doi:10.1038/nature11655

1. Erhai lake-catchment system: recent eutrophication

Lake Erhai (25°36'-25°58' N, 100°05'-100°18' E) is the second largest lake in Yunnan Province, China (Fig. S1) with a lake area of ~250 km^2 and catchment area ~2,250 km^2. Erhai provides a water resource for a catchment population of ~820,000 people, 65% of which are dependent on agriculture or fishing[28]. The lake first exhibited visible signs of eutrophication in 1996 with localised blooms of blue-green algae[28] and since then has been officially characterised as eutrophic (Chinese water class III)[28]. Monitored algae concentration data[29] and water quality data[30] are available for the period 1995-2005 and 1992-2009 respectively (the data for 2006-2009 are from the Bulletin of Environment States of Yunnan Province) (Fig. S4). The records show that levels of blue-green algae and diatoms rose abruptly from 2002, most levels remaining high until at least 2005. Levels of dissolved nutrients in the lake increased up to 2003 before declining to reach new minima in 2008. These curves mirror the rise and fall of the chemical oxygen demand, the long term decline in water transparency (Secchi disk data available only to 2005), and decreasing levels of dissolved oxygen.

2. Land use, climate and lake levels

The Dali (Fig. S1) Prefectural values for annual crop production (Fig 1e) 1952-2009 (continuous except for 1953-1956 which are unrecorded) are from the Dali Year Book and Statistical Bulletin of Dali Economic and Social Development (accessed through the Dali Prefecture Statistics Bureau). Over the past 60 years, the largest environmental changes in the catchment are linked to agriculture, driven by rapidly rising rural and urban populations[28]. Combined crop yields for the two harvest system (early spring rice or maize, and late summer wheat or broad bean) in Dali Prefecture have risen ~3 fold

since 1950, with the highest rates of increase during the 1990s (Fig 1e). Since 1990, the area of crop lands in Dali and Eryuan basins has reduced by more than 15% while average fertilizer applications between 1999 and 2005 have increased by 48% to reach an average application rate of nearly 1000 kg ha^{-1}. Livestock farming is growing rapidly driven by the growing demand for dairy products and meat in the last 10 years, triggered partly by a 1998 government directive to increase milk consumption (http://factsanddetails.com/china). For the Erhai catchment, it is estimated that the numbers of cows in hand will have risen 3-fold between 2002 and 2010[28]. The rising human population has also had a major impact on delivery of nutrients to the lake through untreated sewage. In 2007, it was estimated that nutrient loadings to Erhai were predominantly from non-point sources with largest contributions (>20%) from village domestic wastewater (23% TN, 42% COD, 44% NH_4-N) and urban domestic wastewater (30% COD, 32% NH_4-N)[28]. But contributions from recent livestock farming are higher (41% TP, 23% TN, 21% COD, 20% NH_4-N). Despite agricultural intensification, percentage contributions from modern (2007) agricultural runoff and soil erosion are relatively low (17% TP, 8% TN and 6% COD; 13% TP and 28% TN respectively), largely as a result of the recent additional loadings from livestock farming and urban wastewater[28].

Monthly temperature and precipitation data for Dali (Fig. 1f) between January 1951 and November 2010 were provided by the China Meteorological Administration. While the long-term rainfall trend is stationary (Fig. 1f), the inter-annual variability was relatively high in the 1960s and relatively low in the 1990s. In terms of agricultural crisis, severe drought years in Yunnan were recorded in 1977, 2005-06 and 2009-10: the earliest year corresponding to a measurable drop in crop yields. The annual temperature records show a cooling trend until the early 1990s with subsequent warming up to the present (Fig. 1f).

Lake level data (Fig. 1e) were compiled[30,31] for the period 1952-2007. Lake levels were climatically constrained until a sluice was built on the Xierhe river outflow in 1963. Deepening of the riverbed and construction of a hydroelectric plant led to significantly lower but uncontrolled water levels between 1973 and 1989. Since 1990, water levels have been regulated within legislated minimum and maximum limits[28]. Two periods of exceptionally low levels in 1982-83 and 1988-89 appear to be due to the compounding effect of lowered levels and minima in the annual rainfall record.

3. Lake sediment sampling and dating

A 35 cm core (laboratory code EN5) was collected with a Kajak corer from the deepest part (20 m) of Erhai lake (Fig. S1) in 2007 and analysed for fossil chironomid head capsules at 0.5 cm intervals. A 64 cm sediment core (laboratory code EH1) was collected from the same location (Fig. S1) in 2009 with a Kajak gravity corer. Contiguous sub-samples were sliced in the field at 0.5 cm intervals. These were analysed to create time-series of fossil diatoms, inorganic and organic chemistry. EN5 and EH1 were sampled close to an earlier sampled (2001) and longer pilot core (laboratory code EH2/2001)[32] that was analysed for long term environmental changes.

