# Tropical Indian Ocean Variability Revealed by Self-Organizing Maps

**Tomoki Tozuka[1] Jing-Jia Luo[2], Sebastien Masson[3] and Toshio Yamagata[1,2]**

*[1] Department of Earth and Planetary Science, Graduate School of Science, The University of Tokyo, Tokyo, Japan*

*[2] Frontier Research Center for Global Change/JAMSTEC, Yokohama, Kanagawa, Japan*

*[3] LOCEAN, France*

Climate Dynamics
August 2008, Volume 31, Issue 2-3, pp 333-343

**The final publication is available at:**
**http://link.springer.com/article/10.1007%2Fs00382-007-0356-4**

---

*Corresponding author address: Tomoki Tozuka, Department of Earth and Planetary Science, Graduate School of Science, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan. E-mail: tozuka@eps.s.u-tokyo.ac.jp*

**Abstract**

The tropical Indian Ocean climate variability is investigated using an artificial neural network analysis called Self-Organizing Map (SOM) for both observational data and coupled model outputs. The SOM successfully captures the dipole sea surface temperature anomaly (SSTA) pattern associated with the Indian Ocean Dipole (IOD) and basin-wide warming/cooling associated with ENSO. The dipole SSTA pattern appears only in boreal summer and fall, whereas the basin-wide warming/cooling appears mostly in boreal winter and spring owing to the phase-locking nature of these phenomena. Their occurrence also undergoes significant decadal variation.

Composite diagrams constructed for nodes in the SOM array based on the simulated SSTA reveal interesting features. For the nodes with the basin-wide warming, a strong positive SSTA in the eastern equatorial Pacific, a negative Southern Oscillation, and a negative precipitation anomaly in East Africa are found. The nodes with the positive IOD are associated with a weak positive SSTA in the central equatorial Pacific or positive SSTA in the eastern equatorial Pacific, a positive (negative) sea level pressure anomaly in the eastern (western) tropical Indian Ocean, and a positive precipitation anomaly over East Africa. The warming in the central equatorial Pacific appears to correspond to El Niño Modoki discussed recently. These results suggest usefulness of SOM in studying large-scale ocean-atmosphere coupled phenomena.

**Keywords:** Indian Ocean Dipole, El Niño-Southern Oscillation, Self-Organizing Map, decadal variability, seasonal phase-locking

## 1. Introduction

The tropical Indian Ocean undergoes significant interannual variations owing to large-scale air-sea interaction. El Niño in the Pacific induces basin-wide warming in the tropical Indian Ocean (Klein et al. 1999), whereas the Indian Ocean Dipole (IOD; Saji et al. 1999, Webster et al. 1999), an intrinsic air-sea coupled mode of the tropical Indian Ocean, causes warm (cold) sea surface temperature anomaly (SSTA) to the west (east). These variations in the tropical Indian Ocean not only cause modulations in the monsoons (Annamalai and Murtugudde 2004; Vecchi and Harrison 2004) and the East African short rains (Black et al. 2003; Behera et al. 2005), but also influence extra-tropical climate variability (Guan and Yamagata 2003; Saji and Yamagata 2003). Furthermore, the decadal variation in the IOD causes the decadal variation in the Southern Oscillation by influencing the sea level pressure in the Indo-Pacific domain (Behera and Yamagata 2003) and the relationship with the Indian Summer Monsoon Rainfall and the El Niño/Southern Oscillation (ENSO) as discussed by Ashok et al. (2001). This is why many studies have recently focused on the tropical Indian Ocean (see Yamagata et al. 2004 and Chang et al. 2006 for reviews).

Above studies, however, relied on linear analyses such as composite analysis, correlation analysis, and various kinds of empirical orthogonal function (EOF). Actually, the IOD's existence and dependence on ENSO were discussed on the basis of linear statistics (Allan et al. 2001; Dommenget and Latif 2002; Behera et al. 2003). Since the coupled air-sea phenomena are inherently nonlinear, it is interesting to investigate the variability using a non-linear analysis method. One such method is the self-organizing

map (SOM) (Kohonen 2001), which is only recently introduced to the climate research community (e.g. Richardson et al. 2003; Cheng and Wilson 2006; Leloup et al. 2007a, b). Although non-linear analyses have been applied to ENSO studies (e.g. Wu and Hsieh 2003; An et al. 2006), only Collins et al. (2004) have adopted a nonlinear canonical correlation analysis in the tropical Indian Ocean. However, it was for an empirical model that assesses predictability of SST and SLP over the tropical Indian Ocean.

We here investigate the tropical Indian Ocean climate variability using the SOM for the first time and examine its usefulness in climate research. The present paper is organized as follows. A brief description of the data and the SOM is given in the next section. In section 3, the observed and simulated tropical Indian Ocean climate variability is analyzed using the SOM. In particular, a detailed discussion on the seasonal phase-locking and decadal variations as well as difference in the atmospheric response to the SSTA is given there. The final section summarizes the main results.

## 2. Data and Method

### *2.1 Data and model*

The SST data used in the present study is the HadISST. It is monthly dataset of 1° x 1° horizontal resolution from 1951 to 1999 and is edited by Rayner et al. (2003).

The model data is from a coupled atmosphere-ocean-land general circulation model (GCM) run on the Earth Simulator of Japan Agency for Marine-Earth Science and Technology. The model is called SINTEX-F1 model (Luo et al. 2003; Tozuka et al.

2005), which is an upgraded version of the SINTEX model (Gualdi et al. 2003). The atmospheric component T106L19 ECHAM-4 (Roeckner et al. 1996) is coupled to the oceanic component OPA-8.2 (Madec et al. 1998) using the coupler OASIS 2.4 (Valcke et al. 2000). No measures for flux adjustments are taken in the model. The horizontal resolution of the ocean general circulation model (OGCM) is 2° x 2° with an increased meridional resolution up to 0.5° near the equator. There are 31 levels in the vertical with 19 levels in the upper 400 m. The monthly mean output from the last 200 years of the total 220 years model integration is used for the analysis. Details of the coupled general circulation model (CGCM) can be found in the above references and readers are referred to Behera et al. (2003) and Tozuka et al. (2007) for the model performance in the tropical Indian Ocean. The long time-series of the model data also provides us with the necessary statistical confidence for composite analyses.

*2.2 Method: SOM*

The SOM is an unsupervised artificial neural network and shown to be an effective method for feature extraction and classification. We here used a software package called SOM_PAK 3.1 (Kohonen et al. 1995), and readers are referred to Kohonen (1982, 2001) for more details about the SOM and Leloup et al. (2007a, b) or Richardson et al. (2003) for its application in climate. Here, we give a brief introduction to this method; Fig. 1 is the schematic diagram showing how to apply this method.

We have first prepared input data from both observed and simulated SSTA by calculating average SSTA in 10° (longitude) x 5° (latitude) boxes between 17.5°S and

17.5°N in the tropical Indian Ocean after removing a linear trend using a least-square fit. As a result, the input matrix consisted of 43 grid points and 588 (2400) months of data for observation (model) and 2988 months of data for a combined analysis. Then, the dimension of the two-dimensional SOM array is decided to be 7 x 7 (total of 49 nodes) for the observation and 9 x 9 (total of 81 nodes) for the combined analysis. Each node in the SOM array is associated with a reference vector with dimensions equal to that of the input vector, i.e. 43 in the present study. The topology or lattice structure of the array can be either hexagonal or rectangular, and we have decided to use a rectangular map. Results are not sensitive to this topology.

We have initialized the reference vectors to random values that are evenly distributed in the area of corresponding data vector components. Our qualitative results are not sensitive to the way of initialization. At each step in the training process, a reference vector with the smallest Euclidean distance with the input vector is selected as the "winner" node:

$$c_n = \arg\min \left\| \mathbf{x}_n - \mathbf{m}_{ij} \right\| \quad ,$$

where $c_n$ is an index of the "winner" on the SOM array, $\mathbf{m}_{ij}$ is the reference vector, $\mathbf{x}_n$ denotes the present input vector, and the "arg" denotes index. This "winner" node becomes the center of an update neighborhood, within which reference vectors are restored to the input vector. The "winner" and neighborhood nodes are updated using the following recurrence formula:

$$\mathbf{m}_{ij}(n+1) = \mathbf{m}_{ij}(n) + \alpha(n) \cdot h_{cij}(n) \cdot \left[ \mathbf{x}(n) - \mathbf{m}_{ij}(n) \right] \quad .$$

Here, $\alpha(n)$ is the learning rate, or the rate at which nodes converge to the input vector and decreases linearly between the initial and final step of the training. There are two types of neighborhood function, $h_{cij}$, available in the software; we have decided to use a "bubble" function:

$$h_{cij}(n) = F\left(\sigma(n) - d_{cij}\right) \ ,$$

where $F$ is a step function:

$$F(y) = \begin{cases} 0 & \text{if } y < 0 \\ 1 & \text{if } y \geq 0 \end{cases} \ .$$

Here, $d_{cij}$ is the Euclidean distance between a node and the "winner" node, and $\sigma(n)$ is the neighborhood radius that also linearly decreases between the initial and final step. Any node located within this radius is updated. Note that we have chosen the "bubble" function as it gives better results compared with the "Gaussian" function (Liu et al. 2006). The winning node becomes similar to the input vector and surrounding nodes develop representations of similar, but not the same, patterns. The above procedure is iterated during the training process.

