Originally published as:

SCIENTIFIC REP⚙RTS

Principal nonlinear dynamical modes of climate variability

Dmitry Mukhin[1], Andrey Gavrilov[1], Alexander Feigin[1], Evgeny Loskutov[1] & Juergen Kurths[2,1]

We suggest a new nonlinear expansion of space-distributed observational time series. The expansion allows constructing principal nonlinear manifolds holding essential part of observed variability. It yields low-dimensional hidden time series interpreted as internal modes driving observed multivariate dynamics as well as their mapping to a geographic grid. Bayesian optimality is used for selecting relevant structure of nonlinear transformation, including both the number of principal modes and degree of nonlinearity. Furthermore, the optimal characteristic time scale of the reconstructed modes is also found. The technique is applied to monthly sea surface temperature (SST) time series having a duration of 33 years and covering the globe. Three dominant nonlinear modes were extracted from the time series: the first efficiently separates the annual cycle, the second is responsible for ENSO variability, and combinations of the second and the third modes explain substantial parts of Pacific and Atlantic dynamics. A relation of the obtained modes to decadal natural climate variability including current hiatus in global warming is exhibited and discussed.

Natural variability plays a key role in climate response to different external forcings including carbon dioxide emission. There are several strong internal modes of the climate system, such as El-Niño Southern Oscillation (ENSO)[1], Pacific decadal oscillation (PDO)[2], Atlantic Meridional oscillation[3], which contribute to both regional and global climate, and have a strong impact on global mean temperatures. Recent papers showed that decadal variability of these modes modulates substantially the anthropogenic growth of global mean temperatures[4–6]; in particular, current hiatus in global warming[6,7] is closely connected to the negative PDO phase[6,8,9] that cooled the Pacific ocean and shifted ENSO variability toward domination of La-Niña conditions. Currently it is well understood that the ability to predict the behavior of such natural modes is crucial for correct global climate projections; however, reflecting the decadal variability by climate models is not confidence enough[6,10].

In this work an alternative, empirical way to analyze natural modes of climate is proposed: we attempt to resolve empirically dominant hidden signals which actually govern the observed behavior, so as they could be suitable for a low-dimensional dynamical description of climate variability on the considered time scales. In fact, our aim is to reduce the data dimension by means of expanding measured spatio-temporal observations into a few number of modes, which would form a basis for a dynamical system's phase space construction capturing main features of the observed dynamics.

There is a set of widely used approaches for obtaining principal components by various linear rotations of multivariate data in space domain (see overview of various methods[11,12]); they include both traditional empirical orthogonal function (EOF) decomposition and more advanced rotations, which were successfully applied in climate science: in particular, empirical orthogonal teleconnections[13] and varimax rotation[14]. Also, there are several spatio-temporal extensions of those techniques, based on multichannel singular spectral analysis (MSSA)[15–17] taking into account time-lag correlations in data. In recent papers advances of MSSA-based expansion for empirical forecast of complex system[18,19] as well as for studying synchronization and clustering in multivariate dynamics[20] were demonstrated. Papers[21,22]

[1]Institute of Applied Physics of the Russian Academy of Sciences, 603950 Nizhny Novgorod, Russia. [2]Potsdam Institute for Climate Impact Research, 14412 Potsdam, Germany. Correspondence and requests for materials should be addressed to D.M. (email: mukhin@ipfran.ru)

suggest a method for principal mode extraction combining data rotation and system's evolution operator construction.

A quite apparent inherent drawback of such rotations is regarded to their linearity: many so obtained components contain fractions of the same time scales, or, in other words, modes corresponding to different time scales are usually mixed in different components. Indeed, this effect is essential when there is an evidence of strong nonlinear couplings between time series at different grid points; see the model example in Supplementary Methods. Here we suggest a way for revealing such nonlinear couplings and constructing nonlinear transformation of data to new variables – the nonlinear dynamical modes (NDMs), that is hidden time series holding an essential part of the dynamics. The suggested approach is based on a direct analysis of the data without any prior knowledge about first-principal models. This is in fact a nonlinear generalization of EOF decomposition: instead of constructing linear constraints in the data space, we try to resolve NDMs as most principal nonlinear manifolds (curves) and project data on them. Actually, it is a sort of construction of principal curves[23] allowing an expansion of vector time series $\mathbf{X}(t) \; \mathbb{R} \in^{D}$ (e.g. discretized climatic spatial field) into the set of dominant parametrically defined curves

$$\mathbf{X}(t) = \mathbf{F}_1(p_1(t)) + \dots + \mathbf{F}_q(p_q(t)) + \zeta(t), \tag{1}$$

each of which is a map of some hidden (latent) value of $p_i(t) \in \mathbb{R}$ into the data space $\mathbb{R}^D$. Note that the used method allows reconstructing both latent time series $p_i(t)$ interpreted as internal modes driving observed variability, and maps $\mathbf{F}_i$ needed for analysis of NDMs represented on the data (geographic) grid (hereinafter we will use the abbreviation NDM for both time series $p_i(t)$ and its image $\mathbf{F}_i(p_i(t))$ in data space). This is in contrast to isomap methods based on principal component analysis (PCA) kernels[24], which recently found application to the analysis of climate data[25,26]: they permit only obtaining principal nonlinear components without an inverse transformation.

A geometrical interpretation of a NDM is also useful: generally, each function $\mathbf{F}_i(.)$ in (1) defines a nonlinear curve in $\mathbb{R}^D$. Then the points of the corresponding time series $p_i(t_1), \dots, p_i(t_N)$ are just coordinates of N points along such a curve. These points are specific projections of the observed data (see Supplementary Figure S5 for illustration).

We developed a Bayesian procedure for a reconstruction of principal modes defined by (1); see details in the section Methods and in Supplementary Materials. Note that the proposed method does allow us to fit some predefined functions $\mathbf{F}_i(.)$ as well as to find values of hidden time series p. Also, it includes very important step, which prevents over-fitting: selection of optimal complexity of decomposition (1), from a statistical point of view. This step is based on Occam Razor principal stating that minimal but explaining things model is the best one: it answers the question of how many nonlinear modes can be resolved from available data. Moreover, it yields the optimal nonlinearity degree of the functions $\mathbf{F}_i(.)$ as well as prior correlation properties of p-time series, which control smoothness of the obtained principal curves in data space. In practice, such prior gives us an efficient separation of time scales between NDMs; hence, different NDMs turn out to be responsible for different modes of natural variability. As a result, we obtain a set of statistically significant nonlinear structures $\mathbf{F}_i(.)$ and "noise" $\zeta(t)$ – the time series of the residual field. In the next section results of our nonlinear expansion of SST time series is demonstrated and analyzed.

