Isomap nonlinear dimensionality reduction and bimodality of Asian monsoon convection

Article

Published Version

It is advisable to refer to the publisher's version if you intend to cite from the work. See Guidance on citing.

www.reading.ac.uk/centaur

CentAUR

Central Archive at the University of Reading

Reading's research outputs online

Isomap nonlinear dimensionality reduction and bimodality of Asian monsoon convection

A. Hannachi[1] and A. G. Turner [2]

[1] It is known that the empirical orthogonal function method is unable to detect possible nonlinear structure in climate data. Here, isometric feature mapping (Isomap), as a tool for nonlinear dimensionality reduction, is applied to 1958–2001 ERA-40 sea-level pressure anomalies to study nonlinearity of the Asian summer monsoon intraseasonal variability. Using the leading two Isomap time series, the probability density function is shown to be bimodal. A two-dimensional bivariate Gaussian mixture model is then applied to identify the monsoon phases, the obtained regimes representing enhanced and suppressed phases, respectively. The relationship with the large-scale seasonal mean monsoon indicates that the frequency of monsoon regime occurrence is significantly perturbed in agreement with conceptual ideas, with preference for enhanced convection on intraseasonal time scales during large-scale strong monsoons. Trend analysis suggests a shift in concentration of monsoon convection, with less emphasis on South Asia and more on the East China Sea. **Citation:** Hannachi, A., and A. G. Turner (2013), Isomap nonlinear dimensionality reduction and bimodality of Asian monsoon convection, *Geophys. Res. Lett.*, *40*, doi:10.1002/grl.50351.

1. Introduction

[2] The Asian summer monsoon (ASM) is one of the largest seasonal atmospheric phenomena involving huge moisture transports from ocean to land. This moisture transport ultimately makes the monsoon special as the main source of precipitation affecting livelihoods and infrastructure in the most populated region on Earth. Despite interannual variability of the ASM being a relatively small fraction of seasonal mean rainfall (e.g., for India, the coefficient of variation is around 10%), society is so finely tuned to the monsoon that variations on annual to intraseasonal time scales can cause huge problems relating to flood (infrastructure damage; health) and drought (crop damage; public water supply; hydro-electric generation). The shorter-time scale monsoon intraseasonal variations (MISV hereafter) show the strongest variability, with *active* and *break* conditions leading to more damaging impacts. While slow variations of the atmospheric lower boundary forcing such as snow cover and the El Niño-Southern Oscillation

(ENSO) lend predictability to seasonal rainfall anomalies in the tropics [*Charney and Shukla*, 1981], this predictability is limited by MISV [*Brankovic and Palmer*, 2000]. Recent assessments of state-of-the-art general circulation models (GCMs) show that few show skill at simulating all important characteristics of MISV [*Sperber and Annamalai*, 2008] highlighting the importance of better understanding the dynamics and predictability of the ASM. Furthermore, the relationship between the intraseasonal component of monsoon variability, the seasonal mean, and large-scale forcing conditions is unclear. Among the hypotheses put forward is that of a Lorenz model [*Palmer*, 1994] with chaotic fluctuations between active and break monsoon phases. While some studies have suggested that total seasonal rainfall may be broken down into a seasonal mean component forced by lower boundary conditions and the statistics of inherently unpredictable MISV [*Krishnamurthy and Shukla*, 2000, 2007], others suggest that MISV itself may be somehow related to boundary conditions and thus predictable: the large-scale forcing predisposing the system in a chaotic model to reside in one regime more than the other [*Webster et al*., 1998]. In fact, in relation to rainfall in central India, *Palmer* [1994] suggested that, under a given forcing, one of the predominant locations for convection (central India and the equatorial Indian Ocean Tropical Convergence Zone region) will be favored over the other according to the large-scale forcing.

