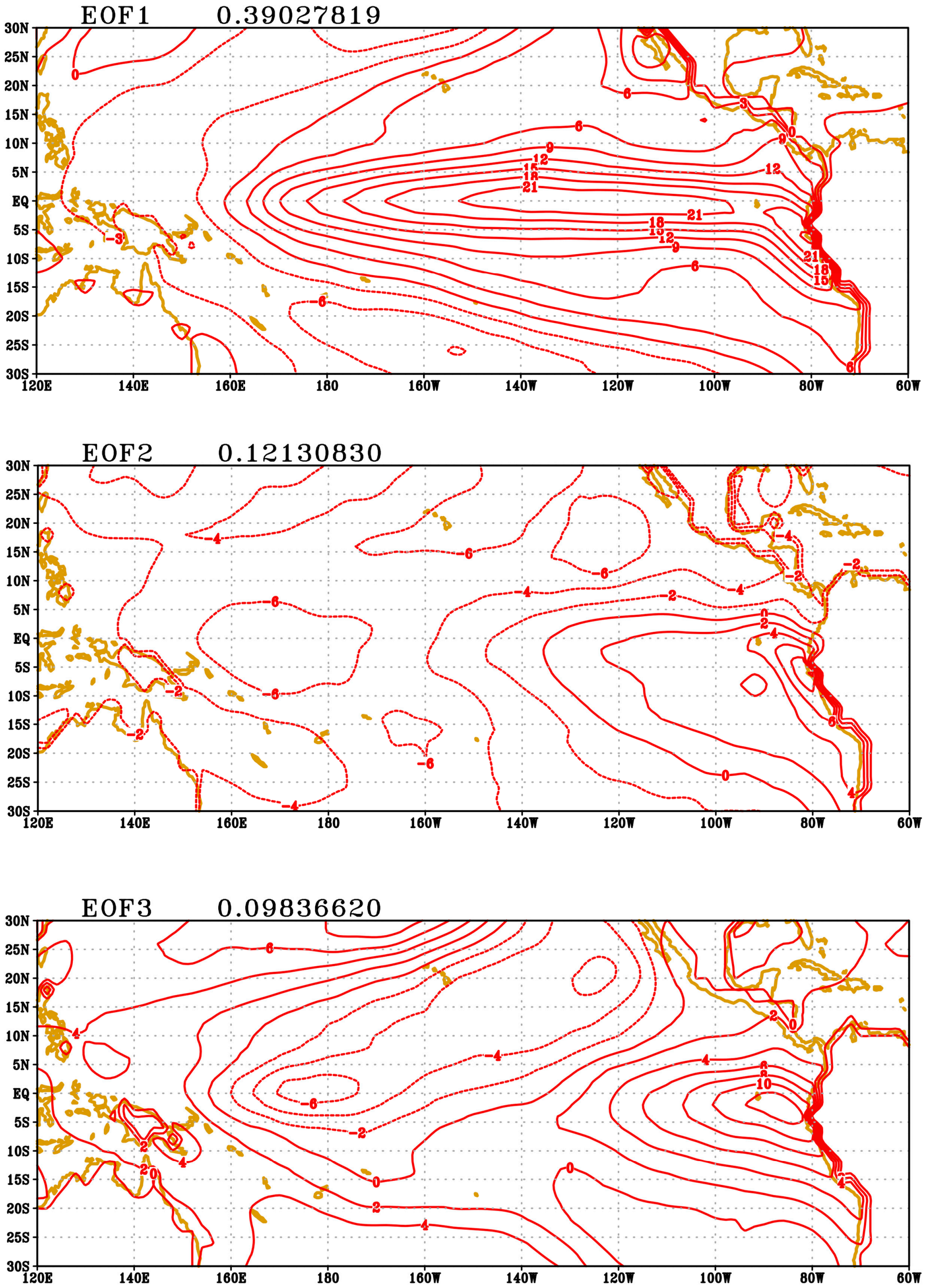