Results

Observed time series and data expansion. For analysis we took the time series of NOAA OI.v2 SST monthly field[27,28]. This space-distributed time series covers the globe with a resolution of $1° \times 1°$ and has a duration of a bit more than 33 years – from Nov. 1981 to present. Altogether we have 44,219 SST signals in different grid points and our aim is to describe them by a moderate number of principal nonlinear dynamical modes (NDMs) p – the set of scalar time series capturing the more the better part of observed variability. To this end, we applied our expansion defined through (1) to this data using the methodology expounded briefly in Methods and in more detail in Supplementary Materials.

We applied our decomposition and found that three principal nonlinear modes can be identified; the corresponding three time series $p(t)$ are shown in Fig. 1a in order of descending (from top to bottom) energy, i.e. captured variance of SST. The line thickness on this figure reflects the 95% Bayesian confidence interval characterizing the uncertainty of the hidden state $p(t)$ estimation (see Supplementary Methods for details). Further expansion of the rest of data, which is not explained by these three nonlinear modes, comes down to traditional EOF decomposition, simply revealing linear relationships between time series. Comparing the amount of total variance of data captured by NDMs vs. traditional linear principal components (PCs) obtained by projecting data to the leading EOFs, we find that the three NDMs explain 85% of whole data variance – the same amount as from approximately 6 linear PCs; see Fig. 1b displaying cumulative variances captured by corresponding number of principal modes. Figure 1c shows the contribution of each NDM to data expanded via EOFs: the rates of presence of NDMs in different linear PCs are plotted. It is clear that each NDM captures its own fracture of PCs; essentially, only about twenty PCs participate in the nonlinear expansion: the subspace determined by them holds all nonlinearity we are able to extract from the data.
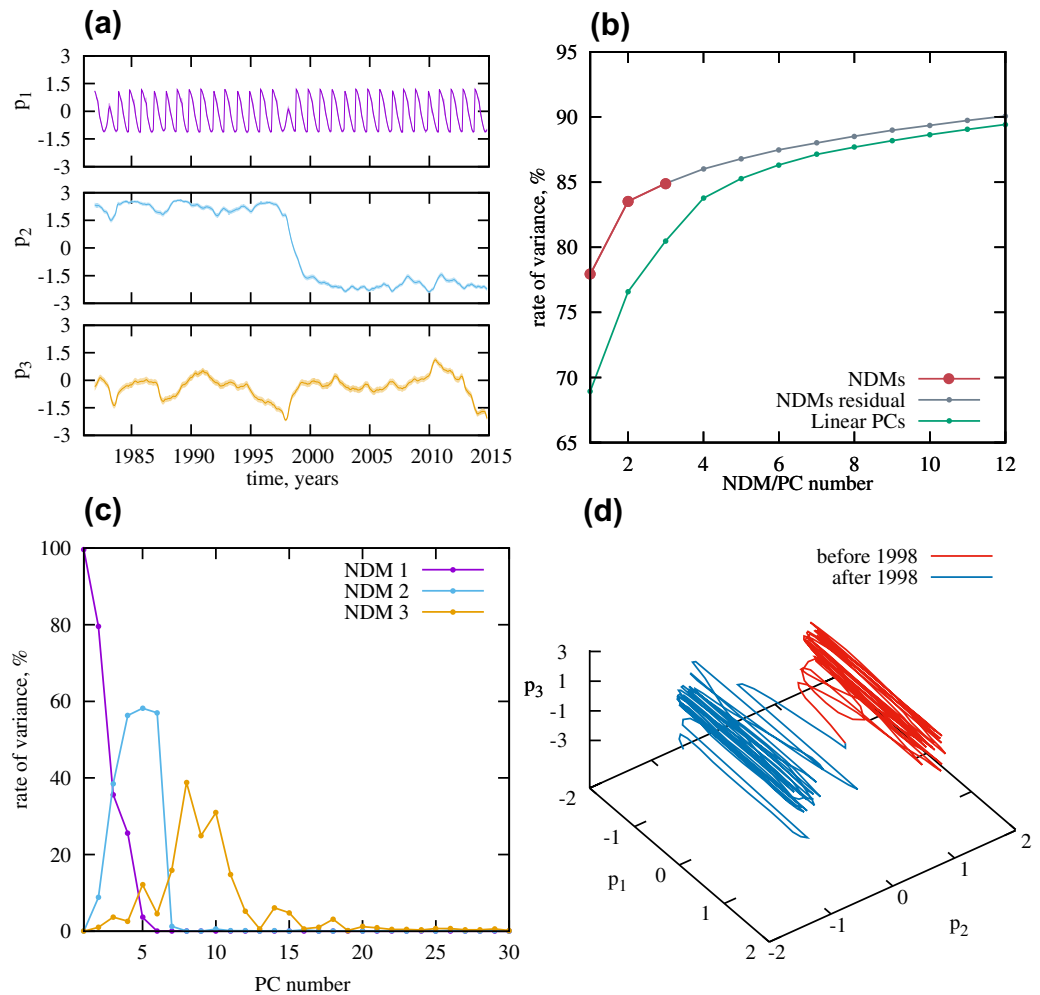
2

Figure 1. Nonlinear dynamical modes (NDMs) reconstructed from data. (**a**) Hidden time series $p(t)$ of NDMs placed in order of descending captured variances. (**b**) Rate of cumulative variances captured by modes. Big red points correspond to NDMs, gray dots – to principal components (PCs) obtained from linear EOF decomposition of residuals. Green dots mark the rate of the variances captured by linear PCs of the data. (**c**) Contributions of NDMs to different linear PCs (rates of variance are shown). (**d**) Projection of system's phase space formed by the reconstructed variables p_1, p_2, p_3. Blue color corresponds to the epoch before 1998, red color – after 1998. Gnuplot (Copyright 1986–1993, 1998, 2004 Thomas Williams, Colin Kelley) was used to create the graphs in this figure.

Since for every ith NDM we have the corresponding function $\mathbf{F}_i(.)$ mapping scalar values of $p_i(t)$ into the data space, it enables us to visualize the time series of NDMs on the geographic grid. Such a visualization is presented in the Supplementary Animation as well as snapshots (see Fig. 2a–c) for each of the three obtained NDMs separately. Also, in both this animation and Fig. 2d, the time series of the rest of data, which is not explained by three modes is shown. The first mode oscillates with approximately a 1-year period, clearly, this reflects the annual cycle in the data. The second and third modes have much larger time scales of order of ENSO and PDO variability; the second mode is apparently ENSO-related: it is excited with different signs during El-Niño and La-Niña events (see, for instance, Fig. 2 with the snapshot of the time series at time instant corresponding to El-Niño episode in 1992). The rest of the data (panels d in both Supplementary Animation and Fig. 2) behaves with shorter than 1 year character times; actually, it is treated as noise signal by our procedure.