The above training is undertaken in two phases. During the first phase, we use a larger initial learning rate $\alpha_o$ and a neighborhood radius $\sigma_o$ so that the reference vectors of the SOM array are ordered. Then, we used a smaller rate $\alpha_o$ and radius $\sigma_o$ to tune the values of reference vectors in the second phase. The training time is significantly longer in the second phase. After these two training phases, the SOM array consists of a number of patterns characteristic of the input data, with similar patterns nearby and dissimilar patterns further apart as shown in Fig. 2.

We tried many different sets of values for the learning rate, neighborhood radius, and training time to make a resulting quantization error to be small. Also, after testing several dimensions from 7 x 7 to 15 x 15, we have found that 7 x 7 (9 x 9) map is sufficient to describe the observed (simulated) SST variability in the tropical Indian Ocean.

## 3. Results

### *3.1 Observational data analysis*

The SOM array thus obtained from the observational data is shown in Fig. 2. Each node in Fig. 2 represents a SSTA pattern in the tropical Indian Ocean. The SSTA patterns with a basin-wide warming (cooling) are mostly located in the lower right (upper left) corner, while those with a positive (negative) SSTA to the west and a negative (positive) SSTA to the east are seen in the lower left (upper right) corner. The latter is known as the IOD (Saji et al. 1999). Neutral conditions with only small SSTA are distributed in the central part.

To obtain more insight into the SSTA patterns in Fig. 2, we have constructed a composite diagram of global SSTA for each node (Fig. 3). It is found that all nodes with basin-wide warming (cooling) are associated with El Niño (La Niña) in the tropical Pacific as in node G7 (A1). On the other hand, nodes representing the IOD are only sometimes associated with ENSO events; G1 is only associated with a weak SSTA in the central equatorial Pacific whereas G2 is associated with a strong SSTA in the eastern equatorial Pacific. This supports Saji et al. (1999), who suggested that IOD

may occur without ENSO events.

As is well known, both IOD and ENSO are phase-locked to the annual cycle. Therefore, a certain SSTA pattern tends to appear more frequently during a certain season (Fig. 4). To quantitatively discuss this, we have divided 7 x 7 SOM array in Fig. 2 into five groups based on the Euclidean distance with the SSTA patterns of four corner nodes (A1, A7, G1, and G7) and zero SSTA, and calculated the frequency of each group during each season (Fig. 5a). The IOD events start around the end of boreal spring, develop in summer, peak in fall, and decay quickly in winter (Saji et al. 1999). This is why the lower left quadrant corresponding to the positive IOD and the upper right quadrant corresponding to the negative IOD are more densely occupied in JJA (Jun.-Aug.) and SON (Sep.-Nov.) (Fig. 4) with frequency of around 15% (20%) for the positive (negative) IOD. On the other hand, ENSO events tend to peak toward the end of a year and its influence on the SSTA of the tropical Indian Ocean peaks about one season later (Klein et al. 1999). Hence, the lower right (upper left) quadrant corresponding to El Niño (La Niña) is more frequently occupied during DJF (Dec.-Feb.) and MAM (Mar.-May) with maximum in boreal spring (about 13%). All these are consistent with the existing views of IOD and ENSO (Chang et al. 2006)

The SOM is also successful in capturing the remarkable decadal variation as shown in Figs. 5b and 6. During the 1950s, the lower left group of the frequency map is occupied only 5.6%, because no major positive IOD event occurs during this decade. Instead, basin-wide warming patterns represented by the lower right group appear frequently (15.7%) owing to the more frequent occurrence of El Niño. This changes

dramatically in the 1960s and the 1970s, when the lower left quadrant is occupied more frequently (24.2% for 1960s and 14.2% for 1970s), indicative of more frequent occurrence of positive IOD events. Furthermore, basin-wide cooling patterns linked with La Niña in the upper left corner are densely populated during the 1970s (26.7%). This is a result of four La Niña events that occurred in this decade. The 1980s are again dominated by the basin-wide warming pattern in the lower right quadrant (23.2%) owing to the 1982/83 and 1986/87 El Niño events with minor IOD activity. In fact, the sum of occurrence for positive and negative IOD-like SSTA is the lowest among five decades with 18.2%. During the 1990s, two positive and negative IOD and El Niño events and one strong 1998/99 La Niña event occurred. Thus, the frequency map is rather uniformly distributed, although the negative IOD SSTA pattern is occupied 20.0% of time.

As a whole, we may conclude that the SOM is successful in detecting the observed SSTA variability in the tropical Indian Ocean. To discuss the variability in more detail, we use the model outputs from SINTEX-F1 model after validating them against the observation.

*3.2 Model output analysis*

To examine the model performance, we have applied a method used by Leloup et al. (2007b), who assessed the skill of CMIP3 models to simulate the ENSO. In brief, we have applied the SOM to the combined (observed and modeled) SSTA data after normalizing each of them separately. This will allow us to check whether the observed

and simulated SSTA patterns are similar or not. Figure 7 shows the 9 x 9 SOM array and Fig. 8 shows the frequency map for both the observation and the model. Again, the SSTA patterns with a basin-wide warming (cooling) are found around the lower right (upper left) corner, while those with the positive (negative) IOD are seen in the lower left (upper right) corner.

All SSTA patterns seen in the observation are captured by the SINTEX-F1 model as a node without any shading does not exist in Fig. 8b; the SSTA patterns associated with El Niño and negative IOD events are well captured by the model. This is remarkable considering the fact that many CMIP3 models show difficulties in capturing even the observed SSTA pattern in the tropical Pacific associated with ENSO events (Leloup et al. 2007b). We note, however, that the percentage of model SSTA that is never seen in the observation amounts to 7.3%. In particular, the nodes in the lower left corner with a strong positive IOD (I1 and I2) in the model are not covered by the observation. This is because the model prefers a strong positive IOD owing to the shallower thermocline bias in the eastern equatorial Indian Ocean (Tozuka et al. 2007).

To check whether the model is successful in simulating the seasonality of SSTA discussed in the previous subsection, we then calculate the seasonal frequency (Fig. 9). As in the observation, the model appears to capture more basin-wide warming and cooling in boreal winter and spring, and more dipole SSTA pattern in boreal summer and fall.

The dominant simulated SSTA pattern also undergoes decadal modulation (Fig. 10) as in the observation (Fig. 5b and 6). For instance, years from 182 to 191 are not

associated with positive IOD SSTA patterns, but associated with frequent basin-wide warming due to El Niño as in the 1950s, whereas years from 42 to 51 are associated with frequent basin-wide cooling as in the 1970s. This is contrasted with some other coupled GCMs, where occurrences of climate modes are too regular. Given that the model has a relatively good skill in simulating the SSTA patterns and their seasonal and decadal variability, we expect that it has a potential to provide some insight into the tropical climate variability. Also, the SOM is shown to be a useful tool to validate model outputs against observations.

To capture the anomaly pattern in SST, rainfall, zonal wind stress, and sea level pressure associated with the basin-wide warming and the positive IOD, composite diagrams are constructed for each node (Fig. 11). We have selected nodes with the positive IOD-like SSTA (H2 and I3) and a node with the basin-wide warming (I9). Nodes such as I1, I2, and H3 are not used for the composites, because these three nodes are associated with SSTA patterns that are never observed (Fig. 8). As expected, H2 and I3 are associated with a warm (cold) SSTA in the western (eastern) tropical Indian Ocean with an easterly wind stress anomaly and I9 is associated with El Niño in the Pacific with a basin-wide warming in the tropical Indian Ocean. The composite diagrams show interesting features in the SSTA patterns for the tropical Pacific; H2 shows a weak positive SSTA in the central equatorial Pacific, whereas I3 and I9 are associated with a strong warming in the eastern equatorial Pacific. The Pacific SSTA pattern of H2 reminds us of the SSTA in 1994 and El Niño Modoki (Ashok et al. 2007; Weng et al. 2007). To investigate the possible link between the IOD and

the warming in the central equatorial Pacific will be an interesting topic to deepen our understanding on the connection between IOD and ENSO (Yamagata et al. 2004).

There is a remarkable difference in the rainfall anomaly over the tropical Indian Ocean, especially near East Africa. It receives more rainfall only when the SSTA over the tropical Indian Ocean is associated with a warm (cold) SSTA to the west (east) as in H2 and I3. In contrast, a weak negative rainfall anomaly is seen in I9 for cases of strong warming over the eastern equatorial Pacific and basin-wide warming over the Indian Ocean. This supports Behera et al. (2005), who demonstrated that the positive IOD causes the enhanced East African short rains.