Sediment chronologies for both short cores were produced from measurements of ^{210}Pb, ^{226}Ra and ^{137}Cs radionuclide activities in contiguous samples using gamma spectrometry at the State Key Laboratory of Lake Science and Environment, Chinese Academy of Sciences. A constant rate of supply (CRS) model was employed to calculate the dates of each sample[33]. The highest ^{137}Cs values exist at 17±1.0 cm in EH1, and 15.75±1.5 cm in EN5 (Fig. S2) which are interpreted as recording the peak release of ^{137}Cs in the northern hemisphere from global nuclear tests dated to 1963[33]. The results show that both ^{137}Cs and ^{210}Pb methods are consistent within errors for the 1963-dated sample in both cores. The CRS model results for both cores, with

2 standard errors (2σ) (Fig. S2), show that the size of model errors rises with age from 3-4 years at 2000, 14-17 years at 1950, to 25-27 years at 1900. All dates in the main text are expressed as estimated means. The chronology of long core EH2 was derived from a combination of interpolated ^{14}C determinations, palaeomagnetic dates and ^{210}Pb analyses[32].

4. Lake sediment techniques

Sample preparation for chironomid head capsule (from non-biting midge) analysis followed standard procedures[34]. A solution of 10% KOH was added to the samples overnight before sieving through a 90 μm mesh. Residues were examined under a stereo-zoom microscope at 25x and all head capsules found were mounted on microscope slides in a solution of Hydromatrix.

Diatom preparations followed standard procedures[35]. At least 300 valves were counted from each sample, and diatom communities were expressed as percentage of each species in the total number of valves counted per sample. Only the species with abundance greater than 1% and which appeared in more than 2 samples were retained for the final statistical analysis. Nomenclature and taxonomy (Fig. S3a) mainly follows Krammer and Lange-Bertalot[36-40], and the current equivalence taxonomy refers to Algaebase (http://www.algaebase.org/). Diatom preservation in EH1 was assessed using an index[41] based on visible evidence of broken and partially dissolved valves. The index for all samples exceeded a value of 0.7 indicating a good level of preservation.

The standard US Environment Protection Agency method (3052) was employed[42] for metal elements analysis using inductively coupled plasma-atomic emission spectrometry (ICP-AES). A representative weighed sediment sample (~125 mg) was placed in a Teflon nitrification tank and 6.0 ml HNO_3, 0.5 ml HCl and 3.0 ml HF were added. The sealed tank was then

placed in a microwave oven (Berghof MWS-3 Digester) and nitrified at 180±5 °C for 15 min. The residue was then transferred into a Teflon beaker and dissolved with 0.5 ml $HClO_4$ by braising in a heating block at about 200 °C and diluted to 25 ml with double-distilled de-ionized water. The solution was then analyzed for metal elements by ICP-AES (Leeman Labs, Profile DV). The accuracy of the analytical determination was established using the reference material GSD-9, supplied by the Chinese Academy of Geological Sciences. The analytical results for all elements were found to be in agreement with the certified values, with accuracy better than 93%. Elemental data are expressed in mg g^{-1} total sediment (Fig. S3b).

Around 2 g of wet sediment was sampled for total organic carbon (TC% sediment mass) and total nitrogen (TN% sediment mass) measurement. The samples were placed in 50 ml of 5% hydrochloric acid overnight to remove any carbonates. The samples were then washed with deionised water, and oven dried at 40 °C. The sediments were ground into a fine powder, and sieved at 80 µm. TC and TN were measured by a FlashEA 1112 Elemental Analyser linked to a Thermo DeltaPlus Advantage mass spectrometer and expressed as percentages of C and N.

Detrended correspondence analysis (DCA) was employed for analysing changes in diatoms using square-root transformations and down-weighting of rare taxa[43]. DCA axis 1 scores were used to represent changes in the diatom communities while Hill's diversity index N2 (HDI)[44] was employed to represent changes in the diatom diversity. DCA and HDI calculations were undertaken in CANOCO software (version 4.5)[43].