Analysis of principal modes. Let us look how the identified modes are manifested in different geographical regions and which climatic phenomena they reflect. We distinguished several SST-based indices corresponding to different regions in the Northern hemisphere, that have a significant impact on the global climate (see Table 1 for details): Nino 3.4 — ENSO SST-based index[1], Pacific Decadal Oscillation (PDO) index derived as a projection of data on leading SST anomaly (SSTA) EOF in Northern Pacific area[29], North Tropical Atlantic (NTA)[30] and Indian Ocean Dipole (IOD)[31] indices. Additionally, we
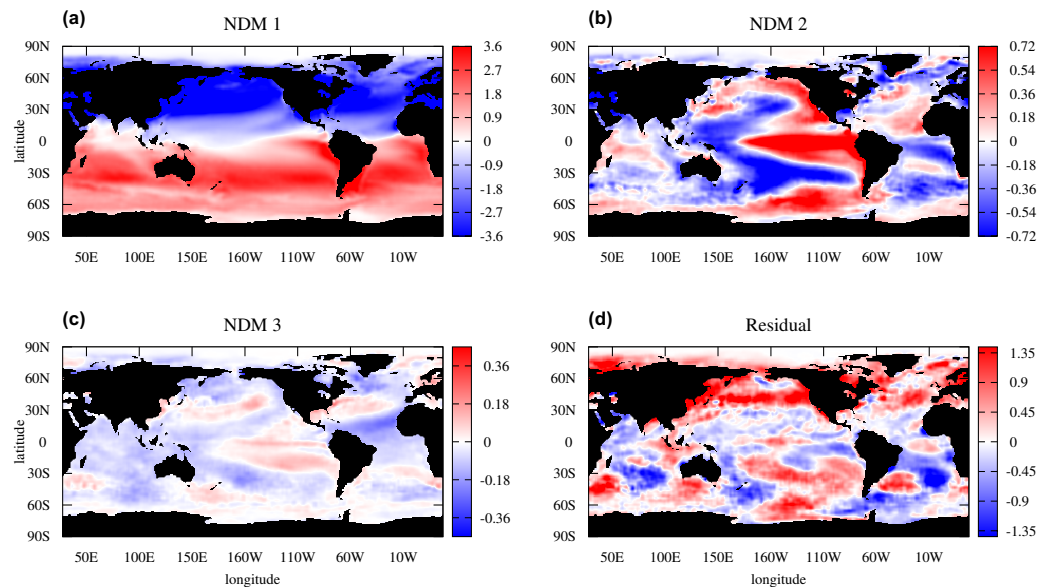
Figure 2. Representation of NDMs in data space. Snapshot corresponding to Jan. 1992 is shown. Captured fractions of SST are marked by color at each grid point. Modes 1, 2 and 3 are displayed on panels (**a–c**). Panel (**d**) shows the rest of the data that is not involved in the nonlinear expansion. Gnuplot (Copyright 1986–1993, 1998, 2004 Thomas Williams, Colin Kelley) was used to create the maps in this figure.

Index	Region	Description	Correlation with SSTA-based index
Nino 3.4	[5°S–5°N; 190°E–120°W]	Averaged data	0, 88
PDO	[20°N–60°N; 140°E–120°W]	Projection on largest SSTA EOF	0, 75
NTA	1. [6°N–18°N; 60°W–20°W] 2. [6°N–10°N; 20°W–10°W]	Averaged data over both 1 and 2 regions	0, 74
NA	[45°N–70°N; 60°W–0°W]	Averaged data	0, 73
IOD	1. [10°S–10°N; 50°E–70°E] 2. [0°N–10°N; 90°E–110°E]	Difference between averaged data in regions 1 and 2	0, 38

Table 1. Indices derived from SST field reconstructed from obtained NDMs. Correlations of the indices with "true" ones, calculated from full SSTA field, are presented in the last column.

introduced the North Atlantic (NA) index characterizing the temperature of the northern Atlantic ocean; it is defined as averaged SST anomalies over a certain rectangle area specified in Table 1.

We calculated all these indices from each of three NDMs. The largest (first) NDM gives an almost periodical, slightly modulated signal in every NDM-based index (see Supplementary Figure S1). It corresponds to the annual cycle – the response of the climate to the 1-year periodic forcing. In doing so, the annual cycle is completely separated from data: there is no trace of harmonics of the 1-year periodical signal in the power spectra of the other modes and it is obtained in a natural way, without any periodical constraints in the procedure (see figures in Supplementary Figure S2). As a confirmation, Fig. 3a shows a clear global annual cycle pattern as a result of the difference of the average SST produced by the first NDM between summer and winter months during the whole period of observations.