In H2 and I3, a distinct dipole pattern is seen in SLPA with a negative (positive) SLPA to the west (east). In particular, the large SLPA over the Maritime Continent suggests that the IOD events affect the Southern Oscillation through its influence on the surface pressure field near Darwin, Australia as shown by Behera and Yamagata (2003) for the observation. There is another interesting difference between H2 and I9 even though both of them show a positive SSTA in the equatorial Pacific. The SLPA in I9 shows an east-west asymmetry over the Pacific, indicative of negative Southern Oscillation Index (SOI) (Walker, 1924). However, no statistically significant SLPA is found in H2. This suggests that the westerly wind anomaly seen along the equatorial Pacific in H2 is triggered by the significant positive SLPA around the maritime continent generated by the IOD.

## 4. Conclusions

Using the Self-Organizing Map (SOM) analysis, we have investigated the tropical Indian Ocean climate variability. The SOM successfully captured the dipole SSTA pattern associated with the Indian Ocean Dipole (IOD) and basin-wide warming (cooling) associated with El Niño (La Niña) in the observation data of the HadISST. The IOD pattern appears during boreal summer and fall, while the ENSO pattern shows up more often during boreal winter and spring, confirming that these phenomena are phase-locked to the annual cycle. Furthermore, the occurrence of these patterns undergoes significant decadal variation. The positive IOD pattern appears more frequently during the 1960s, 1970s, and 1990s, whereas the basin-wide warming pattern appears more often in the 1950s, 1980s, and 1990s.

Then, we have examined the outputs from 200 years integration of SINTEX-F1 model that successfully captured the above features. Composite diagrams constructed for nodes in the SOM array for the simulated SSTA have revealed interesting differences. For the nodes with basin-wide warming, a strong positive SSTA in the eastern equatorial Pacific, together with a negative Southern Oscillation is seen. These nodes are also associated with a weak negative precipitation anomaly in East Africa. On the other hand, the nodes with the positive IOD are associated with a weak positive SSTA over the central equatorial Pacific or a positive SSTA over the eastern equatorial Pacific. These nodes show a positive (negative) sea level pressure anomaly in the eastern (western) tropical Indian Ocean. In contrast to the nodes with basin-wide warming, they are associated with a positive precipitation anomaly over East Africa. In general, the SOM derives results similar to observation.

The SOM approach seems to be a powerful analysis method in climate research; it is successful in describing the SSTA variability in the tropical Indian Ocean and useful in validating model outputs against observations. As is clear from our results, the SOM allows us to construct composite diagrams without either introducing indices based on a priori knowledge or obtaining principal components by the EOF analysis. Also, we are away from a danger of discussing a statistical artifact as in the EOF analysis, since artificial nodes that appear in a SOM array is never occupied (cf. Liu et al. 2006). Of course, a following method similar to the SOM can be used to check whether a pattern extracted by EOF is real or not. After extracting first seven modes using the EOF method, the observed SSTA patterns are projected onto these modes based on the Euclidean distance. The frequency of projections on each EOF mode is shown in Fig. 12. As expected, grave modes with larger variance have higher frequency. More importantly, the nodes for the second EOF mode representing the IOD are found, suggesting that the IOD is not a statistical artifact of the EOF method (cf. Behera et al. 2003). However, this method is limited by the orthogonality condition. Moreover, the EOF method cannot separate modes when several modes have similar variance as sample vectors become a random mixture of the true eigenvectors (North et al. 1982). In contrast, the SOM method can separate modes even when they have the same variance (see Appendix). Although linear statistical analyses cannot detect asymmetry between a positive phase and a negative phase of a particular mode, the SOM method can. This is another advantage. Actually, we see asymmetry in the positive and negative IOD. The difference in the basin-wide warming and cooling (Fig. 2) is also

derived, and Leloup et al. (2007a) were also successful in showing asymmetry in El Niño and La Niña using the SOM. Furthermore, differences among events can only be captured by the SOM. For instance, as shown in Figs. 2 and 3, the positive IOD is associated with the maximum in positive SSTA over the central equatorial Indian Ocean in node G1, whereas that is located in the Arabian Sea in node G2. All of these suggest that the SOM is a very useful tool in discussing large-scale ocean-atmosphere coupled phenomena. In the present article, we have provided one such concrete example for the tropical Indian Ocean climate variability.

**Acknowledgements**

This study is benefited from discussions with Drs. S. K. Behera and I. Iskandar. Constructive comments from three reviewers helped us to improve our manuscript. We are indebted to Dr. R. Zhang for data management. The SINTEX-F1 model was run on the Earth Simulator. The SOM_PAK software was provided by the Neural Network Research Centre at the Helsinki University of Technology and is available at http://www.cis.hut.fi/research/som_pak. The present research is supported by the 21st century COE grant from the Ministry of Education, Culture, Sports, Science, and Technology of Japan for the "Predictability of the Evolution and Variation of the Multi-scale Earth System: An Integrated COE for Observational and Computational Earth Science" of the University of Tokyo, and the Japan Society for Promotion of Science through Grant-in-Aid for Scientific Research (A) 17204040.

**Appendix**

To examine whether SOM and EOF can capture two climate modes with the same variance, we have generated an artificial SSTA data by

$SSTA(x,y,t) = pc1(t) \times EOF1(x,y) + pc2(t) \times EOF2(x,y) + noise(x,y,t)$ .

Here, $pc1(t)$ and $pc2(t)$ are the first two normalized principal components, and $EOF1(x,y)$ and $EOF2(x,y)$ are the normalized eigenfunctions or spatial patterns of the first two EOF modes of the tropical Indian Ocean shown in Fig. 12a. Note that we made the random noise $noise(x,y,t)$ to contribute to about 10% of the total variance and each of the first two modes explains about 45% of the total variance as they are normalized.

When the SOM method is applied to this data, an SOM array, which is very similar to Fig. 2, is obtained (figure not shown). However, as is clear from Fig. A1, the EOF method failed to capture the original SSTA pattern for the first two modes shown in Fig. 12. The first EOF mode in Fig. A1 has a larger SSTA in the southeastern tropical Indian Ocean, whereas that in Fig. 12a has a larger SSTA in the west. We note that this mixture occurs even when more than three modes have a similar variance and rotated EOF could not provide a remedy for this particular problem (see also Behera et al. 2003). A similar pitfall was also met by Dommenget and Latif (2002).

**References**

Allan NJ, Chambers D, Drosdowsky W, Hendon H, Latif M, Nicholls N, Smith I, Stone R, Tourre Y (2001) Is there an Indian Ocean Dipole, and is it independent of El Niño-Southern Oscillation? CLIVAR Exchange 6: 18-22

An SI, Ye Z, Hsieh WW (2006) Changes in the leading ENSO modes associated with the late 1970s climate shift: Role of surface zonal current. Geophys Res Lett 33: doi:10.1029/2006GL026604

Annamalai H, Murtugudde R (2004) Role of the Indian Ocean in regional climate variability. In Earth's climate: The ocean-atmosphere interaction. Geophys Monogr Ser 147: 213-211

Ashok K, Guan Z, Yamagata T (2001) Impact of the Indian Ocean Dipole on the decadal relationship between the Indian monsoon rainfall and ENSO. Geophys Res Lett 28: 4499-4502

Ashok K, Behera SK, Rao SA, Weng H, Yamagata T (2007) El Niño-Modoki and its possible teleconnections. J Geophys Res, in press

Behera SK, Rao SA, Saji HN, Yamagata T (2003) Comments on "A cautionary note on the interpretation of EOFs". J Clim 16: 1087-1093

Behera SK, Yamagata T (2003) Influence of the Indian Ocean Dipole on the Southern Oscillation. J Meteor Soc Japan 81: 169-177

Behera SK, Luo JJ, Masson S, Delecluse P, Gualdi S, Navarra A, Yamagata T (2005) Paramount impact of the Indian Ocean Dipole on the East African Short Rains: A CGCM Study. J Clim 18: 4514-4530

Black E, Slingo J, Sperber KR (2003) An observational study of the relationship between excessively strong short rains in coastal East Africa and Indian Ocean SST. Mon Wea Rev 131: 74-94

Chang P, Yamagata T, Schopf P, Behera SK, Carton J, Kessler WS, Meyers G, Qu T, Schott F, Shetye S, and Xie SP (2006) Climate fluctuations of tropical coupled systems-The role of ocean dynamics. J Clim 19: 5122-5174

Cheng P, Wilson RE (2006) Temporal variability of vertical nontidal circulation pattern in a partially mixed estuary: Comparison of self-organizing map and empirical orthogonal functions. J Geophys Res 111: doi:10.1029/ 2005JC003241

Collins DC, Reason CJC, Tangang F (2004) Predictability of Indian Ocean sea surface temperature using canonical correlation analysis. Clim Dyn 22: 481-497