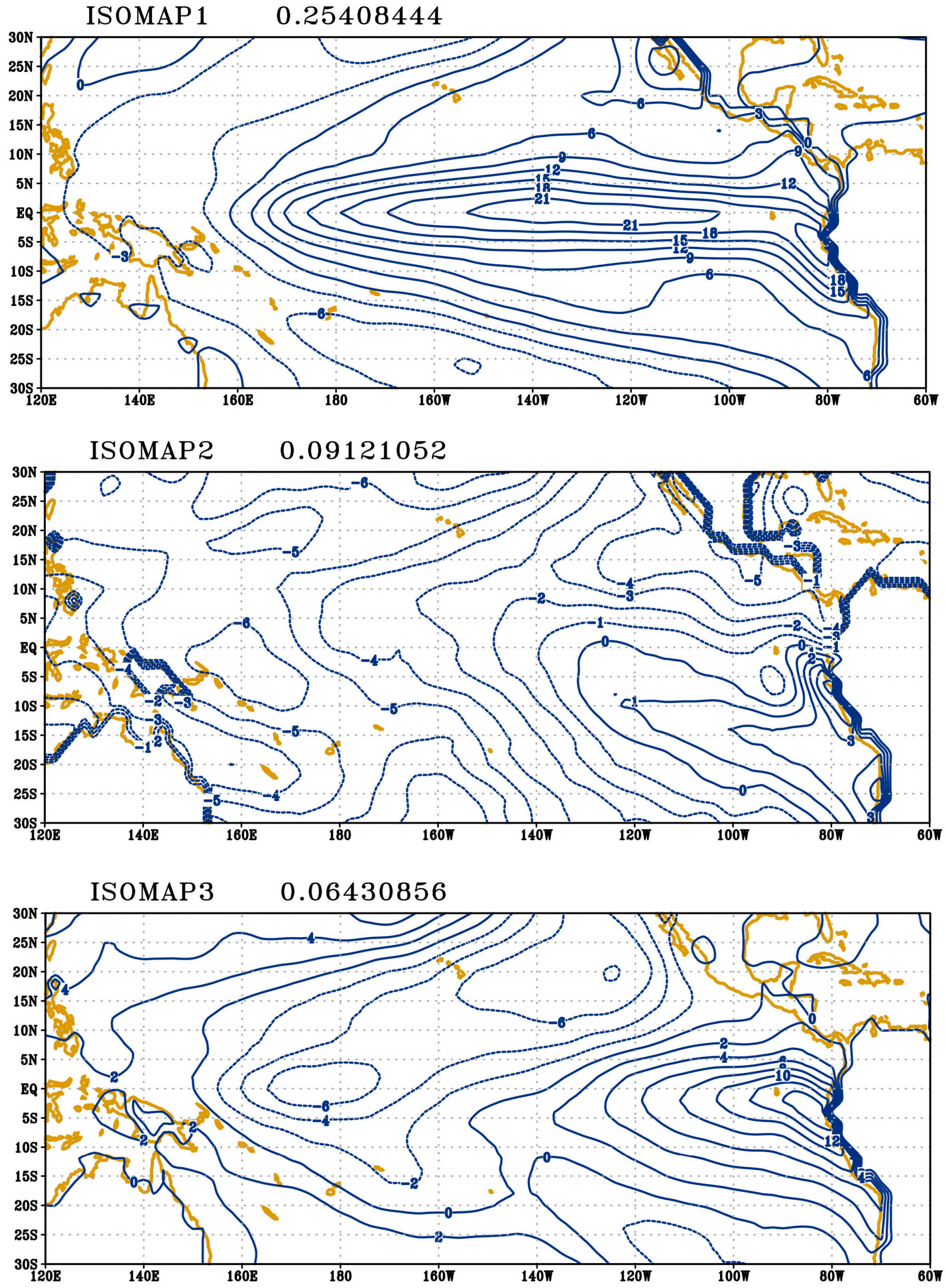