The left panels of Fig. 4 displays the NDM-based indices calculated from the combinations of NDMs yielding highest correlations with the corresponding original indices (black lines) derived from SSTA time series. Corresponding correlation coefficients are presented at the right column of Table 1. Purple columns on right panels of Fig. 4 show correlations of indices obtained from different NDMs (mean annual periodical signal was removed from the first NDM before calculations) with certain original index. For comparison, correlations of different linear PCs (after removing mean annual signal) with the original indices are shown as blue transparent columns on the same panels. We infer that the second NDM contributes, in a first place, to ENSO dynamics: it provides a high correlation of 0.88 with the original Nino 3.4 index. Note that the first NDM, which is responsible mainly for the annual cycle, also participates in ENSO variability having little more than 0.5 correlation with Nino 3.4 (see also Supplementary Figure S3). Moreover, it is seen from Fig. 1a that the hidden time series of the first and second NDMs are
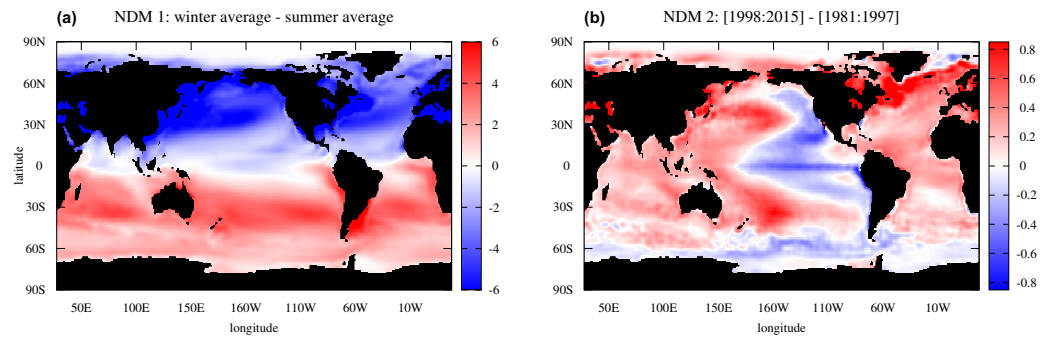
Figure 3. Spatial patterns explained by NDM 1 and NDM 2. (**a**) Annual cycle pattern: difference between SST captured by NDM 1 averaged over winter and summer months during the entire time interval of data. (**b**) Negative PDO pattern: difference between SST captured by NDM 2 averaged over the intervals [1998–2015] and [1981–1997]. Gnuplot (Copyright 1986–1993, 1998, 2004 Thomas Williams, Colin Kelley) was used to create the maps in this figure.
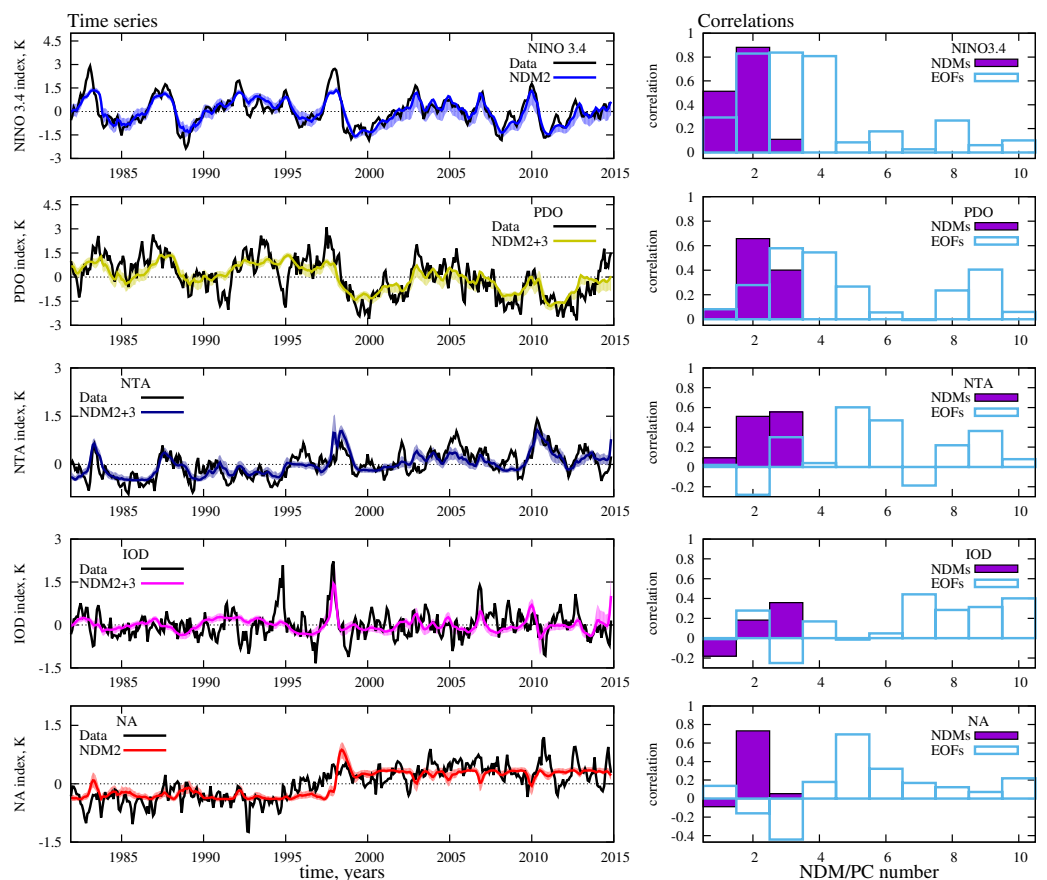

Figure 4. Reconstruction of climate indices enumerated in Table 1. Left panels: comparing original indices (black color) calculated from the SSTA field and those reconstructed from different combinations of NDM 2 and NDM 3. Bayesian confidence intervals of the reconstruction (see Supplementary Methods) are shown by shadow around each reconstructed index. Right panels: purple columns show correlations of certain index (see legend on figures) with different NDMs (number of NDM is on the abscissa). For comparison, correlations of the same indices with linear PCs – projections of data on EOFs – are shown by blue transparent columns. Gnuplot (Copyright 1986–1993, 1998, 2004 Thomas Williams, Colin Kelley) was used to create the graphs in this figure.

connected with each other: there are apparent jumps of both phase and amplitude in the almost periodic time series of the first NDM, which occurred during strong El-Niño events of 1982–83 and 1997–98. This fact can be interpreted as an evidence of a nonlinear interaction between ENSO and annual climate

response – two modes with substantially different time scales. See also the comparison of ENSO indices derived from the first vs. second NDM in supplementary Figure S3. Anyhow, it is clear that the second NDM is dominant in ENSO variability; a similar situation takes place in the NA region: this NDM has a much higher contribution there than other NDMs. As for the two other considered indices – PDO and NTA – the third NDM becomes important as well: the second and the third NDMs together explain a substantial part of the SST variability in both regions. The Indian ocean dynamics is reflected less well: the three obtained NDMs give only a 0.38 correlation with the original IOD index.

However, contributions to the considered regions of linear PCs are spread over many PCs, as it can be seen from the right panels of Fig. 4. For instance, there are at least three PCs equally significant for ENSO dynamics as well as many PCs that contribute to PDO and Atlantic regions. Indeed, this is the result of linearity of the EOF-based rotation of data, leading to a mixing of different modes in the obtained components in case when nonlinear couplings are essential. Supplementary Figure S4 shows an advantage of our nonlinear approach over the traditional EOF decomposition to separate different time scales by reconstructed modes.

Climate shift of 1997–1998.   The most interesting event that plumped into our analyzed epoch is the 1997–1998 El-Niño episode – one of the strongest since records began (see the behavior of Nino 3.4 index on top panel of Fig. 4). There is an attractive hypothesis[6] that this event plays a decisive role in both switching over the climate to a negative phase of PDO and a successive drastic change of the global warming scenario. Soon after the event a growth of the average atmospheric temperature had been suspended setting up the current hiatus in global warming[6,7]; before this the atmosphere was heated since the mid seventies during the positive PDO phase. It was shown[6,8–10] that the significant part of heat goes into deep ocean during negative PDO that prevents the atmosphere from warming; at the same time melting ice and rising sea level is continued[32] indicating that the global warming is in progress. But the regime of the warming was changed and, to all appearances, this is the result of internal dynamics of the climate system.

This climate shift manifests itself in our second NDM as a distinct jump in its hidden time series $p(t)$ (Fig. 1a) on time interval comparable with one year. This is reflected in the time series of PDO and two Atlantic indices reconstructed from this NDM, as an abrupt change of averages (see corresponding panels of Fig. 4). Since this event is associated with a PDO phase change, we see how our second NDM, where the corresponding shift is observed, captures the PDO pattern. To this end, we calculated averages of SST captured by this single mode, over the period both before and after the shift, and then plotted the difference between these two patterns (Fig. 3b). The negative PDO pattern is manifested clearly in the Pacific region, that is in good correspondence with figure 9 from paper[6], where similar differences are shown, but calculated from full surface temperature data. Thus, we can conclude that the second NDM, driven by only a single scalar time series, properly reflects the nature of the climate shift and confirms the hypothesis of the triggering role of the 1997–98 El-Niño episode[6].