[3] The ASM is a highly nonlinear and high dimensional phenomenon; one way to understand ASM dynamics is to find ways to reduce the dimensionality of the system in a way that could help capture the main features of its nonlinear behavior. *Sperber et al*. [2000] (SP00 hereafter) used empirical orthogonal function (EOF) analysis to reduce the dimensionality of National Centers for Environmental Prediction-National Center for Atmospheric Research reanalysis winds, identifying a common mode of variability on intraseasonal and interannual time scales. SP00 showed the interesting result that a probability density function (PDF) of an intraseasonal principal component time series could be translated towards negative or positive values according to seasonal mean conditions. However, only a small subset of MISV can be perturbed in this way by large scale forcing, and SP00 found no bimodality, the PDF being Gaussian, suggesting that the dominant modes of MISV are due to (inherently stochastic) internal processes, hence placing a limit on their predictability as part of the monsoon. Instead, *Straus and Krishnamurthy* [2007] showed that bimodality only exists under certain conditions. Clear bimodality of the South and East Asian summer monsoon activity, however, has not been established with certainty. In an earlier study by the authors [*Turner and Hannachi*, 2010, TH10 hereafter], one-dimensional Gaussian mixture

[1]Department of Meteorology, Stockholm University, Stockholm, Sweden.

[2]NCAS-Climate, Department of Meteorology, University of Reading, Reading, UK.

Corresponding author: A. Hannachi, Department of Meteorology, Stockholm University, Stockholm, Sweden. (a.hannachi@misu.su.se)

model analysis was performed on the leading mode of an intraseasonal outgoing longwave radiation (OLR) index of ASM convection. The OLR was unimodal but skewed, and the skewness was interpreted using a mixture model in terms of two intraseasonal monsoon regimes, namely active and break phases. TH10 suggested a preference for break conditions over India during seasonally weak monsoons. Note that this is not simply a trivial point, since the weak monsoon season may be caused by a country-wide seasonal anomaly related to large-scale forcing such as ENSO [*Krishnamurthy and Shukla*, 2000, 2007], even in the absence of any active/break activity.

[4] In this paper, we advance on earlier studies by using the isometric feature mapping (Isomap) method [*Tenenbaum et al.*, 2000] on sea-level pressure (SLP) to reduce the dimensionality of the ASM while maintaining nonlinear components. We then apply a multivariate Gaussian mixture model to estimate the PDF of the ASM within the obtained Isomap low-dimensional space. Isomap is based on interpoint distances rather than explained variance (as in EOFs) and is therefore more suited to study the nonlinear structure of the ASM. The data and methodology are described in section 2. Section 3 discusses the results and the implications are presented in the last section.

2. Data and Methodology

2.1. Data

[5] We have used daily SLP and 850 hPa wind fields from the ERA-40 project [*Uppala et al.*, 2005] of the European Centre for Medium-Range Weather Forecasts over the ASM region (50–145°E, 20°S–35°N) for the period 1958–2001. Daily June–September (JJAS) anomalies to the seasonal cycle are computed after first removing any linear long-term trend. One degree gridded Indian rainfall data covering the same period are also used [*Rajeevan et al.*, 2006]. In addition, to characterize the large-scale seasonal mean ASM, we have used the dynamical monsoon index (DMI) proposed by *Webster and Yang* [1992]. The DMI is a proxy for the heating of the atmospheric column over a broad region of the Asian monsoon and is defined as the JJAS average of anomalous zonal wind shear between the lower (850 hPa) and upper (200 hPa) troposphere, averaged over 40–110°E, 5–20°N. The index is scaled to zero-mean and unit-variance.

2.2. Isomap and Mixture Model

[6] Isomap is a technique for nonlinear dimensionality reduction based on preserving geodesic proximities using a non-Euclidean metric, and as such maintains nonlinear features of the original data that are lost in traditional EOF analysis [*Tenenbaum et al.*, 2000]. If the data lie on a nonlinear manifold, the geodesic metric measures precisely the interpoint distances between points on this manifold. The Isomap algorithm has three main steps. The first step is to use the available interpoint (usually Euclidean) distances, d_{ij}, for all i and j, to construct neighboring points. This is done here by selecting the K-nearest (for some value of K) points to a target point. The neighborhood is then defined as a weighted graph where the weight of the edges is represented by the distances d_{ij}, for all i and j. The second step consists of defining the geodesic distance δ_{ij} between any two points using the shortest path following the graph constructed in step 2. Once the dissimilarity matrix $\boldsymbol{\Delta} = (\delta_{ij})$

is obtained, the last step of Isomap consists of applying the classical multidimensional scaling (MDS) procedure [*Borg and Groenen*, 1997] to find the embedding space and the associated principal coordinates.