Figure 5. The first three spatial eigen modes (EOFs) of the PCA. The number in the top legend of every figure is the ratio of the explained variance.


ISOMAP1 0.25408444

ISOMAP2 0.09121052

ISOMAP3 0.06430856

Figure 6. The first three spatial eigen modes (EOFs) of the ISOMAP. The number in the top legend of every figure is the ratio of the explained variance.

The first 20 leading components were considered as data points for classification, and SSVM results are presented in Table 1 (the PCA) and Table 2 (the ISOMAP). We performed 20-times 5-fold cross-validation to obtain these results. The ISOMAP results were slightly better than the PCA results. We defined two classes problem as El Niño/non El Niño or La Niña/non La Niña, because the simple SVM was the 2-class classifier. To classify El Niño/non El Niño was little easier than to classify La Niña/non La Niña. The clues were already in Figures 2 and 3, because the blue La Niña points were closer each other and more difficult to distinguish. No special reason existed for taking the 20 dimensionalities to form data points for the classification. We tested from 1–20 dimensionalities for classification, and the accuracy was approximately 90% with 2–20 dimensionalities, whereas it was approximately 84% with 1 dimensionality. Removal of the normal events from the testing would have led to considerably better classification results and 99% accuracy. All these test results support that the Niño 3.4 index is a good index for defining ENSO events.

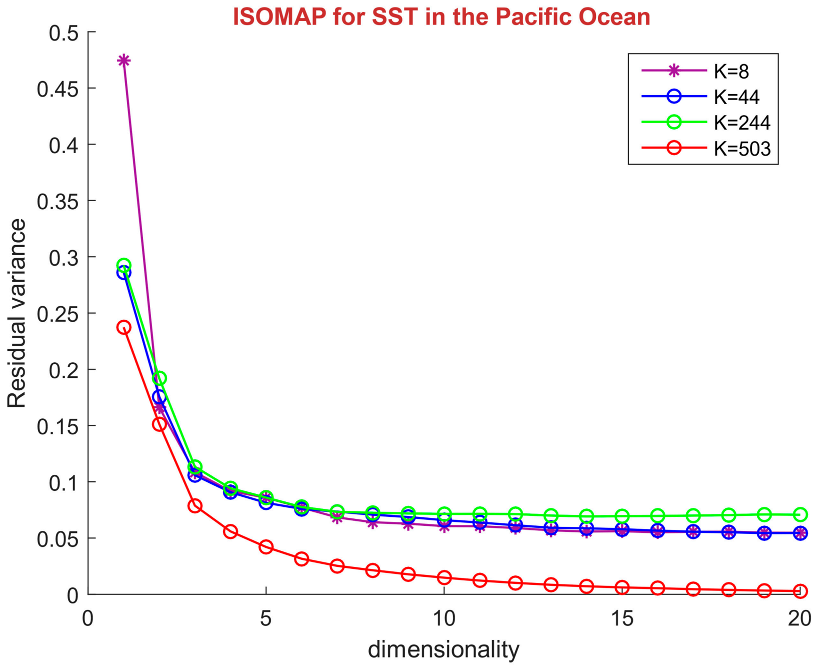


Figure 7. ISOMAP residual variances. The different nearest neighbor numbers in ISOMAP calculation are shown in different colors.

Table 1. The training errors and the testing errors from the 20-dimensionality PCA with SVM.

PCA with SVM	Training Error	Testing Error
El Niño and non El Niño	0.0525	0.1115
La Niña and non La Niña	0.0525	0.1001

Table 2. The training errors and the testing errors from the 20-dimensionality ISOMAP with SVM.

ISOMAP (44 *) with SVM	Training Error	Testing Error
El Niño and non El Niño	0.0385	0.0922
La Niña and non La Niña	0.0545	0.0801

* 44 is the number of neighbors.

4. Conclusions and Discussion

The ISOMAP could help identify extreme El Niño events and easily measure the differences between any two events from the reconstructed space and data points. The distances far or close in the ISOMAP-reconstructed space provide the measurement of the similarities between data points and were more accurate than the distances in the traditional PCA-reconstructed space. The ISOMAP residual variances provided the reference values to decide whether the number of lower dimensionalities was sufficient for the classification. The ISOMAP results could be easily used to perform clustering. Moreover, the ISOMAP allowed grouping of some events or clarified why two events belonging to the El Niño class were different. Although studies have indicated that no two El Niño events are identical, measuring the extent of differences between two events in the same class was possible because of the ISOMAP tool. Meanwhile, we are also proceeding the test that if it is possible to define or predict the ENSO event through SSVM with ISOMAP-reconstructed points instead of the Niño 3.4 index.