Dommenget D and Latif M (2002) A cautionary note on the interpretation of EOFs. J Clim 15: 216-225

Guan Z, Yamagata T (2003) The unusual summer of 1994 in East Asia: IOD teleconnections. Geophys Res Lett 30: doi: 10.1029/2002GL016831

Klein SA, Soden BJ, Lau NC (1999) Remote sea surface temperature variations during ENSO: Evidence for a tropical atmospheric bridge. J Clim 12: 917-932

Kohonen T (1982) Self-organized information of topologically correct features maps. Biol Cyber 43: 59-69

Kohonen T, Hynninen J, Kangas J, Laaksonen J (1995) SOM_PAK, The self-organizing map program package version 3.1, Laboratory of Computer and Information Science, Helsinki University of Technology, Finland, 27pp

Kohonen T (2001) Self-Organizing Maps, 3rd edn. Springer-Verlag Berlin Heidelberg New York, 501pp

Leloup JA, Lachkar Z, Boulanger JP, Thiria S (2007a) Detecting decadal changes in ENSO using neural networks. Clim Dyn 28: 147-162

Leloup JA, Lengaigne M, Boulanger JP (2007b) Twentieth century ENSO characteristics in the IPCC database. Clim Dyn: doi:10.1007/s00382-007-0284-3

Liu Y, Weisberg RH, Mooers CNK (2006) Performance evaluation of the self-organizing map for feature extraction. J Geophys Res 111: doi:10.1029/2005JC003117

Luo JJ, Masson S, Behera SK, Gualdi S, Navarra A, Delecluse P, Yamagata T (2003) South Pacific source of the decadal ENSO-like variation as reproduced by a coupled GCM. Geophys Res Lett 30: doi: 10.1029/2003GL018649

Madec G, Delecluse P, Imbard M, Levy C (1998) OPA version 8.1 ocean general circulation model reference manual. Tech Rep/Note, 11: LODYC/IPSL, Paris, France, pp 91

North GR, Bell LT, Cahalan RF (1982) Sampling errors in the estimation of empirical orthogonal functions. Mon Wea Rev 110: 699-706

Rayner NA, Parker DE, Horton EB, Folland CK, Alexander LV, Rowell DP, Kent EC, Kaplan A (2003) Global analysis of SST, sea ice and night marine air temperature since the late nineteenth century. J Geophys Res 108: doi:10.1029 /2002JD002670

Richardson AJ, Risien C, Shillington FA (2003) Using self-organizing maps to identify

patterns in satellite imagery. Prog Oceanogr 59: 223-239

Roeckner E, coauthors (1996) The atmospheric general circulation model ECHAM4: model description and simulation of present day climate. Max-Plank Institute fur Meteorologie Rep 218: Hamburg, Germany, pp90

Saji NH, Goswami BN, Vinayachandran PN, Yamagata T (1999) A dipole mode in the tropical Indian Ocean. Nature 401: 360-363

Saji NH, Yamagata T (2003) Possible impacts of Indian Ocean dipole events on global climate. Clim Res 25: 151-169

Tozuka T, Luo JJ, Masson S, Behera SK, Yamagata T (2005) Annual ENSO simulated in a coupled ocean-atmosphere model. Dyn Atmos Ocean 39: 41-60

Tozuka T, Luo JJ, Masson S, Yamagata T (2007) Decadal modulations of the Indian Ocean Dipole in the SINTEX-F1 coupled GCM. J Clim 20: 2881-2894

Valcke S, Terray L, Piacentini A (2000) The OASIS coupler user guide version 2.4. Tech Rep TR/CMGC/00-10, CERFACS, Toulouse, France, pp 85

Vecchi GA, Harrison DE (2004) Interannual Indian rainfall variability and Indian Ocean sea surface temperature anomalies. In Earth's climate: The ocean-atmosphere interaction. Geophys Monogr Ser 147: 247-259

Walker GT (1924) Correlations in seasonal variations of weather. IX Mem India Meteorological Department 24: 275-332

Webster PJ, Moore A, Loschnigg J, Leban M (1999) Coupled ocean-atmosphere dynamics in the Indian Ocean during 1997-98. Nature 401: 356-360

Weng H, Ashok K, Behera SK, Rao SA, Yamagata T (2007) Impacts of recent El Niño Modoki on dry/wet conditions in the Pacific rim during boreal summer. Clim Dyn 29: 113-129

Wu A, Hsieh WW, Nonlinear Interdecadal changes of the El Niño-Southern Oscillation. Clim Dyn 21: 719-730

Yamagata T, Behera SK, Luo J-J, Masson S, Jury MR, Rao SA (2004) Coupled ocean-atmosphere variability in the tropical Indian Ocean. In Earth's climate: The ocean-atmosphere interaction. Geophys Monogr Ser 147: 189-211

**Figure captions**

**Figure 1:** Schematic diagram showing how to apply the SOM to obtain Fig. 2.

**Figure 2:** 7 x 7 SOM array of observed SSTA in the tropical Indian Ocean (40°-120°E, 17.5°S-17.5°N). Each node represents a SSTA pattern in the tropical Indian Ocean. All nodes are divided into five groups based on the Euclidean distance with the SSTA pattern of four corner nodes (A1, A7, G1, and G7) and zero SSTA, and they are shown by dashed boxes.

**Figure 3:** Composite diagrams of observed SSTA for nodes A1, A6, A7, G1, G2, and G7 in the SOM array of Fig. 2. Shading indicates anomalies exceeding 90% significance when tested by two-tailed t-test.

**Figure 4:** Seasonal frequency map of the 7 x 7 SOM array showing how frequently (in %) each SSTA pattern shown in Fig. 2 is observed in each season. A node without any shading indicates that the SSTA pattern of the node is never seen in a particular season.

**Figure 5:** Bar diagram showing (a) seasonal and (b) decadal frequency of each group. The 7 x 7 SOM array in Fig. 2 is divided into five groups based on the Euclidean distance with the SSTA pattern of four corner nodes (A1, A7, G1, and G7) and zero SSTA.

**Figure 6:** Decadal frequency maps of the 7 x 7 SOM array. They show how frequently (in %) each SSTA pattern shown in Fig. 2 is observed in each decade.

**Figure 7:** As in Fig. 2, but for the combined normalized SSTA including both 588 months of observed SSTA and 2400 months of simulated SSTA.

**Figure 8:** Frequency map of the 9 x 9 SOM array shown in Fig. 7 for (a) HADISST and (b) SINTEX-F1 model.

**Figure 9:** Bar diagrams showing how frequently each group is seen in each season for (a) the observation and (b) the SINTEX-F1 model. The 9 x 9 SOM array in Fig. 7 is divided into five groups based on the Euclidean distance with the SSTA pattern of four corner nodes (A1, A9, I1, and I9) and zero SSTA.

**Figure 10:** As in Fig. 9, but shows how frequently each group is seen in each decade in the SINTEX-F1 model.

**Figure 11:** Composite diagrams of simulated SSTA, rainfall anomaly, zonal wind stress anomaly, and SLPA for nodes H2, I3, and I9 in the SOM array of Fig. 7. Shading indicates anomalies exceeding 95% significance for precipitation and 99% significance for other variables when tested by two-tailed t-test.

**Figure 12:** (a) Spatial patterns of the first seven EOF modes for observed SSTA in the tropical Indian Ocean (40°-120°E, 17.5°S-17.5°N). Values in parentheses show the variance contribution of each mode. (b) Frequency map of the first seven EOF modes. For example, the value and shading corresponding to "1" on the x-axis and "-1.0" on the y-axis indicate that the frequency of occurrence of the first EOF mode with a basin-wide cooling is 25.0%. "0" on the x-axis indicates the frequency of occurrence that the SSTA is closest to zero SSTA in terms of the Euclidean distance.

**Figure A1:** First two EOF modes obtained from the artificial SSTA data. Values in parentheses show the variance contribution of each mode.