Determining teleconnections.   Detecting principal curves in data space allows investigating teleconnections – the linkages of dynamics at widely-spaced regions of the globe (this is a typical notion for long-range connections in climatology[33]). Really, the second NDM captures about 7% of the whole observed SST variability. If we exclude from consideration the annual cycle, contained in the first NDM that holds about 78%, we see that the second NDM holds about 27% of SST anomalies (in this case the anomalies are regarded as data after subtraction of the first NDM). But this mode attaches the dynamics at all grid points to a one-dimensional curve and therefore, different points, and hence, different regions and different SST-based indices are coupled by this mode in a very simple way. In Fig. 5 (points) the curves derived from the second NDM become exhibited: they determine the dominant dependencies between the behavior in the Nino 3.4 region and other considered regions. They all consist of two branches – one corresponds to the period before the shift and the other – after the shift; note, that the shift led to significant changes in these curves. Adding the third NDM gives a bit spread dependencies, but the main shapes of them remain the same: thin lines in Fig. 5 denote couplings derived from the sum of the second and third NDMs which together hold more than 30% of SST anomalies.

Discussion

In this work we expanded space-distributed global SST time series into a set of principal nonlinear modes. Each of these modes couples the dynamics at all grid points by a parametrically defined curve in data space. The proposed method uses *a priori* information about the dynamical nature of the reconstructed components by using certain restriction on the time scales inherent of the modes. The complexity of such nonlinear transformations as well as the characteristic time scale of every mode are both estimated to be relevant to available data: the Bayesian approach to optimal model selection is used. Moreover, the number of nonlinear modes we are able to resolve from data is also derived from a Bayesian criterion, that gives us a set of statistically significant nonlinear components driving the observed variability. The analysis of the 33-year time series of the SST field yielded three nonlinear modes underlying the dynamics, each of them is determined by single scalar time series. The most powerful one is associated with the annual cycle; it completely separates the 1-year oscillation from the rest of the data. The second mode describes a major part of ENSO-related variability. Eventually, combinations of the second and the third
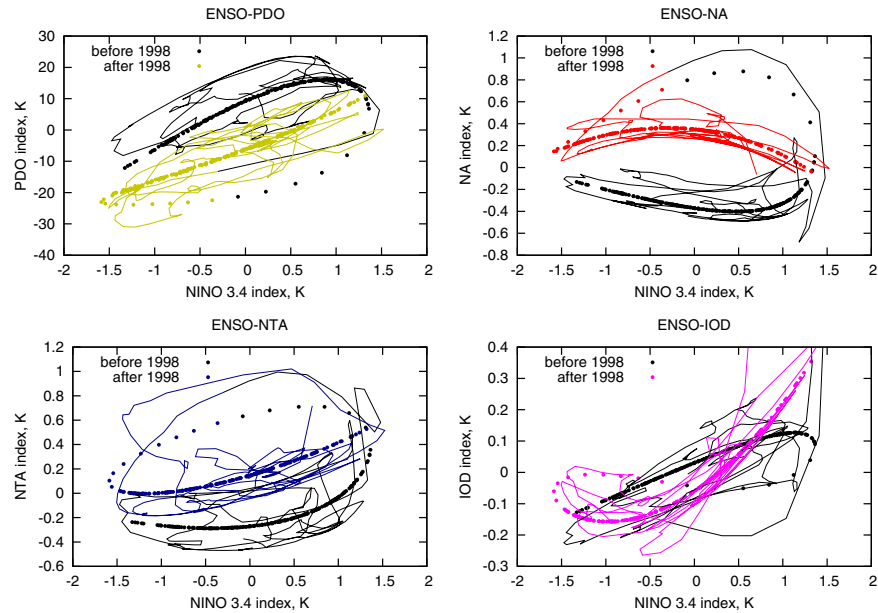
Figure 5. Couplings between ENSO and dynamics in other regions, revealed by NDMs. The dependencies derived from NDM 2 are shown by points, those reconstructed from superposition of NDM 2 and NDM 3 – by thin lines. Black color embraces the period before the 1997–1998 climate shift, other colors – after the shift. Colors after the shift are the same as in Fig. 4; they mark the coupling of ENSO and corresponding indices from that figure. Gnuplot (Copyright 1986–1993, 1998, 2004 Thomas Williams, Colin Kelley) was used to create the graphs in this figure.

modes explain PDO, NTA and Northern Atlantic SST anomalies: they give more than 0.7 correlations with the corresponding indices.

The obtained expansion is interesting from the point of view of global climate analysis including a description of different phases of global warming regime. Really, the resulting NDMs reflect three main components of global climate variability, which have characteristic time scales less than the duration of the analyzed time series: annual cycle, ENSO mode and decadal mode associated with PDO behavior. In particular, the single second NDM correctly reproduces the climate shift of 1997–98–the change of PDO phase: the nature of the shift is clearly manifested in SST variability captured by this NDM. Further, if we consider the set of time series p driving the obtained NDMs as a source for phase variables of some dynamical subsystem governing the observed variability, it will be possible to construct the phase space in different ways. Figure 1d displays a projection consisting of three time series $p_1(t)$, $p_2(t)$, $p_3(t)$. The climate shift in this space looks like a transition between two areas. However, we analyzed here quite short time series, which includes only the single transition. Further expansion of longer data taking into consideration several such climate shifts in the past–including PDO transitions from a positive to a negative phase and back–would be the proper step on the way of constructing a dynamical model suitable for forecasting future events on decadal scales, which strongly affect global warming scenario. That is a challenging current problem in climate science: modern global climate models hardly predict internal decadal variability[6], including current hiatus in global warming.

Methods
According to Eq. (1), the main aim of the proposed expansion is to decompose multivariate (space-distributed) time series consisting of points $\mathbf{X}(t_n):=\mathbf{X}^n$ into a set of principal modes, each of them is represented as a one-dimensional parametrically defined curve $\mathbf{F}_i(p^n):=\mathbf{F}_i^n$. Below, we state briefly the main points of the method we suggest; more technical details can be found in the Supplementary Methods.

Bayesian reconstruction. As in traditional EOF decomposition, we use a recursive procedure for a successive reconstruction of terms (modes) in Eq. (1). The first step is finding the leading mode $\Phi:=\mathbf{F}_1$, having the largest contribution to data variance, by minimization of mean square deviations of residuals between data vectors $\mathbf{X}^n$ and mode values Φ^n. The main problem here is the extremely high dimension of climate data defined on spatial grid, which makes such a minimization very hard computationally. Therefore, a preliminary data truncation would be reasonable. For this purpose we firstly rotate the data to the basis of EOFs which gives the set of PCs – the new variables $\mathbf{Y}^n=\mathbf{V}^T\mathbf{X}^n$ – which are ranked in accordance with their variances (EOFs are the columns of the orthogonal matrix $\mathbf{V}$). After such a rotation

the main problem is to determine the subspace in the whole space of PCs, where the sought nonlinear curve $\mathbf{F}_i(.)$ actually lies. Let us then construct the nonlinear transformation only in the space of the d first PCs, while the other PCs are treated as noise:

$$
\begin{aligned}
Y_1^n &= \Phi_1(p^n, \mathbf{a}) + \zeta_1^n \\
&\vdots \\
Y_d^n &= \Phi_d(p^n, \mathbf{a}) + \zeta_d^n \\
Y_{d+1}^n &= \zeta_{d+1}^n \\
&\vdots \\
Y_D^n &= \zeta_D^n.
\end{aligned}
\tag{2}
$$

Here $\mathbf{a}$ are the unknown parameters in the Φ representation, D is the total number of PCs, ζ are the residuals, which are assumed to be Gaussian and uncorrelated. Thus, for reconstruction of this leading mode we should find proper values of both the latent variables p^n and the parameters $\mathbf{a}$. Note that it is easy to show that in the special case of linear Φ this strictly corresponds to the traditional EOF rotation: minimization of squared residuals under condition $|\mathbf{a}| = 1$ gives the largest EOF in the $\mathbf{a}$ vector and the corresponding PC in p time series.