[7] Multidimensional scaling is a geometric method for reconstructing a configuration from its interpoint distances and enables visualizing proximities in low-dimensional spaces. Given a matrix of interpoint distances that are not necessarily Euclidean, between different pairs of the n data points $\mathbf{x}_k$:

$$\mathbf{D} = (d_{ij}) = (\|\mathbf{x}_i - \mathbf{x}_j\|), \tag{1}$$

where $k = 1, \ldots n$ within a high-dimensional space, the objective is to find a low-dimensional embedding space of the coordinates or configuration of the data points (i.e., of the data matrix $\mathbf{X} = (\mathbf{x}_1, \ldots, \mathbf{x}_n)^T$ that produces the matrix $\mathbf{D}$). The solution according to the classical metric problem [*Borg and Groenen*, 1997] is to compute first the double centered distance matrix:

$$\mathbf{A} = -\frac{1}{2}\left[\mathbf{I}_n - \frac{1}{n}\mathbf{1}\mathbf{1}^T\right]\mathbf{D}_2\left[\mathbf{I}_n - \frac{1}{n}\mathbf{1}\mathbf{1}^T\right]^T, \tag{2}$$

where $\mathbf{D}_2 = (d_{ij}^2)$, $\mathbf{I}_n$ is the identity matrix of order n, and $\mathbf{1} = (1, \ldots, 1)^T$ is the vector of length n containing ones. A singular value decomposition of $\mathbf{A}$, i.e. $\mathbf{A} = \mathbf{U}\boldsymbol{\Lambda}\mathbf{U}^T$, is then obtained and the principal coordinates (or data) matrix $\mathbf{X}$ can be estimated using

$$\mathbf{X} = \mathbf{U}\boldsymbol{\Lambda}^{\frac{1}{2}}. \tag{3}$$

[8] Since, in general, the dissimilarities are not Euclidean, the diagonal matrix $\boldsymbol{\Lambda}$ will not be positive, and therefore only the leading positive eigenvalues are considered in (3). When the dissimilarities are Euclidean, the MDS problem becomes equivalent to the standard empirical orthogonal function (EOF) method.

[9] In addition, to estimate the ASM PDF, we have used the multi-Gaussian mixture model (TH10) within the Isomap low-dimensional ASM space. Within this framework, the PDF $F(\mathbf{x})$ is written as the convex combination of two bivariate Gaussian distributions $g_{1,2}(\mathbf{x})$ as

$$F(\mathbf{x}) = \alpha g_1(\mathbf{x}) + (1 - \alpha)g_2(\mathbf{x}) \tag{4}$$

with α being the mixing proportion of the first component and where $g_{1,2}(\mathbf{x})$ are characterized each by its mean and covariance matrix (see TH10 for more details).

[10] EOFs, which maximize variance, and also MDS, which is a proximity-preserving method, are linear projective techniques and are therefore unable to detect nonlinear structures in the data. The idea is then to compute interpoint distances between data points using, not the global Euclidean metric, but the geodesic distances based on neighboring points. All the data points are then linked by a graph, using the nearest points, and the distance between any pair of points is computed by looking for the shortest path, on this graph, linking these points. The classical example in Isomap is that of the Swiss Roll [*Tenenbaum et al.*, 2000], for which the shortest Euclidean distance between two points that are far apart on the manifold does not represent the real geodesic distance. Here, the interpoint distance matrix $\mathbf{D}$ between all JJAS daily observation pairs of the SLP anomalies is first computed based on the Euclidean distance. To construct the neighborhood graph, we connect the data using the nearest $K = 12$ points to define neighborhood.
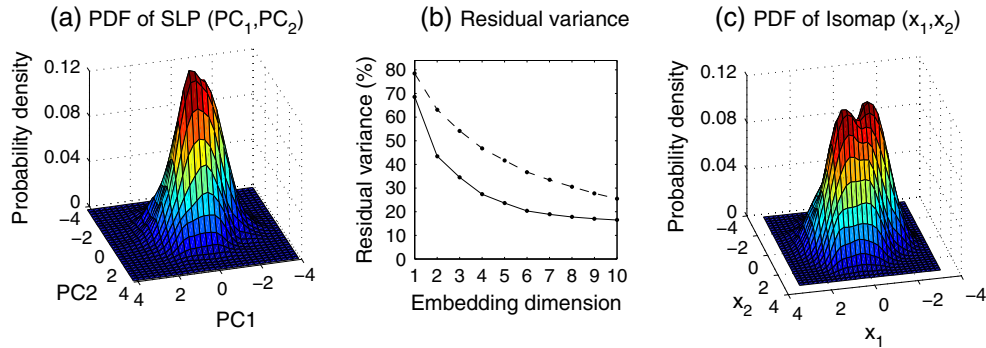
Figure 1. (a) Gaussian kernel PDF estimate using the leading two PCs of SLP anomalies. (b) Residual variance obtained from Isomap (continuous) and EOFs (dashed) versus the embedding dimension. (c) The Gaussian kernel PDF of the leading two Isomap time series x_1 and x_2.