Besides solving the El Niño problem, the ISOMAP method can also be used to perform composite analysis when similar cases are to be selected among other meteorological problems. The ISOMAP can be used as a tool for diagnosing different atmospheric circulations. Of course, the ISOMAP can be a score measurement for evaluating the NWP outputs and observation. For example, the SST NWP model output can be projected on the leading components of the ISOMAP SST observational data and some forecasts with a high probability to be true can be analyzed. Moreover, the same procedure can be used to trace the NWP perturbations and detect the growing unstable behaviors. The ISOMAP trajectory behaviors showing in Figure 4 are only similar to Lorenz 63 model trajectory in shapes. It is worth doing more researches in ISOMAP trajectory behavior to check the profound meaning of sensitivity to initial conditions in Lorenz 63 model or other NWP models.

In this study, the number of nearest neighbors was 44 for the ISOMAP calculation. In fact, we tested the neighbor number from 8 to 60, and the number 44 was selected as it could provide the best SSVM classification results. We used the monthly SST data and do not know whether the selected neighbor number is related to the El Niño period of 2–10 years. With a neighbor number greater than 480, the classification results are similar to those of the PCA method. Actually, if we use Euclidean space to measure the distance and take all points connected (any one point connecting to others), the ISOMAP is degenerated to MDS which is identical to the PCA [18]. The reconstructed data points are closer to each other (same as Figure 2), and the advantage of using the ISOMAP is lost.

The traditional PCA is sensitive to the counting domain chosen. If we are concerned about the ENSO, the SST in Pacific Ocean, we should perform a PCA calculation in the tropical Pacific region. If the global domain SST is used to calculate the PCA, global EOF structures that are difficult to explain are obtained. However, the ISOMAP can maintain the local structures even it is not wise to take this kind of calculation. Because the PCA or ISOMAP extracts the counting domain maximum variance patterns out, the bigger domain contains the irrelevant pattern with Pacific Ocean SST pattern. The ISOMAP builds the nearest neighbors first and it has the limitation of the irrelevant pattern.

During this study, the SST ISOMAP-reconstructed point returned to the cluster of the last La Niña event since August 2021. At that time, according to the Niño 3.4 index, La Niña was defined as 3-month running mean SST anomalies lower than 0.5 °C for at least 5 consecutive months. Thus, NOAA declared the recent La Niña event until January 2022. Are there any methods to determine the ISOMAP point trajectory in the moment of August 2021 toward the La Niña event or back to the normal year? The ISOMAP-reconstructed points from the NWP ensemble results can probably predict the SST after the moment of August 2021 would be a La Niña event.

Funding: This research is grateful for support from the Ministry of Science and Technology, Taiwan, ROC (grant # MOST 110-2634-F-008-008) and Central Weather Bureau, Taiwan, ROC.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The SST data are from IRI web, and the link is https://iridl.ldeo.columbia.edu/SOURCES/.NOAA/.NCDC/.ERSST/ (accessed on 27 April 2022).

Acknowledgments: The author thanks three anonymous reviewers for their constructive and helpful comments. The author appreciates reviewers for pointing out some mistakes in the original manuscript. During this research, the author thanks Bo-Wen Shen for his guidance and encouragement.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

The data matrix is $\mathbf{Y}_{KN}$. It can be formed by the SST anomaly, that is, the SST value subtracting its climatological mean. Moreover, M is the dimension of the space, and N is the dimension of time. During the PCA calculation, we solved the eigen problem of

$\mathbf{Y}^T\mathbf{Y}$ or $\mathbf{Y}\mathbf{Y}^T$. In meteorology, temporal eigen components are called principal components, and spatial eigen components are called empirical orthogonal functions. $\mathbf{Y}^T\mathbf{Y}$ or $\mathbf{Y}\mathbf{Y}^T$ is the covariance matrix of the data $\mathbf{Y}$. However, we can count the distance between any two temporal points (pairwise) and take the sum of all spatial points to build the square distance matrix $\mathbf{D}_{NN}$. This distance matrix can also be used for counting the eigen problem. This method is often called multidimensional scaling.