## Figures

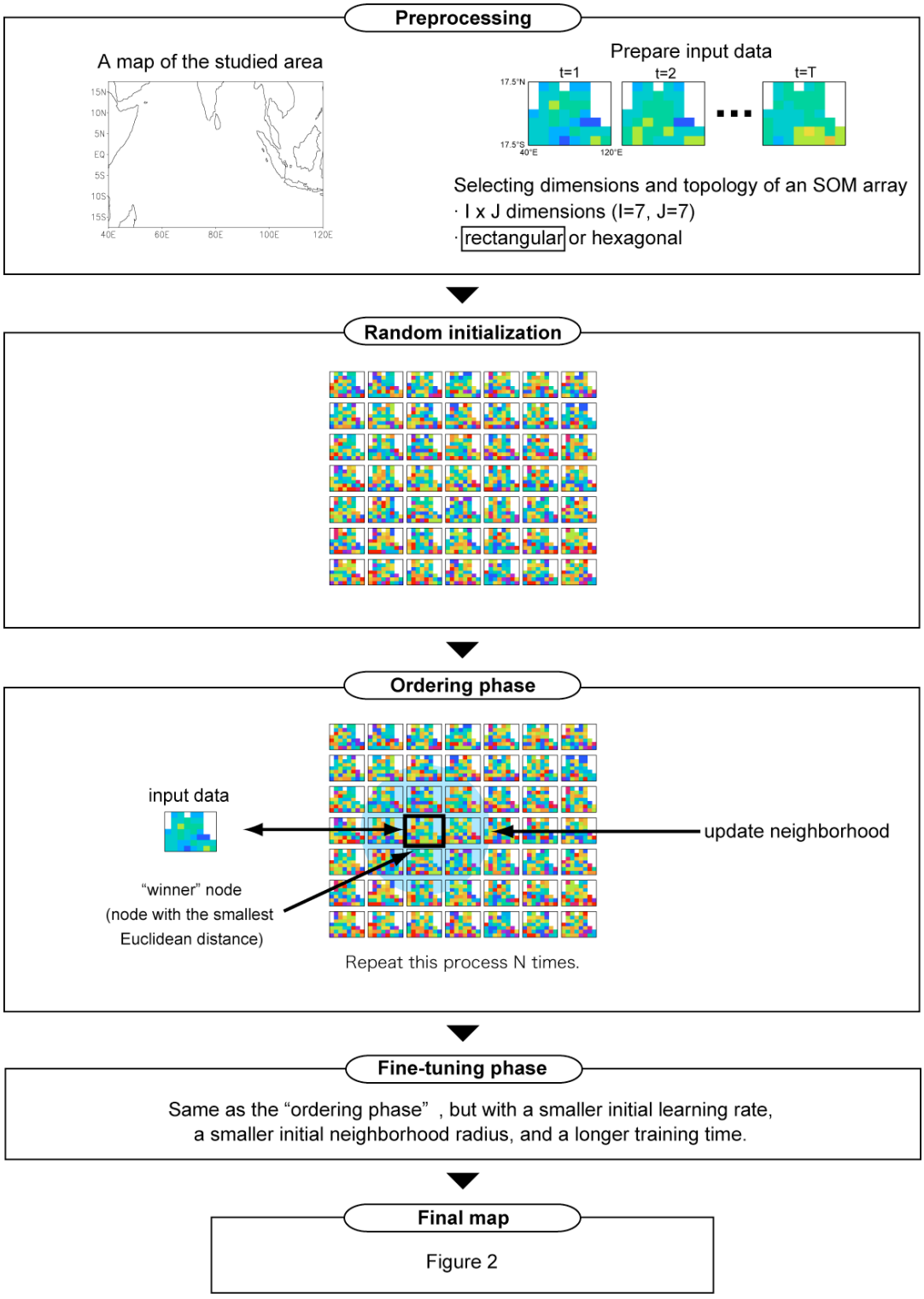


**Figure 1:** Schematic diagram showing how to apply the SOM to obtain Fig. 2.

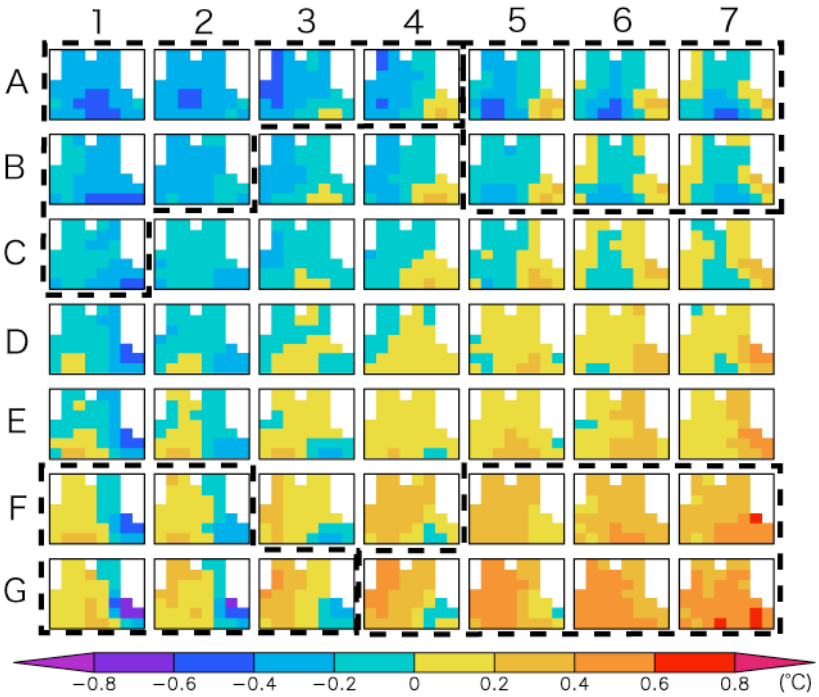


**Figure 2:** 7 x 7 SOM array of observed SSTA in the tropical Indian Ocean (40°-120°E, 17.5°S-17.5°N). Each node represents a SSTA pattern in the tropical Indian Ocean. All nodes are divided into five groups based on the Euclidean distance with the SSTA pattern of four corner nodes (A1, A7, G1, and G7) and zero SSTA, and they are shown by dashed boxes.

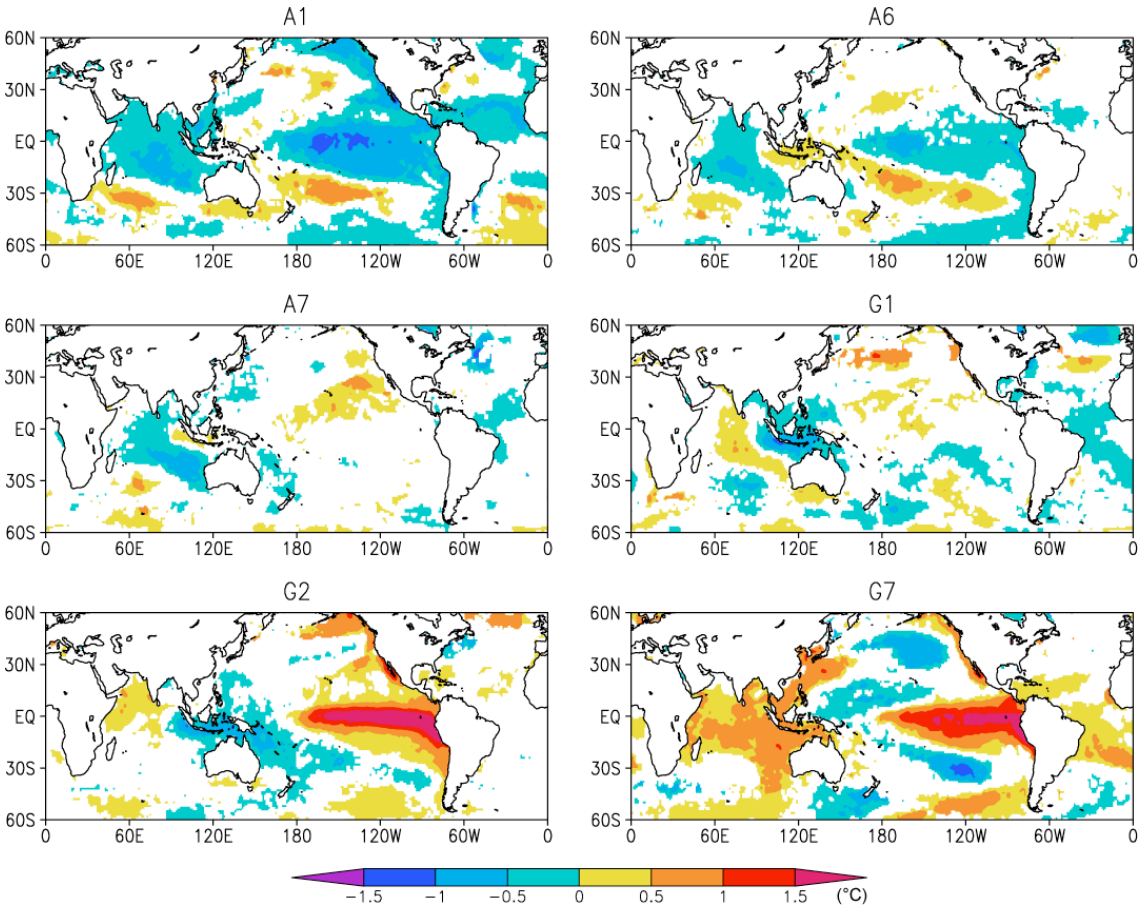


**Figure 3:** Composite diagrams of observed SSTA for nodes A1, A6, A7, G1, G2, and G7 in the SOM array of Fig. 2. Shading indicates anomalies exceeding 90% significance when tested by two-tailed t-test.