Once computed, the geodesic dissimilarity matrix $\mathbf{\Delta} = (\delta_{ij})$ is submitted to an MDS decomposition.

3. Results

3.1. Isomap Structure of the Monsoon

[11] As in earlier studies examining principal components of 850 hPa wind and OLR in the monsoon domain (SP00 and TH10, respectively), the PDF of the two leading PCs of SLP anomalies (Figure 1a) is also unimodal. We next investigate the Isomap time series, which still contains nonlinearities. Figure 1b shows the Isomap residual variance of the SLP anomalies as a function of the embedding dimension. The residual variances associated with the leading two Isomap components are 69% and 44%. The same residual variance is also plotted for the leading 10 EOFs (Figure 1b). These values are larger than their Isomap analogs, because Isomap attempts to follow the nonlinear manifold of the data, unlike the EOFs that are obtained by looking for optimal linear subspaces. Therefore, Isomap is more suited to extracting nonlinear features compared to EOFs. Figure 1b shows also a clear elbow at embedding dimension $d = 2$. This elbow is normally taken [*Tenenbaum et al.*, 2000] as the dimension of the nonlinear manifold of the data. Accordingly, and for simplicity, we focus on two-dimensional Isomap embedding to look for possible nonlinearity in the ASM. Figure 1c shows the Gaussian kernel PDF estimate of the leading two Isomap time series to be clearly bimodal, unlike that of the leading two PCs (Figure 1a). We have used the optimal kernel width [*Silverman*, 1981] for the PDF estimation. The bimodality of the obtained PDF (Figure 1c) is robust to changes in the smoothing parameter, the number K used in the Isomap as well as changes in the ASM domain (not shown). We have also investigated other fields such as OLR where the bimodality was particularly strong, but we avoided using this field here due to uncertainties, especially prior to the satellite era when it is wholly modeled. Further support is provided by the mixture model (TH10) discussed next.

[12] To explain the above results further, we show in Figure 2a the Gaussian kernel estimate [*Silverman*, 1981] of the PDF within the Isomap plane along with a two-component mixture model (dashed line). The centers of the individual bivariate Gaussians of the mixture model are also shown, as small filled circles, along with their associated

covariance ellipses. Each ellipse delimits around 84% of the total mass of the corresponding component. The weights of the left and right components (Figure 2a) of the mixture model are respectively $\alpha = 0.45$ and $1 - \alpha = 0.55$.

[13] In order to identify the ASM conditions associated with the PDF modes, we have used a composite analysis of the nearest 300 ASM states to the centers of these components. The obtained ASM composites based on SLP data are shown in Figures 2b and 2c. Superimposed on these maps are the associated 850 hPa wind composites. The first ASM phase (Figure 2b) corresponds to the left-hand ellipse of Figure 2a and shows high pressure anomalies over most of the domain with particular enhancement over the East China Sea and also at the head of the Bay of Bengal in the South Asian monsoon trough region. This is accompanied by an anticyclonic low-level circulation over the East China Sea, weakening of the monsoon trough and obvious weakening of the Somali jet. As seen in Figure 2d, this phase corresponds to monsoon break conditions over India, with negative rainfall anomalies over the Western Ghats and in a band across central India, while there are positive anomalies near the Himalayan foothills.