In the framework of the Bayesian paradigm the cost-function used for learning model (2) is constructed in a form of a probability density function (PDF) of unknown variables under the condition of data $\mathbf{X}^1, ..., \mathbf{X}^N$: the learning model (2) means finding the global maximum of this PDF over unknown $\mathbf{a}$ and $p^1, ..., p^N$. This PDF can be expressed through the Bayes theorem:

$$
P(\mathbf{a}, p^1, ..., p^N | \mathbf{X}^1, ..., \mathbf{X}^N) \propto P(\mathbf{X}^1, ..., \mathbf{X}^N | \mathbf{a}, p^1, ..., p^N) \times P_{pr}^p(p^1, ..., p^N) P_{pr}^a(\mathbf{a}).
\tag{3}
$$

Here $P(\mathbf{X}^1, ..., \mathbf{X}^N | \mathbf{a}, p^1, ..., p^N) := L(\mathbf{a}, p^1, ..., p^N)$ is a likelihood function, i.e. the probability that the data $\mathbf{X}$ are obtained by both the model (2) with the parameters $\mathbf{a}$ and the series $p^1, ..., p^N$, and the subsequent rotation by matrix $\mathbf{V}$. This function can be easily written under the assumption of normality and whiteness of noise ζ (see Supplementary Methods for details); its maximum is reached at minimal mean square errors of ζ. But the crucial point is the specification of a priori information, which reflects all desirable properties of the solution, in a form of prior probability densities of $\mathbf{a}$ and $p^1, ..., p^N$.

The main *a priori* assumption we suggest to use is the dynamical nature of the hidden mode p time series; in other words, we require every current point p^n to be connected with the previous one p_{n-1}. Namely, we introduce a prior probability density for $p^1, ..., p^N$ in the following form:

$$
P_{pr}^p(p^1, ..., p^N) \propto \exp\left[-\frac{1}{2\sigma_p^2}\left\{ \sum_{n=1}^{N-1}(p^{n+1} - bp^n)^2 \right\} - \frac{1}{2}p^{1^2} \right], \quad \sigma_p^2 + b^2 = 1
\tag{4}
$$

Actually, Eq. (4) determines a prior ensemble of p time series consisting of various realizations of colored noise so as most of them have an auto-correlation time $\tau > -\frac{1}{\log b}$. The second condition in (4) provides unit prior variance of each $p^n \sim N(0, 1)$. Though the ensemble so defined is quite general, it provides an efficient restriction of the class of possible solutions by excluding from consideration short-scale signals.

Under the prior constraint (4) we introduced a representation of the functions $\Phi(.)$ in (2) as a superposition of polynomials $\Pi_i(p)$ which are mutually orthogonal in the probability measure defined by (4):

$$
\Phi_i(p) = \alpha_i^0 \Pi_0(p) + \alpha_i^1 \Pi_1(p) + ... + \alpha_i^m \Pi_m(p)
$$

$$
\int \Pi_i(p) \Pi_j(p) P_{pr}^p(p)\, dp = \delta_{ij}
$$

$$
\mathbf{a} = (\alpha_1^0, ..., \alpha_d^0, \alpha_1^1, ..., \alpha_d^1, ..., \alpha_1^m, ..., \alpha_d^m),
\tag{5}
$$

where δ_{ij} is Kronecker delta. Such a representation facilitates a model learning procedure substantially, since the problem becomes linear with respect to the parameters $\mathbf{a}$. At the same time it allows an increasing model power by simply adding more orthogonal polynomials in (5). The idea how to set a prior PDF for the parameters of such a representation is quite apparent: it should be the most general, but permitting functions $\Phi_i(.)$ to have *a priori* the same variances as the corresponding time series of PC Y_i (correlations between different PCs Y_i are zero, so we need not to care about covariances). Thus, it is reasonable to define this PDF as a product of Gaussian functions of each α_i^j with variance $\sigma_{\alpha_i^j} = \frac{\langle Y_i^2 \rangle}{m+1}$, where $\langle . \rangle$ denotes average over time. The same restrictions of parameters were used, for instance, in[34,35] regarding to external layer coefficients of artificial neural networks used for fitting an evolution operator.

Optimal model selection.   It is clear that the proposed method provides a solution that strongly depends on three parameters: d - the number of PCs involved in the nonlinear transformation, m - the number of orthogonal polynomials determining the degree of nonlinearity, and τ - the characteristic autocorrelation time of the reconstructed mode. All these values determine the complexity of the data transformation, and therefore should be relevant to the available statistics. For example, if we take a very large m and $\tau = 0$, we would obtain the curve passing through every point Y_1^n, ..., Y_d^n, and our single mode p^1, ..., p^N would capture the whole variability in the subspace of the d first PCs. But, indeed, in this case we would get an overfitted, statistically unjustified model. Thus, a criterion of optimality is needed allowing the selection of the best model from the model set defined by different values of (d, m, τ). We define optimality E through a Bayesian evidence function, i.e. the probability density of data $\mathbf{X}^1$, ..., $\mathbf{X}^N$ given the model (d, m, τ):

$$E = -\log P\left(\mathbf{X}^1, \ ..., \mathbf{X}^N | d, \ m, \ \tau\right). \tag{6}$$

Actually, E is the minimal amount of information required for transferring the given series $\mathbf{Y}^1$, ..., $\mathbf{Y}^N$ by the model (d, m, τ); so, the same criterion could be derived in the framework of the minimal description length principle[36]). The model corresponding to the smallest E, or, which is the same, providing the maximum probability of reproducing the observed time series, is regarded to be optimal.