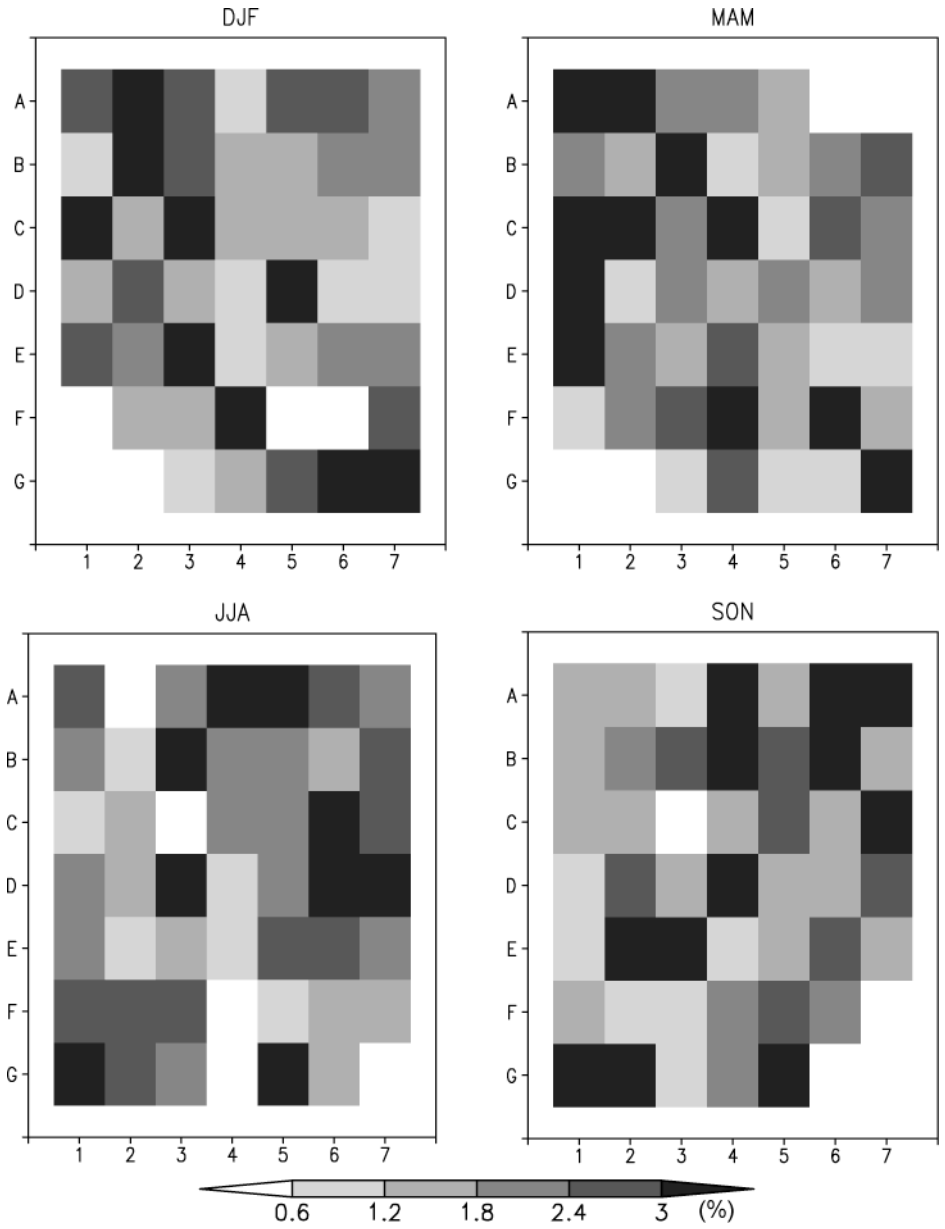


**Figure 4:** Seasonal frequency map of the 7 x 7 SOM array showing how frequently (in %) each SSTA pattern shown in Fig. 2 is observed in each season. A node without any shading indicates that the SSTA pattern of the node is never seen in a particular season.

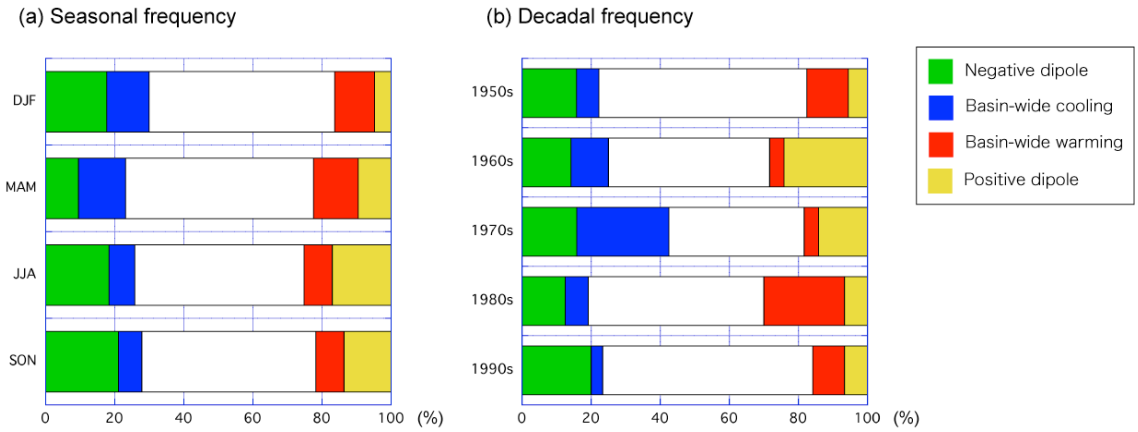


**Figure 5:** Bar diagram showing (a) seasonal and (b) decadal frequency of each group. The 7 x 7 SOM array in Fig. 2 is divided into five groups based on the Euclidean distance with the SSTA pattern of four corner nodes (A1, A7, G1, and G7) and zero SSTA.

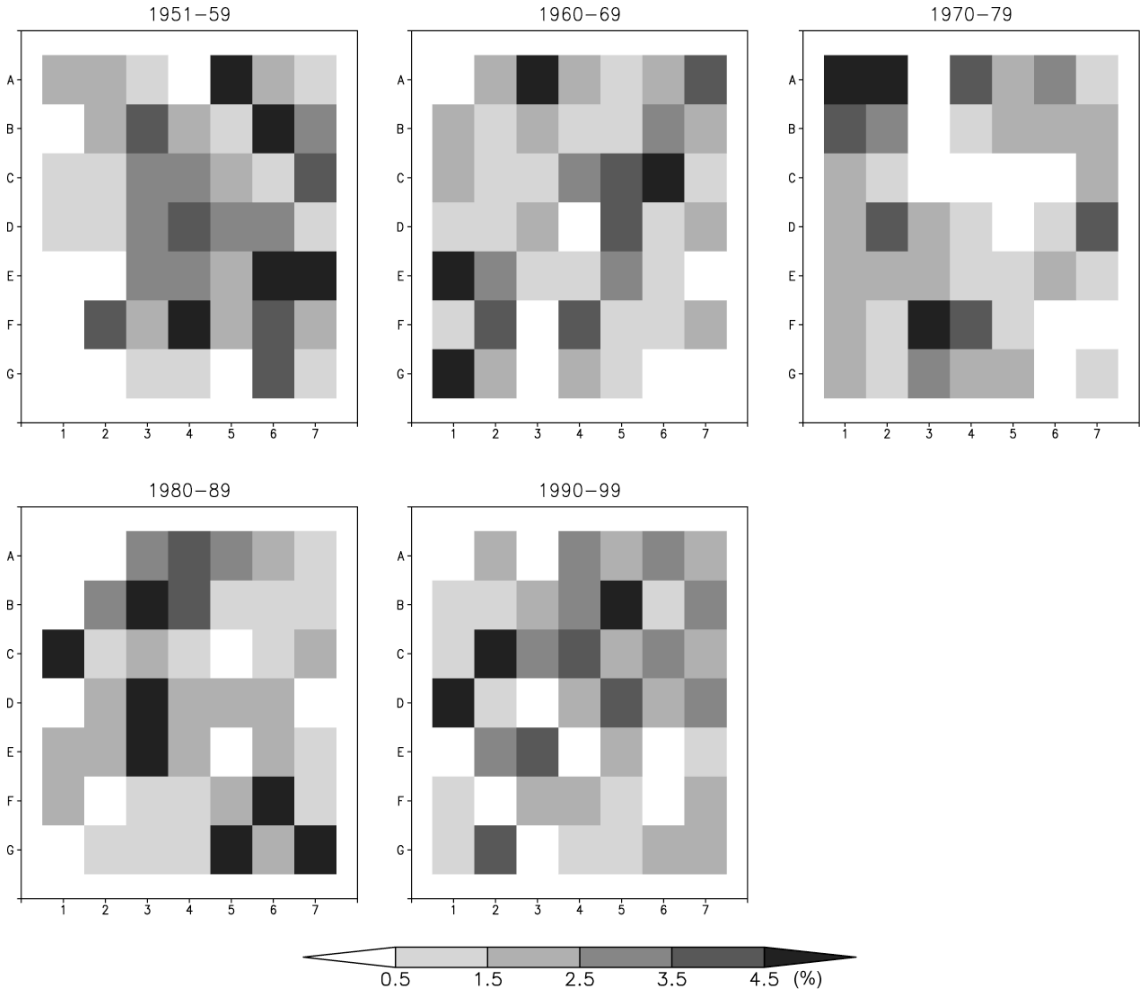


**Figure 6:** Decadal frequency maps of the 7 x 7 SOM array. They show how frequently (in %) each SSTA pattern shown in Fig. 2 is observed in each decade.