[14] The second ASM phase (Figure 2c), which corresponds to the right-hand ellipse in Figure 2a, shows low pressure anomalies over much of the domain, most notably the anomalously strong low pressure trough and associated cyclonic circulation over the Philippine and East China Seas. In the Indian Ocean region, the flow regime is consistent with the active phase of the South Asian monsoon, with an enhanced Somali jet and circulation around the monsoon trough. Opposite rainfall conditions to those in the break phase are shown in Figure 2e, with wet anomalies over the Western Ghats and central India, and those of opposite sign to the north and south. The break and active rainfall conditions shown in Figures 2d and 2e are well consistent with observed composites [e.g., *Krishnamurthy and Shukla*, 2000, 2007]. The strong center of action of variability over the Western North Pacific shown in Figures 2b and 2c is consistent with analysis performed by *Sperber and Annamalai* [2008], regressing OLR against an ISV index (their Figures 5d and 5h indicating reduced and enhanced convection over the Western North Pacific, respectively).

[15] TH10 used an OLR index over the same domain and identified similar circulation regimes over the Western North Pacific using EOF analysis. However, no signals were noted in the Somali jet (the fundamental characteristic of the
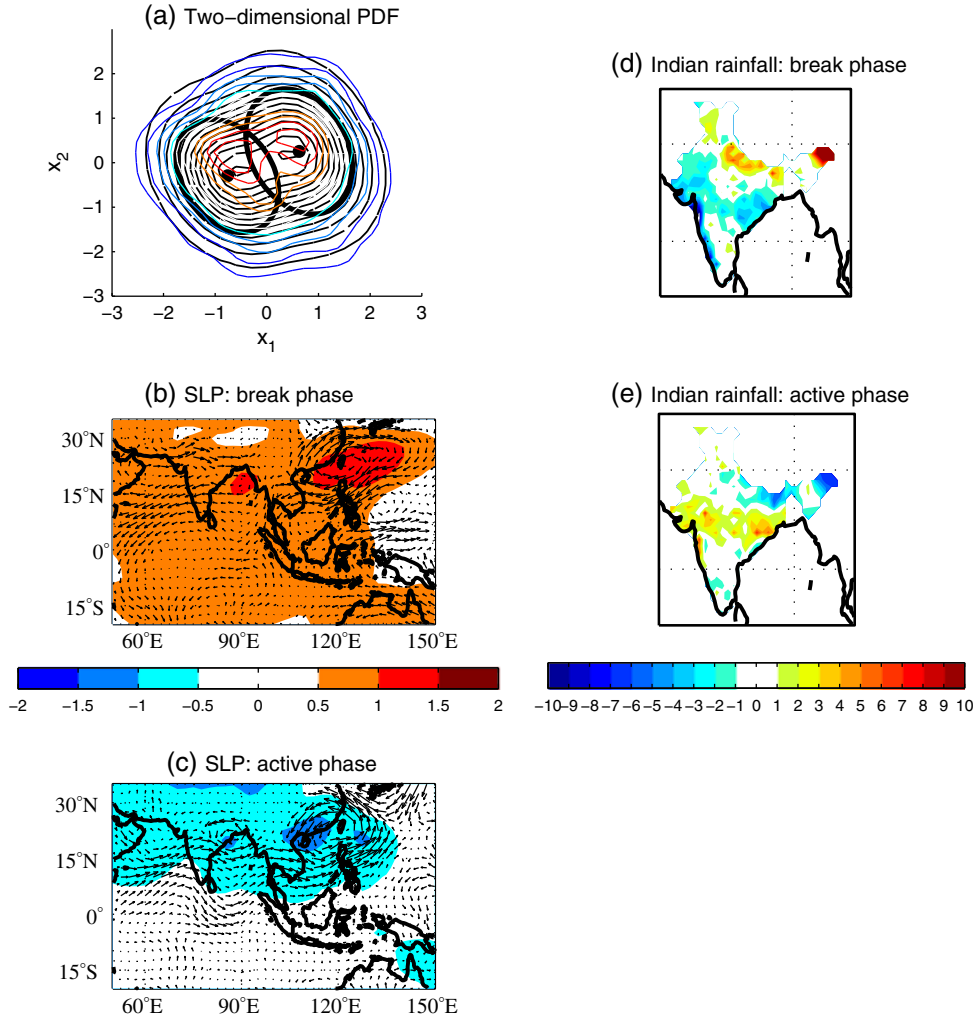
Figure 2. (a) Kernel (continuous) and mixture model (dashed) PDFs of x_1 and x_2. The centers of the mixture components and associated covariance ellipses are also shown. The composites of the 300 closest points to the left and right centers shown in Figure 2a of the SLP, and 850 hPa wind anomalies are shown in Figures 2b and 2c, respectively, and for rainfall anomalies in Figures 2d and 2e, respectively. The SLP contour interval 0.5 hPa and the maximum speed is 1.8 m/s. Rainfall units are mm/day.