The Bayesian evidence function can be calculated by the integral of likelihood over all parameters $\mathbf{a}$ and latent variables p^1, ..., p^N with the prior probability measure:

$$P\left(\mathbf{X}^1, \ ..., \mathbf{X}^N | d, \ m, \ \tau\right) = \int L\left(\mathbf{a}, p^1, \ ..., p^N\right) P_{pr}^p\left(p^1, \ ..., p^N\right) P_{pr}^a\left(\mathbf{a}\right) d\mathbf{a} dp^1 ... \ dp^N. \tag{7}$$

We calculate this integral by the Laplace method assuming that the main contribution comes from the neighborhood of the maximum of the integrand over all variables; see details in Supplementary Methods. The eventual optimality can be expressed in the following form:

$$E = \psi\left(\overline{\mathbf{a}}, \overline{p}_1, \ ..., \overline{p}_N\right) + \frac{1}{2} \log\left|\nabla_p \nabla_p^T \psi\left(\overline{\mathbf{a}}, \overline{p}_1, \ ..., \overline{p}_N\right)\right|$$

$$+ \frac{1}{2} \log\left|\nabla_a \nabla_a^T \psi\left(\overline{\mathbf{a}}, \overline{p}_1, \ ..., \overline{p}_N\right)\right|$$

$$\psi\left(\overline{\mathbf{a}}, \overline{p}_1, \ ..., \overline{p}_N\right) = -\log L\left(\overline{\mathbf{a}}, \overline{p}_1, \ ..., \overline{p}_N\right) - \log P_{pr}^p\left(\overline{p}_1, \ ..., \overline{p}_N\right) - \log P_{pr}^a\left(\overline{\mathbf{a}}\right)$$

$$\left(\overline{\mathbf{a}}, \overline{p}_1, \ ..., \overline{p}_N\right) = \text{argmin}\left[\psi\left(\mathbf{a}, p^1, \ ..., p^N\right)\right]. \tag{8}$$

Here $\left(\overline{\mathbf{a}}, \overline{p}_1, \ ..., \overline{p}_N\right)$ are estimated values corresponding to the minimum of the function $\psi(.)$ – minus log of the cost-function; $\nabla_p \nabla_p^T$ and $\nabla_a \nabla_a^T$ denote operators of second derivations over all p^n and $\mathbf{a}$ correspondingly. The last two terms in (8) penalize the model complexity; they provide the existence of an optimum of E over the model set. Thus, for estimating the optimality of the given model with (d, m, τ), we should (i) learn the model by finding the values $\left(\overline{\mathbf{a}}, \overline{p}_1, \ ..., \overline{p}_N\right)$ and then (ii) calculate penalty terms as determinants of $\psi(.)$ second derivation matrices in the domains of parameters $\mathbf{a}$ and latent variables p^1, ..., p^N in the point $(\overline{\mathbf{a}}, \overline{p}^1, \ ..., \overline{p}^N)$.

Main steps of the algorithm.   The main steps of the proposed algorithm are the following.

1. Rotating given data $\mathbf{X}^n \in \mathbb{R}^D$ to PCs $\mathbf{Y}^n \in \mathbb{R}^D$: $\mathbf{Y}^n = \mathbf{V}^T \mathbf{X}^n$; columns of the $\mathbf{V}$ matrix are EOFs of the $\mathbf{X}$ time series.
2. Finding the optimal model in the space of PCs $\mathbf{Y}$. This step includes finding the optimal dimension d of the subspace for a nonlinear expansion, estimation of the optimal degree of nonlinearity m as well as characteristic autocorrelation time τ of the hidden mode. Here we also obtain concrete parameters $\left(\overline{\mathbf{a}}, \overline{p}_1, \ ..., \overline{p}_N\right)$ of the optimal mode.
3. Subtracting the obtained mode corresponding to the maximum of the cost-function from the data vectors:

$$\mathbf{X}_{\mathbf{new}}^n = \mathbf{X}^n - \mathbf{V} \begin{pmatrix} \Phi_1\left(\overline{p}_n, \overline{\mathbf{a}}_1\right) \\ \vdots \\ \Phi_d\left(\overline{p}_n, \overline{\mathbf{a}}_d\right) \\ 0 \\ \vdots \\ 0 \end{pmatrix} \tag{9}$$

The vector of the obtained mode image in the PC space is complemented by $D-d$ zeros, since only d PCs are involved in the nonlinear transformation. In fact, at this step we set the new data vectors $\mathbf{X}_{\mathrm{new}}^n$ – the residuals between initial ones and values of the mode. After that we are going to find the next mode: the same procedure beginning from i.1 is applied to the new data $\mathbf{X}^n = \mathbf{X}_{\mathrm{new}}^n$.

The iteration of the procedure i.1–i.3 is stopped when we obtain time series p of the new mode equal to constant zero, or, equivalently, we find that the optimal degree m of polynomials (5) is equal to zero. It means that we cannot resolve any more nonlinearity in the data, and the noise is most probably significant; in other words, all the best we can further do is traditional EOF decomposition of the residuals. In particular, the expansion of SST time series into three NDMs presented in the current paper gives $d = 5$, $m = 3$ for the first NDM, $d = 4$, $m = 6$ for the second and $d = 8$, $m = 8$ for the third. Eventually, the entire nonlinear expansion of data can be written as follows:

$$\mathbf{X}^n = \mathbf{V}_1 \overline{\mathbf{F}}_1\left(p_1^n\right) + \ldots + \mathbf{V}_q \overline{\mathbf{F}}_q\left(p_q^n\right) + \zeta^n, \qquad (10)$$

where q is the total number of obtained NDMs. Note that each of the NDMs is defined in its own subspace, which is rotated relative to the initial data space: the orthogonal matrices $\mathbf{V}_i$ are different for different NDMs.