5. Lake sediment data

The full record of diatoms (Fig. S3a) for EH1 shows 67 diatom species. About 70% of the recent diatom community is dominated by planktonic species, for example *Aulacoseira ambigua* (Grunow) Simonsen, *Cyclostephanos dubius* (Fricke) Round, *Cyclotella ocellata* Pantocsek, and *Fragilaria crotonensis* Kitton. The community shows major changes centred on 7 cm depth (dated ~ 2001). Below this depth, species such as *Aulacesira ambigua*, *Fragilaria crontonensis* are far less abundant. One notable component of the spring bloom in Erhai is *Fragilaria crotonensis* which is related to lake circulation and stratification[45] and often used as a eutrophication indicator. This species increased from ~17 cm depth (~1960), while other typically eutrophic species, for example *Cyclostephanos dubius* and *Aulacesira ambigua*, increased gradually over the length of the core. The fossil diatoms from long core EH2 below a depth of 3-5 cm are dominated by planktonic species, typically ~80%. *Cyclostephanos dubius* is the dominant species typically >70% in most of samples. *Fragilaria* species are distributed throughout EH2 but their abundance is small except in the uppermost samples. The most striking change in diatom assemblage over last 750 years is a shift from centric species to *Fragilaria* species at the top of the core similar to the observed changes in EH1.

The sedimentary changes in metal elements from EH1 (Fig. S3b) also show a clear change at the depth 6.0-7.5 cm (dated 2000-2003) with calcium (Ca) and phosphorus (P) rising while the other elements decline. Below this depth, most of the elements show stationary or gradually changing trends. The trends in Ca and P are similar to total organic carbon (TC %) and total nitrogen (TN %), strongly suggesting that all four elements are linked to increased biological productivity as a result of eutrophication (Fig. 1c and d). Rises across the regime shift are ~3-fold for Ca (Fig. 1c), TN % and TC % (Fig. 1d) but far less pronounced for P with only a ~1.3 fold increase between 2000 and 2009.

Given the evidence for increasing concentrations of dissolved TP up to 2000 (Fig. S4b) and increasing levels of aquatic productivity up to 2000 and beyond, a relatively low increase in sediment P is consistent with recycling of P from the upper sediments as anoxia develops (Fig. 1b and Fig. S5). The reduced percentage values of other elements (Fig. S3b) after 2000 that have mainly detrital origins from the metamorphic and basic igneous geologies[32] in the catchment (Al, Ba, Be, Co, Cr, Cu, Fe, Li, K, Mg, Mn, Na, Ni, Pb, Sr, Ti, V, Zn) is due to the diluting effect of higher absolute fluxes of biological material to the sediment. There is no obvious change in land use or hydrology that could account for an absolute decrease in detrital fluxes in the period 2000-2003.

6. Evidence for bistability and critical transitions

DCA data for EH2 cover the period of continuous and intensive agriculture recorded at Erhai[32,46] over ~750 years from just before the start of the Ming Dynasty (AD 1346). Comparing the normalised DCA variability using z scores shows that the long core EH2 data lie mainly within the range ±1 standard deviations, whereas the abrupt recent changes in recent short core EH1 data exceed -2 standard deviations (Fig. S6). This is evidence that over the past 750 years the aquatic system has behaved as a relatively stable, though fluctuating system, until very recent times. Probability density functions (Gaussian kernel density estimation) for DCA in EH1 indicate the development of bimodality over the period of early warning signals 1971-2000 (Fig 2a). Probability density functions using the full range of standard selectors for bandwidth (using the Stats package downloaded from http://www.r-project.org/) show similar results: Fig. 2a shows results using pilot estimates of derivatives[47].

We tested the hypothesis for a critical transition at ~2001 in two ways. First, we searched for critical transitions in the DCA and HDI time-series 1980-2009 through sequential Student's t-test[48]. Significant break-points were identified

from a sequential analysis of mean values using Student's t-test set to detect regime shifts lasting at least five years (i.e cut-off length =5). The results for both DCA (Fig. 2b) and HDI with the Huber parameter[48] set to 1 showed statistically significant ($p \leq 0.01$) break-points indicating a transition in the time series at ~2001.

Second, we fitted a suite of autoregressive integrated moving average [ARIMA (p,d,q)] models where p is the autoregressive (AR) order, q is the moving average (MA) order and d is the differencing part to the DCA and HDI time series for the period 1883-2000 (i.e. up until the point of the hypothesized critical transition) using the Stats package downloaded from http://www.r-project.org/. We fitted ARIMA models ranging in orders from (1,1,0) to (4,2,2) for the DCA time series and from (1,0,0) to (4,1,2) for the HDI time series. We selected the optimum model for each time series based on the lowest Akaike Information Criterion (AIC) to predict the evolution of DCA and HDI for the period of 2001-2010. Changing the d value to account for trends did not improve the models for DCA and made only small differences to models for HDI. Lowest AIC values were found for models (2,1,1) for DCA and (2,0,1) for HDI (Table S1). For both time-series, the predicted values for 2001-2010 diverge from the observed data within 95% probability levels (DCA shown in Fig 2c) showing that the abrupt changes observed in ~2001 cannot be predicted from previous observations using a linear model. The results of the sequential statistical tests and ARIMA models allow rejection of the null hypothesis that there are no significant break-points in the DCA and HDI data.