South Asian monsoon circulation) and indeed rainfall conditions over India were associated with opposite phases of the Western North Pacific circulation shown here. The result presented here is more consistent with current understanding relating the monsoon wind and precipitation, with rainfall anomalies over South Asia in some way related to perturbations in the Somali jet. The difference between the results in this study and those in TH10 may be due to the methods used. As we have described earlier, the use of Isomap is advantageous over the use of EOFs, particularly in the context of nonlinearities in the ASM. In addition, the introduction of satellite data to the reanalysis product in the late 1970s may have affected OLR (used in TH10) far more than SLP (used here).

3.2. Relationship Between Intraseasonal Regimes and the Seasonal Mean Monsoon

[16] SP00 did not find bimodality in the monsoon intraseasonal variability (MISV). However, they found that the PDF mean of the third PC time series was systematically perturbed left or right during weak and strong monsoon years categorized in terms of seasonal mean all-India rainfall. Here, we investigate the relationship between MISV and the large-scale seasonal mean monsoon using the DMI. To characterize this relationship, we simply stratify, as in TH10, the daily SLP anomalies within Isomap space, according to whether the DMI is larger (smaller) than 1 (–1) standard deviations from the mean.

[17] Figures 3a and 3b show the Gaussian kernel estimate of the two-dimensional PDF using the leading two Isomap time series during weak (Figure 3a) and strong (Figure 3b) seasons as measured by the DMI. Although the PDF during both phases combined (not shown) is not bimodal but skewed towards the right-hand mode of Figure 2a, i.e., the active phase, it is clear that during strong DMI seasons (Figure 3b) we have a maximum value of probability of the active monsoon phase. During weak DMI seasons (Figure 3a), on the other hand, we have maximum probability associated with an increased frequency of the break monsoon phase. In addition, the skewness of the PDF shows that beside the peak probability being associated with active (break) conditions, we still observe finite and relatively

4
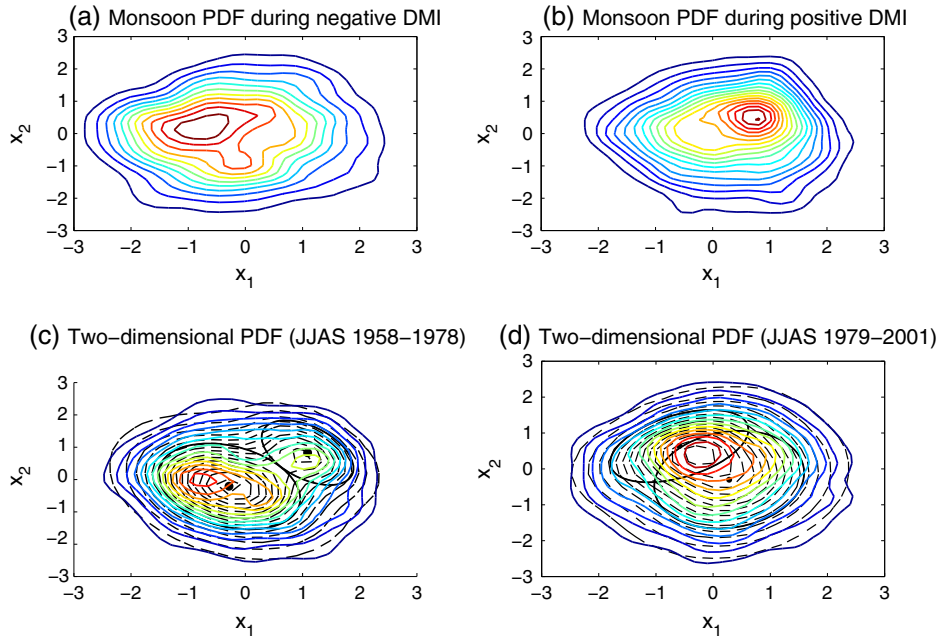
Figure 3. Kernel PDF estimate of the monsoon during weak (a) and strong (b) DMI seasons, and the two-dimensional PDF estimates of the Isomap time series obtained using the non-detrended SLP anomalies for the (c) first and the (d) second halves of the record.