References

1. Trenberth, K. E. The Definition of El Niño. *Bulletin of the American Meteorological Society* **78,** 2771–2777, doi: 10.1175/1520-0477(1997)078%3C2771:tdoeno%3E2.0.co;2 (1997).
2. Mantua, N. J., Hare, S. R., Zhang, Y., Wallace, J. M. & Francis, R. C. A Pacific Interdecadal Climate Oscillation with Impacts on Salmon Production. *Bulletin of the American Meteorological Society* **78,** 1069–1079, doi: 10.1175/1520-0477(1997)078%3C1069:apicow%3E2.0.co;2 (1997).
3. Knight, J. R., Folland, C. K. & Scaife, A. A. Climate impacts of the atlantic multidecadal oscillation. *Geophysical Research Letters* **33,** L17706, doi: 10.1029/2006GL026242 (2006).
4. Wu, Z., Huang, N. E., Wallace, J. M., Smoliak, B. V. & Chen, X. On the time-varying trend in global-mean surface temperature. *Climate Dynamics* **37,** 759–773, doi: 10.1007/s00382-011-1128-8 (2011).
5. DelSole, T., Tippett, M. K. & Shukla, J. A Significant Component of Unforced Multidecadal Variability in the Recent Acceleration of Global Warming. *J. Climate* **24,** 909–926, doi: 10.1175/2010jcli3659.1 (2010).
6. Trenberth, K. E. & Fasullo, J. T. An apparent hiatus in global warming? *Earth's Future* **1,** 19–32, doi: 10.1002/2013EF000165 (2013).
7. Easterling, D. R. & Wehner, M. F. Is the climate warming or cooling? *Geophysical Research Letters* **36,** L08706, doi: 10.1029/2009GL037810 (2009).
8. Balmaseda, M. A., Trenberth, K. E. & Källén, E. Distinctive climate signals in reanalysis of global ocean heat content. *Geophysical Research Letters* **40,** 1754–1759, doi: 10.1002/grl.50382 (2013).
9. Meehl, G. A., Arblaster, J. M., Fasullo, J. T., Hu, A. & Trenberth, K. E. Model-based evidence of deep-ocean heat uptake during surface-temperature hiatus periods. *Nature Climate Change* **1,** 360–364, doi: 10.1038/nclimate1229 (2011).
10. England, M. H. *et al.* Recent intensification of wind-driven circulation in the Pacific and the ongoing warming hiatus. *Nature Climate Change* **4,** 222–227, doi: 10.1038/nclimate2106 (2014).
11. Jolliffe, I. *Principal Component Analysis* (Springer-Verlag, 1986), 2 edn.
12. Navarra, A. & Simoncini, V. *A Guide to Empirical Orthogonal Functions for Climate Data Analysis* (Springer-Verlag, 2010).
13. Van Den Dool, H. M., Saha, S. & Johansson, A. Empirical orthogonal teleconnections. *Journal of Climate* **13,** 1421–1435 (2000).
14. Vejmelka, M. *et al.* Non-random correlation structures and dimensionality reduction in multivariate climate data. *Climate Dynamics* **44,** 2663–2682, doi: 10.1007/s00382-014-2244-z (2015).
15. Keppenne, C. L. & Ghil, M. Adaptive filtering and prediction of noisy multivariate signals: An application to subannual variability in atmospheric angular momentum. *Intl. J. Bifurcation Chaos* **3,** 625–634 (1993).
16. Plaut, G. & Vautard, R. Spells of low-frequency oscillations and weather regimes in the Northern Hemisphere. *J. Atmos. Sci.* **51,** 210–236 (1994).
17. Ghil, M. *et al.* Advanced spectral methods for climatic time series. *Rev. Geophys.* **40,** 1003–1, doi: 10.1029/2000rg000092 (2002).
18. Mukhin, D. *et al.* Predicting Critical Transitions in ENSO models. Part II: Spatially Dependent Models. *J. Climate* **28,** 1962–1976, doi: 10.1175/jcli-d-14-00240.1 (2015).
19. Kondrashov, D. & Berloff, P. Stochastic modeling of decadal variability in ocean gyres. *Geophys. Res. Lett.* **42,** 1543–1553, doi: 10.1002/2014gl062871 (2015).
20. Groth, A. & Ghil, M. Multivariate singular spectrum analysis and the road to phase synchronization. *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics* **84,** 1–10 (2011).
21. Kwasniok, F. Reduced Atmospheric Models Using Dynamically Motivated Basis Functions. *Journal of the Atmospheric Sciences* **64,** 3452–3474 (2007).
22. Kwasniok, F. Optimal Galerkin approximations of partial differential equations using principal interaction patterns. *Physical Review E* **55,** 5365–5375 (1997).
23. Hastie, T. *Principal Curves and Surfaces: Ph.D Dissertation.* Ph.D. thesis, Stanford Linear Accelerator Center, Stanford University (1984).
24. Scholkopf, B., Smola, A. & Muller, K. R. Nonlinear Component Analysis as a Kernel Eigenvalue Problem. *Neural Comp.* **10,** 1299–1319 (1998).
25. Gámez, A. J., Zhou, C. S., Timmermann, A. & Kurths, J. Nonlinear dimensionality reduction in climate data. *Nonlinear Processes in Geophysics* **11,** 393–398 (2004).
26. Ross, I., Valdes, P. & Wiggins, S. ENSO dynamics in current climate models: an investigation using nonlinear dimensionality reduction. *Nonlinear Processes in Geophysics* **15,** 339–363, doi: 10.5194/npg-15-339-2008 (2008).
27. Reynolds, R. W., Rayner, N. A., Smith, T. M., Stokes, D. C. & Wang, W. An Improved *In Situ* and Satellite SST Analysis for Climate. *Journal of Climate* **15,** 1609–1625, doi: 10.1175/1520-0442(2002)015%3C1609:aiisas%3E2.0.co;2 (2002).
28. IGOSS nmc: Global SST data from Reynolds and Smith. Date of access: 17/11/2014. Available online at http://iridl.ldeo.columbia.edu/SOURCES/.IGOSS/.nmc/.Reyn_SmithOIv2/.monthly/.sst/.
29. Chao, Y., Ghil, M. & McWilliams, J. C. Pacific interdecadal variability in this century's sea surface temperatures. *Geophysical Research Letters* **27,** 2261–2264, doi: 10.1029/1999GL011324 (2000).

30. Penland, C. & Matrosova, L. Prediction of tropical Atlantic sea surface temperatures using linear inverse modeling. *Journal of Climate* **11,** 483–496 (1998).
31. Saji, N. H., Goswami, B. N., Vinayachandran, P. N. & Yamagata, T. A dipole mode in the tropical Indian Ocean. *Nature* **401,** 360–363, doi: 10.1038/43854 (1999).
32. Trenberth, K. E. & Fasullo, J. T. Tracking Earth's Energy. *Science* **328,** 316–317, doi: 10.1126/science.1187272 (2010).
33. Glantz, M., Katz, R. & Nicholls, N. *Teleconnections Linking Worldwide Climate Anomalies* (Cambridge University Press, 1991).
34. Molkov, Y. I., Mukhin, D. N., Loskutov, E. M., Timushev, R. I. & Feigin, A. M. Prognosis of qualitative system behavior by noisy, nonstationary, chaotic time series. *Phys. Rev. E* **84,** 036215, doi: 10.1103/PhysRevE.84.036215 (2011).
35. Molkov, Y. I., Loskutov, E. M., Mukhin, D. N. & Feigin, A. M. Random dynamical models from time series. *Phys. Rev. E* **85,** 036216, doi: 10.1103/PhysRevE.85.036216 (2012).
36. Molkov, Y. I., Mukhin, D. N., Loskutov, E. M., Feigin, A. M. & Fidelin, G. A. Using the minimum description length principle for global reconstruction of dynamic systems from noisy time series. *Phys. Rev. E* **80,** 046207, doi: 10.1103/PhysRevE.80.046207 (2009).

Acknowledgements

Author Contributions

D.M. conceived the project, developed the method, analyzed the results, and wrote the manuscript, A.G. developed and performed the numerical algorithm, wrote Supplementary Materials and analyzed the data, E.L. analyzed the data and plotted the graphs, A.F. set the line of research and analyzed the results, J.K. analyzed the results and corrected the manuscript. All authors reviewed the manuscript.

Additional Information

Supplementary information accompanies this paper at http://www.nature.com/srep

Competing financial interests: The authors declare no competing financial interests.

How to cite this article: Mukhin, D. *et al.* Principal nonlinear dynamical modes of climate variability. *Sci. Rep.* **5**, 15510; doi: 10.1038/srep15510 (2015).