The detailed arrangements are as follows:

(1) The covariance matrix for the PCA is

$$\mathbf{Y}^T\mathbf{Y} = \mathbf{C}_{NN} = \begin{pmatrix} \sum\limits_{k=1}^{K} y_{k,\ t=1}^T y_{k,\ t=1} & \sum\limits_{k=1}^{K} y_{k,\ t=1}^T y_{k,\ t=2} & \cdots & \sum\limits_{k=1}^{K} y_{k,\ t=1}^T y_{k,\ t=N} \\ \sum\limits_{k=1}^{K} y_{k,\ t=2}^T y_{k,\ t=1} & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \sum\limits_{k=1}^{K} y_{k,\ t=N}^T y_{k,\ t=N} \end{pmatrix}$$

(2) The distance matrix for MDS is

$$\mathbf{D} = \begin{pmatrix} \sum\limits_{k=1}^{K} (y_{k,\ t=1} - y_{k,\ t=1})^2 & \sum\limits_{k=1}^{K} (y_{k,\ t=1} - y_{k,\ t=2})^2 & \cdots & \sum\limits_{k=1}^{K} (y_{k,\ t=1} - y_{k,\ t=N})^2 \\ \sum\limits_{k=1}^{K} (y_{k,\ t=2} - y_{k,\ t=1})^2 & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \sum\limits_{k=1}^{K} (y_{k,\ t=N} - y_{k,\ t=N})^2 \end{pmatrix}$$

We can choose the number of neighbors (e.g., 44 neighbors in this study) required to plot the weighted graph and construct the geodesic distance matrix. This geodesic distance matrix can be implemented using Dijkstra's algorithm [20], one of the shortest path algorithms. This means that the direct distance (shortcut) between number 1 and number 45 is replaced by the shortest path through the 44 neighbors to neighbor 45. The new geodesic distance matrix is called $\mathbf{D}_G$ in this article. The ISOMAP method first establishes the weighted graph and constructs the geodesic distance matrix and then solves the MDS problem.

The data point coordinate is calculated as the PCA/ISOMAP temporal principal component multiplied by its square root of the eigenvalue (step 6 in Algorithm 1). We employ subscript M to represent the number of eigen components used. For example, we take the first three leading components to form the distance matrix $\mathbf{D}_M$ as follows:

$$\mathbf{D}_M = \begin{pmatrix} d_{M(t=1,\ t=1)} & \cdots & \cdots & d_{M(t=1,\ t=N)} \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ d_{M(t=N,\ t=1)} & \cdots & \cdots & d_{M(t=N,\ t=N)} \end{pmatrix}$$

The element of $\mathbf{D}_M$, the distance between any two-time interval steps is defined as follows:

$$d_{M(t,\ t)} = \sum_{i=1}^{M} (\sigma_i^{1/2} pca_{i\ (t,\ t)} - \sigma_i^{1/2} pca_{i\ (t,\ t)})^2$$

where *pca* is the temporal eigenmode from the PCA/ISOMAP (steps 4 and 5 in the Algorithm 1), and σ is the corresponding eigenvalue.

The coefficient of correlation $R^2(\mathbf{D}_M, \mathbf{D}_G)$ is calculated using all the matrix elements of $\mathbf{D}_M$ and $\mathbf{D}_G$ as follows:

$$R^2 = \frac{\left(\sum (d_M - \overline{d}_M)(d_G - \overline{d}_G)\right)^2}{\sum (d_M - \overline{d}_M)^2 \sum (d_G - \overline{d}_G)^2}$$

Then, the residual variance is defined as

$$1 - R^2(\mathbf{D}_M, \mathbf{D}_G)$$

We can count $M = 1, 2, \ldots, 20$ consecutively to obtain residual variances of the leading 20 eigen components. In PCA's residual variance calculation, we use the $\mathbf{D}$ instead of $\mathbf{D}_G$.