large probabilities of break (active) conditions. In particular, the maximum probabilities erring towards active or break phases agree well with the conceptual paradigm put forward by *Palmer* [1999] regarding the change of regime frequency under forcing changes in a chaotic system. *Palmer* [1994] also postulated that seasonal mean conditions relate to preference for a particular phase of MISV, a result confirmed here. So our results suggest a reconciliation between the conceptual model proposed by *Palmer* [1994], which suggests an idealized upper bound on monsoon predictability, and SP00, which suggests a lower bound (namely that only a subset of the intraseasonal variability can be perturbed by the large scale).

3.3. Trends in Monsoon Regimes

[18] We also examine evidence for the existence of any trend in the nonlinearity of the monsoon phases by considering (non-detrended) SLP anomalies for pre- and post-1979 separately. An Isomap analysis of these two data sets is performed, and the kernel and mixture PDFs are computed (Figures 3c and 3d). The bimodality is clear in the first period, with 79% and 21% weights for the break and active phases, respectively (Figure 3c). In the second period, however, the PDF is unimodal (Figure 3d), but we have fitted a two-Gaussian mixture because it provides a better fit. The weights are now 33% and 68% for the top left and bottom right phases, respectively (Figure 3d). The top left regime (not shown) is similar to the break phase. The other regime (not shown) has a low pressure center over eastern China and the Philippine Sea, the pattern looking rather like the quadrupole of MISV convection seen in observations by *Annamalai and Slingo* [2001]. The trend itself could be described as an eastward shift of the region of intensive monsoon conditions, similar to that seen in *Annamalai et al.* [2013] who saw increases in convection in the Western North Pacific at the expense of the South Asian monsoon.

We note caution however, in describing trends as a result of the reanalysis product, given the changes in methodology of its composition over time. However, there are no comprehensive directly-observed daily SLP products available for the region.

4. Summary and Discussion

[19] We have investigated monsoon intraseasonal variability using nonlinear dimensionality reduction based on Isomap of ERA-40 daily SLP anomalies over the ASM region for summer (JJAS) 1958–2001. The Isomap projection technique is based on computing local geodesic distances between atmospheric states. The data points are first connected by a graph based on the 12 nearest points, then distances between any two states are computed based on the previous graph. MDS is then used to get an Isomap embedding. A kernel PDF estimate is fitted to the leading two embedded Isomap time series and reveals bimodality. A two-component bivariate Gaussian mixture model is fitted to the data, and the two monsoon phases are obtained via a composite analysis using the 300 closest points to the centers of this mixture model. The bimodality obtained here is important in understanding monsoon predictability. It suggests, for example, a probabilistic way to define active and break phases. It also suggests that monsoon dynamics may be explained by low-order chaos. In addition, the bimodality can be used to assess interaction between modes of monsoon variability in GCMs.

[20] The first mode corresponds to suppressed monsoon conditions, associated with high SLP anomalies particularly over the monsoon trough regions of the East China Sea and head of the Bay of Bengal. Anticyclonic circulation anomalies over the East China Sea lead to reduced flow across Southeast Asia. The Somali jet is notably weakened and rainfall over India is typical of a monsoon break. The second

mode is associated with enhanced conditions, with low pressure anomalies in the South Asian monsoon trough and East China Sea, where they are particularly deep. The low-level flow is characterized by a cyclonic circulation over these low pressure centers and an anomalously strong Somali jet that extends across Southeast Asia.

[21] We have also investigated the relationship between intraseasonal regimes of preferred monsoon convection and the large-scale seasonal mean monsoon heating using the DMI. Besides the *expected* increase in probability of active and break phases of MISV during strong and weak DMI years, respectively, we still maintain non-negligible probabilities of large events in the other direction. The trend analysis of the ASM indicates that in the second half of the record, there has been a change in the nonlinear structure of the monsoon phases, the active phase being replaced by a dipolar phase which has the structure of active conditions over east China and the East China Sea, and with break conditions over India. This suggests that if the present trend in SLP is to continue, then we may witness a shift in the preferred location of active monsoon convection from India further east, to east China and the East China Sea. This is somewhat consistent with observed trends and modeling projections shown in *Annamalai et al.* [2013], who suggested that enhanced convection in the Western North Pacific was concurrent with sea surface temperature warming there, leading to Rossby-forced circulation changes over South Asia. However, we restate our earlier caution in the interpretation of trends from reanalysis data sets.