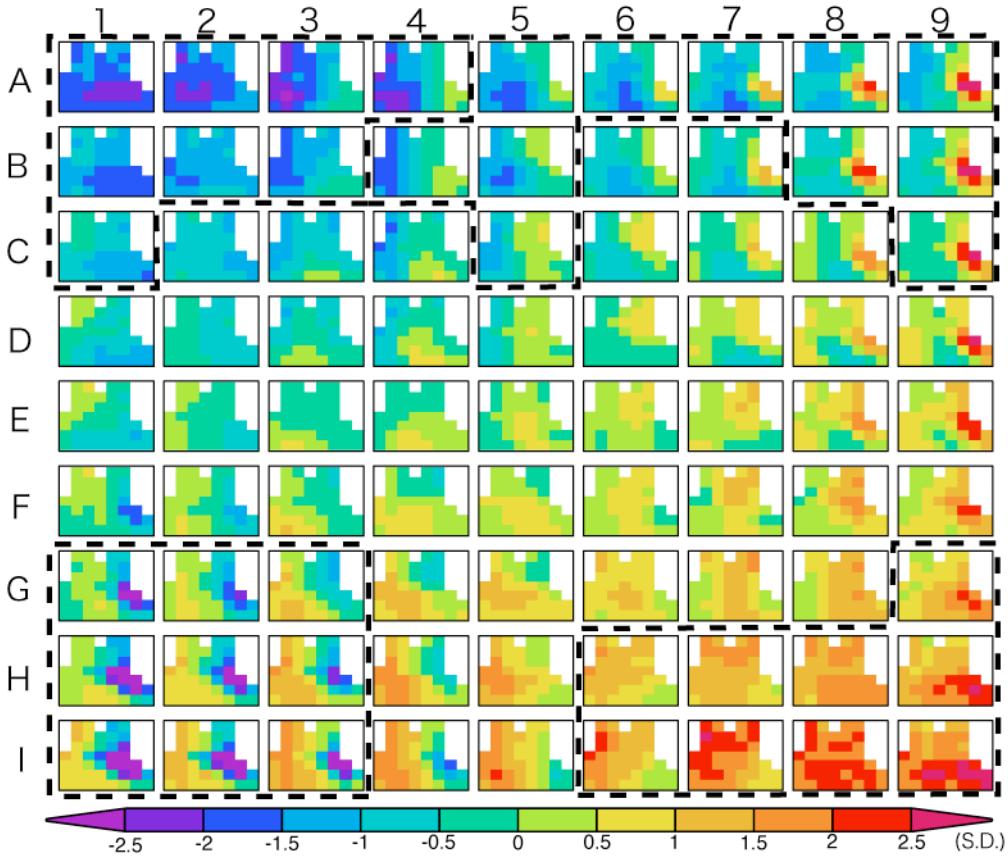


**Figure 7:** As in Fig. 2, but for the combined normalized SSTA including both 588 months of observed SSTA and 2400 months of simulated SSTA.

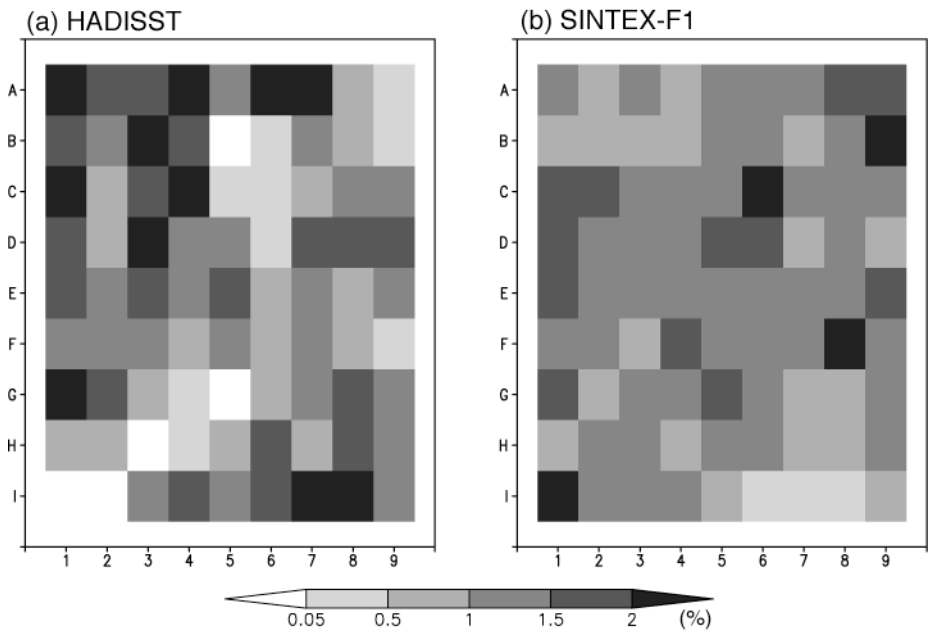


**Figure 8:** Frequency map of the 9 x 9 SOM array shown in Fig. 7 for (a) HADISST and (b) SINTEX-F1 model.

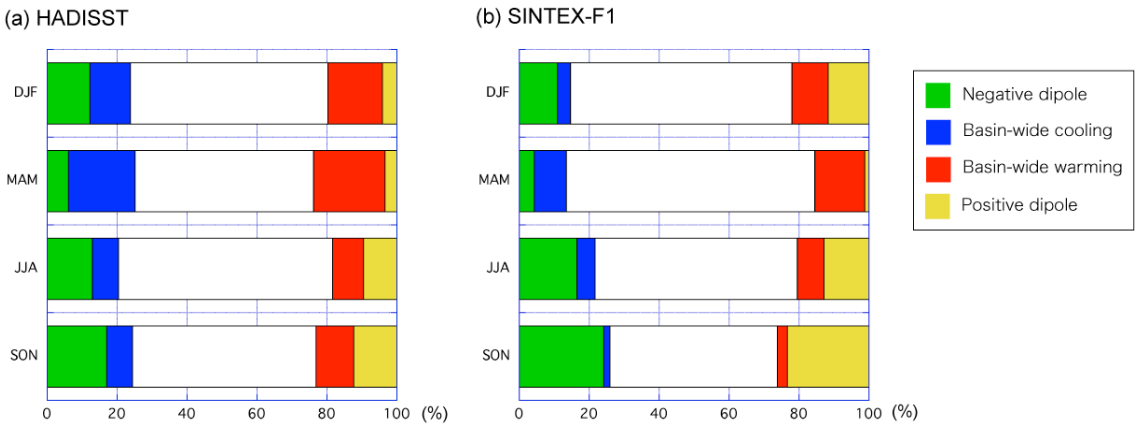


**Figure 9:** The 9 x 9 SOM array in Fig. 7 is divided into five groups based on the Euclidean distance with the SSTA pattern of four corner nodes (A1, A9, I1, and I9) and zero SSTA. Bar diagrams show how frequently each group is seen in each season for (a) the observation and (b) the SINTEX-F1 model.