References

1. Alpaydin, E. *Introduction to Machine Learning*, 3rd ed.; MIT Press: Cambridge, MA, USA, 2014; p. 640.
2. Bishop, C.M. *Pattern recognition and machine learning*; Springer-Verlag Press: New York, NY, USA, 2006; p. 738.
3. Hsieh, W.W. *Machine learning methods in the environmental sciences: Neural networks and kernels*; Cambridge university press: New York, NY, USA, 2009; p. 349.
4. Richman, M.B. Rotation of principal components. *J. Climatol.* **1986**, *6*, 293–335. [CrossRef]
5. Tipping, M.E.; Bishop, C.M. Probabilistic principal component analysis. In *Technical Report, NCRG/97/010*; Neural Computing Research Group, Aston University: Birmingham, UK, 1997; p. 13.
6. Tipping, M.E.; Bishop, C.M. Mixtures of probabilistic principal component analyzers. *Neural Comput.* **1999**, *11*, 443–482. [CrossRef] [PubMed]
7. Roweis, S. EM algorithms for PCA and SPCA. In *Advances in Neural Information Processing Systems*; Jordan, M.I., Kearns, M.J., Solla, S.A., Eds.; MIT Press: Cambridge, MA, USA, 1998; Volume 10, pp. 626–632, 1107.
8. Schölkopf, B.; Smola, A.; Müller, K.-R. Nonlinear component analysis as a kernel eigenvalue problem. *Neural Comput.* **1998**, *10*, 1299–1319. [CrossRef]
9. Hyvärinen, A.; Oja, E.E. Independent component analysis: Algorithms and applications. *Neural Netw.* **2000**, *13*, 411–430. [CrossRef]
10. Hannachi, A.; Unkel, S.; Trendafilov, N.T.; Jolliffe, I.T. Independent component analysis of climate data: A new look at EOF rotation. *J. Clim.* **2009**, *22*, 2797–2812. [CrossRef]
11. Pires, C.; Hannachi, A. Independent subspace analysis of the sea surface temperature variability: Non-Gaussian sources and sensitivity to sampling and dimensionality. *Complexity* **2017**, *2017*, 1–23. [CrossRef]
12. Tenebaum, J.B.; Silva, V.D.; Langford, J.C. A global geometric framework for nonlinear dimensionality reduction. *Science* **2000**, *290*, 2319–2323. [CrossRef]
13. Kao, H.-Y.; Yu, J.-Y. Contrasting Eastern-Pacific and Central-Pacific types of ENSO. *J. Clim.* **2009**, *22*, 615–632. [CrossRef]
14. Kug, J.-S.; Jin, F.-F.; An, S.-I. Two types of El Niño and warm pool El Niño. *J. Clim.* **2009**, *22*, 1499–1515. [CrossRef]
15. Ashok, K.; Behera, S.K.; Weng, H.; Yamagata, T. El Niño Modoki and its possible teleconnection. *J. Geophys. Res.* **2007**, *112*, C11007. [CrossRef]
16. Cox, T.F.; Cox, M.A.A. *Multidimensional Scaling*; CRC press: Boca Raton, FL, USA, 2001; p. 328.
17. Tripathy, B.K.; Sundareswaren, A.; Ghela, S. *Unsupervised learning approaches for dimensionality reduction and data visualization*; CRC Press: Boca Raton, FL, USA, 2021; p. 160.
18. Marsland, S. *Machine Learning: An Algorithmic Perspective*, 2nd ed.; CRC Press: Boca Raton, FL, USA, 2014; p. 457.
19. Hannachi, A.; Turner, A.G. ISOMAP nonlinear dimensionality reduction and bimodality of Asian monsoon convection. *Geophys. Res. Lett.* **2013**, *40*, 1653–1658. [CrossRef]
20. Neapolitan, R.E. *Foundations of Algorithms*, 5th ed.; Jones & Bartlett Learning Press: Burlington, MA, USA, 2015; p. 676.
21. Lee, Y.-J.; Mangasarian, O.L. 2000: SSVM: A smooth support vector machine for classification. *Comput. Optim. Appl.* **2000**, *20*, 5–22. [CrossRef]
22. Lorenz, E.N. Deterministic nonperiod flow. *J. Atmos. Sci.* **1963**, *20*, 130–141. [CrossRef]
23. Tziperman, E.; Cane, M.A.; Zebiak, S.E. Irregularity and locking to the seasonal cycle in an ENSO prediction model as explained by the quasi-periodicity route to chaos. *J. Atmos. Sci.* **1995**, *52*, 293–306. [CrossRef]
24. Wang, B.; Fang, Z. Chaotic oscillations of the tropical climate: A dynamic theory for ENSO. *J. Atmos. Sci.* **1996**, *53*, 2786–2802. [CrossRef]