References

Annamalai, H., and J. M. Slingo (2001), Active/break cycles: Diagnosis of the intraseasonal variability of the Asian summer monsoon, *Clim. Dyn.*, *18*, 85–102.

Annamalai, H., J. Hafner, K. P. Sooraj, and P. Pillai (2013), Global warming shifts the monsoon circulation, drying South Asia, *J. Climate,* published online, doi:10.1175/JCLI-D-12-00208.1.

Borg, I., and P. Groenen (1997), *Modern Multidimensional Scaling. Theory and Applications*, Springer Verlag, New York.

Brankovic, C., and T. N. Palmer (2000), Seasonal skill and predictability of ECMWF PROVOST ensembles, *Q. J. R. Meteorol. Soc.*, *126*(567), 2035–2067.

Charney, J. G., and J. Shukla (1981), Predictability of monsoons, in *Monsoon Dynamics*, edited by Lighthill, J., and R. P. Pearce, pages 99–109, Cambridge University Press, Cambridge.

Krishnamurthy, V., and J. Shukla (2000), Intraseasonal and interannual variability of rainfall over India, *J. Climate*, *13*(24), 4366–4377.

Krishnamurthy, V., and J. Shukla (2007), Intraseasonal and seasonally persisting patterns of Indian monsoon rainfall, *J. Climate*, *20*(1), 3–20, doi: 10.1175/JCLI3981.1.

Palmer, T. N. (1994), Chaos and predictability in forecasting the monsoons, *Proc. Indian Natl. Sci. Acad.*, *60*, 57–66.

Palmer, T. N. (1999), A nonlinear dynamical perspective on climate prediction, *J. Climate*, *12*(2), 575–591.

Rajeevan, M., J. Bhate, J. D. Kale, and B. Lal (2006), High resolution daily gridded rainfall data for the Indian region: Analysis of break and active monsoon spells, *Curr. Sci.*, *91*, 296–306.

Silverman, B. W. (1981), Using kernel density estimates to investigate multimodality, *J. R. Stat. Soc. Ser. B Methodol.*, *43*(1), 97–99.

Sperber, K. R., and H. Annamalai (2008), Coupled model simulations of boreal summer intraseasonal (30-50 day) variability, part 1: Systematic errors and caution on use of metrics, *Clim. Dyn.*, *31*(2-3), 345–372, doi: 10.1007/S00382-008-0367-9.

Sperber, K. R., J. M. Slingo, and H. Annamalai (2000), Predictability and the relationship between subseasonal and interannual variability during the Asian summer monsoon, *Q. J. R. Meteorol. Soc.*, *126*(568), 2545–2574.

Straus, D. M., and V. Krishnamurthy (2007), The preferred structure of the interannual Indian monsoon variability, *Pure Appl. Geophys.*, *164*(8-9), 1717–1732, International Conference on Mesoscale Processes in Atmosphere Ocean and Environmental Systems FEB 14-17, 2006 Indian Inst Technol, Delhi, India, doi:10.1007/S00024-007-0248-X.

Tenenbaum, J. B., V. de Silva, and J. C. Langford (2000), A global geometric framework for nonlinear dimensionality reduction, *Science*, *290*(5500), 2319–2323.

Turner, A. G., and A. Hannachi (2010), Is there regime behavior in monsoon convection in the late 20th century?, *Geophys. Res. Lett., 37.*, doi: 10.1029/2010GL044159.

Uppala, S. M., et al. (2005), The ERA-40 re-analysis, *Q. J. R. Meteorol. Soc.*, *131*(612), 2961–3012, doi:10.1256/4j.04.176.

Webster, P. J., and S. Yang (1992), Monsoon and ENSO - selectively interactive systems, *Q. J. R. Meteorol. Soc.*, *118*(507), 877–926.

Webster, P. J., V. O. Magana, T. N. Palmer, J. Shukla, R. A. Tomas, M. Yanai, and T. Yasunari (1998), Monsoons: Processes, predictability, and the prospects for prediction, *J. Geophys. Res. Oceans*, *103*(C7), 14451–14510.