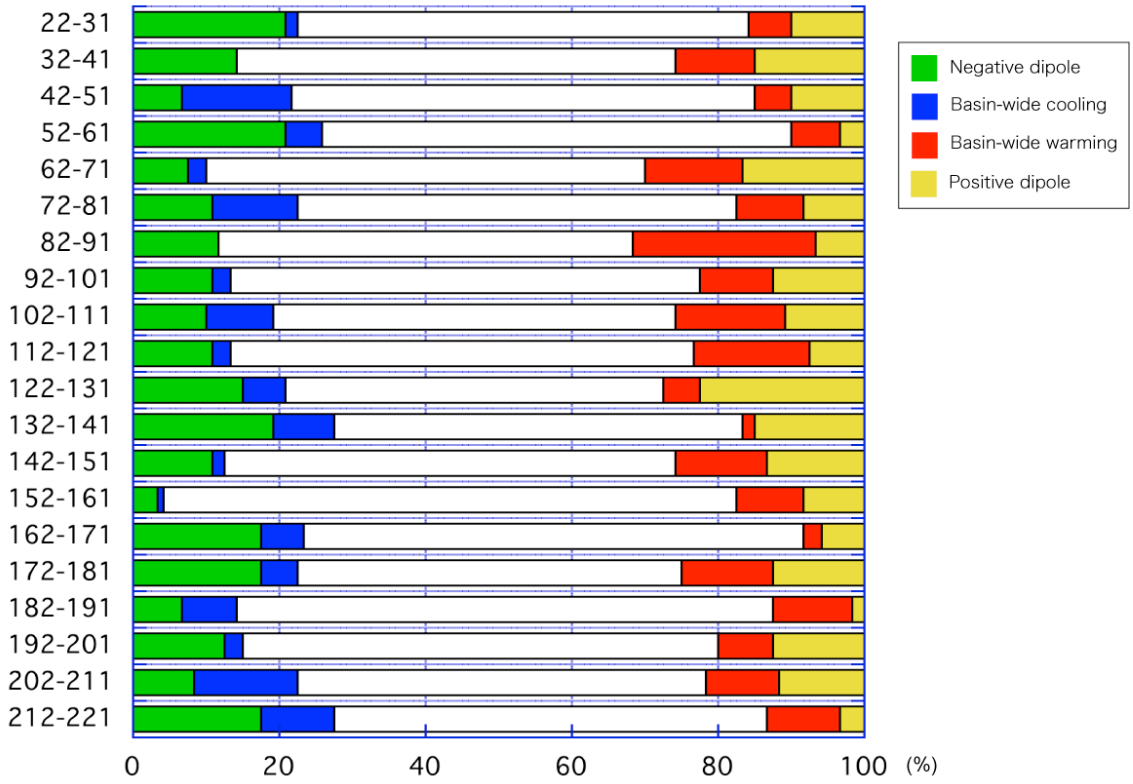


**Figure 10:** As in Fig. 9, but shows how frequently each group is seen in each decade in the SINTEX-F1 model.

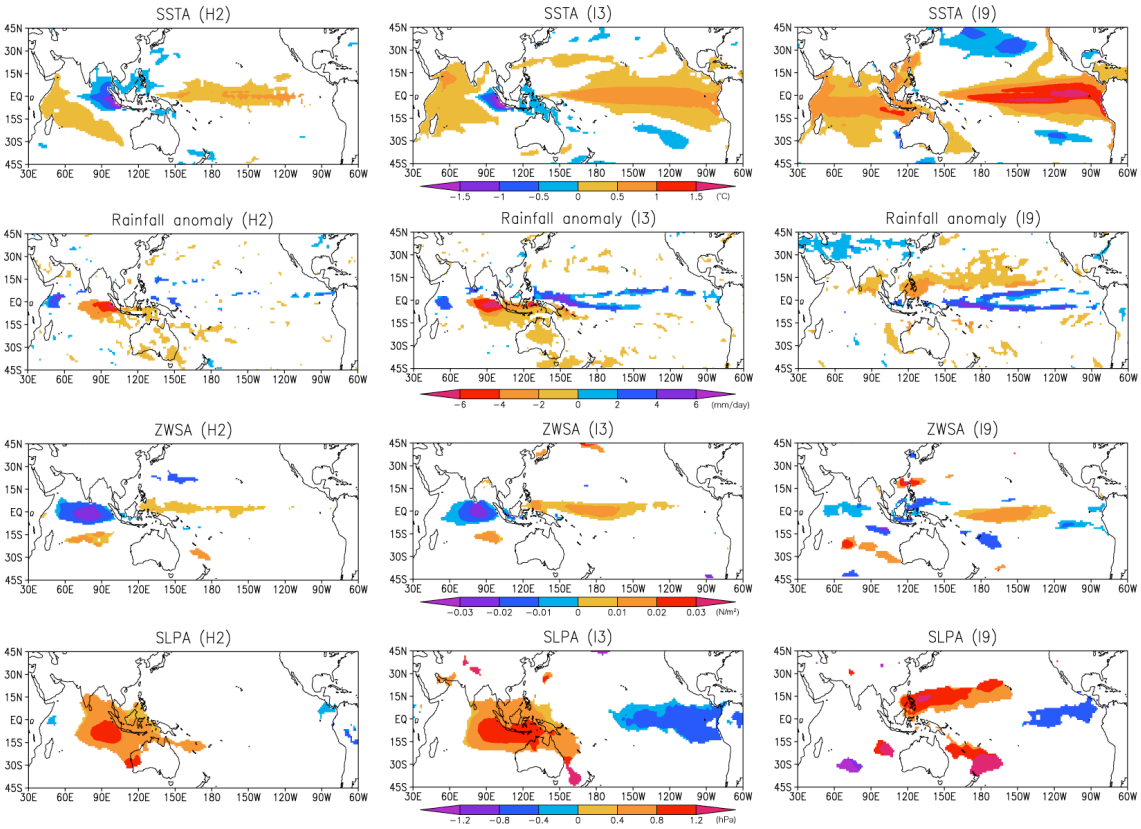


**Figure 11:** Composite diagrams of simulated SSTA, rainfall anomaly, zonal wind stress anomaly, and SLPA for nodes H2, I3, and I9 in the SOM array of Fig. 7. Shading indicates anomalies exceeding 95% significance for precipitation and 99% significance for other variables when tested by two-tailed t-test.

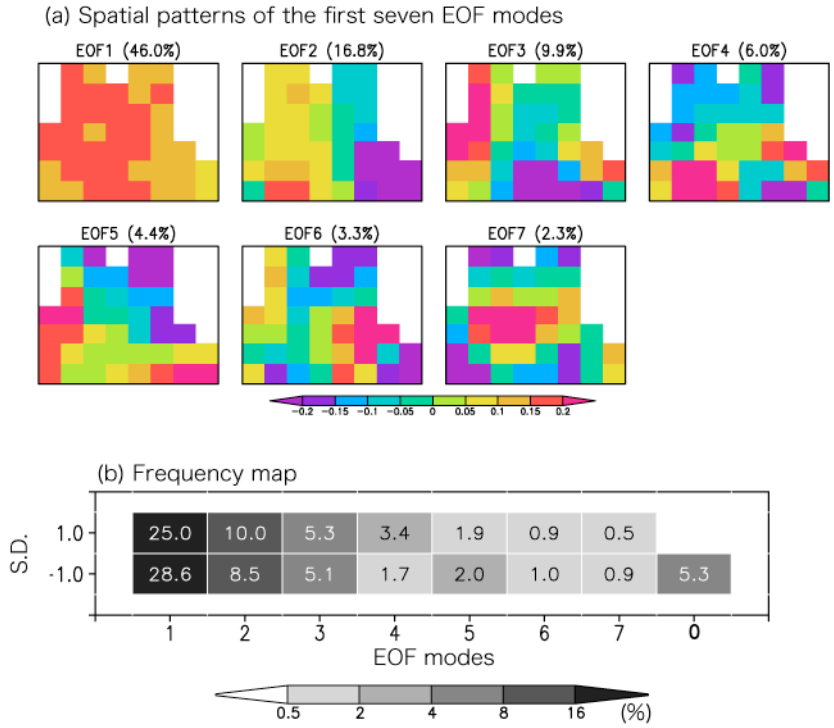


**Figure 12:** (a) Spatial patterns of the first seven EOF modes for observed SSTA in the tropical Indian Ocean (40°-120°E, 17.5°S-17.5°N). Values in parentheses show the variance contribution of each mode. (b) Frequency map of the first seven EOF modes. For example, the value and shading corresponding to "1" on the x-axis and "-1.0" on the y-axis indicate that the frequency of occurrence of the first EOF mode with a basin-wide cooling is 25.0%. "0" on the x-axis indicates the frequency of occurrence that the SSTA is closest to zero SSTA in terms of the Euclidean distance.

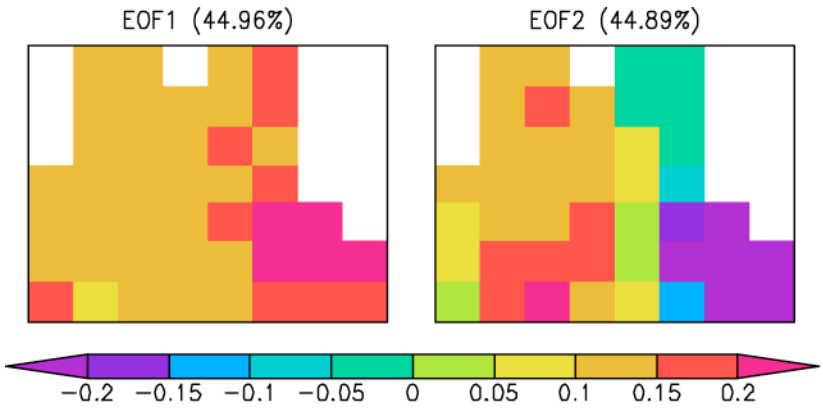


**Figure A1:** First two EOF modes obtained from the artificial SSTA data. Values in parentheses show the variance contribution of each mode.