We have also tested the null hypothesis that nonstationary drivers cause the observed changes in DCA and HDI curves, using regression models. Linear multiple regression (Table S2a) for the period 1994-2007 (with the proposed critical transition at the mid-point 2000-2001) shows only a weak explanation of DCA or HDI in terms of crop yield, lake water level, precipitation and

temperature (statistically insignificant levels for DCA p≤0.062 and HDI p≤0.078). In simple regression models (Table S2b), only lake water level shows statistically significant explanatory power but with a *negative* and counterintuitive relationship that cannot easily be explained in simple cause-and-effect terms. On this basis we reject this null hypothesis in favour of a critical transition brought about by complex nonlinear interactions, a new positive feedback loop and transgression of an internal threshold.

7. Phase diagrams and hysteresis

Monitored values for TP and TN were plotted against the DCA and HDI curves (representing changes in diatom communities) for the period 1992-2008 (the most recent publicly available data). The DCA versus TP plot shows two clusters of points, each with approximately linear trends (Fig. 2d), indicating that more than one algal state is possible for a specific level of TP. Two clusters of points are also evident for DCA versus TN, HDI versus TP and TN, and Bacillariophyta versus TP and TN, but in each case the evidence for hysteresis is weaker than for DCA versus TP. We note that while phase space plots can provide evidence for alternate steady states, evidence for hysteresis can only exist where the driving variable is in decline following a regime shift. In many modern environments where stresses are continuously increasing, it may prove difficult to demonstrate that a fold bifurcation has taken place. Phase space patterns between lowest and highest water level, annual temperature and annual rainfall against annual DCA and HDI show weak evidence for alternate steady states and hysteresis. Phase space patterns between annual DCA/HDI and lowest lake levels in the period 1952-2007 show some evidence for alternate steady states. For DCA, the two clusters appear approximately linear, but less so for HDI. Phase space plots for high water level and climate variables against DCA and HDI show no evidence for alternate steady states.

8. Early warning signal metrics

Calculations of variance, autocorrelation and skewness (Fig. 3) were undertaken on linearly interpolated (1-year steps) diatom DCA and HDI data in order to obtain equivalent time series[18]. The time series prior to the regime shift (before 2001) were detrended with a single exponential smoothing method (Minitab software) to filter out long-term trends. The default weight was chosen for smoothing, which Minitab computes by fitting an ARIMA (0, 1, 1) model to the data. The residuals between interpolated data and fitted data were used to calculate standard deviations, lag-1 autocorrelation and skewness under half-time series moving window size (59 yr). The variance is represented by the standard deviation data.

We tested the sensitivity of the early warning metrics to different data and statistical treatments: a) using non-interpolated DCA and HDI data, b) using correspondence analysis (CA) as an alternative to DCA, c) calculating early warning metrics using different window sizes, d) using alternative smoothing functions, e) assessing the effect of chronological errors.

The pattern of residuals (Fig. S7a) and rising variance (Fig. S7b) in the non-interpolated diatom DCA are similar to the results for the interpolated DCA data (Fig. 3b-c), but the rising variance for HDI is weaker (Fig. S7b). Correspondence analysis (CA) was conducted on the diatom data using CANOCO software (version 4.5) with species square-root transformations and down-weighting of rare taxa. Post-1980s rising variance and skewness of the CA (Fig. S7c) is clear, but autocorrelation shows a declining trend before the transition (Fig. S7d). The importance of sliding window size was assessed with additional calculations using 30 yr and 39 yr (1/4 and 1/3 time-series) windows. The different moving window sizes show similarly rising variance (Fig. S7e) as observed for the 59 yr window (Fig. 3c). Gaussian kernel smoothing was conducted with a curve fit creator in Excel and the results (Fig. S7g, h) show

the same trends for variance and autocorrelation as for the single exponential smoothing method. However, the autocorrelation values (with 5 as a smoothing bandwidth) are much higher than autocorrelation in the single exponential smoothing method (with ARIMA (0, 1, 1) weight). Two additional sets of calculations were undertaken for DCA data based on mean lake sediment dates ±2 standard errors for each sample. Both sets of residuals show increasing variability towards the present (Fig. S7i) and both trends in variance are rising (Fig. S7j). But compared to the rising variance based on mean dates the start of the rise is put back by ~10 yr for - 2 standard errors and brought forwards ~10 yr for +2 standard errors. Overall the rising trend in variance is robust, but the start of rising variance is subject to chronological accuracy.

9. Phosphorus model

We modelled eutrophication using a simple nondimensional model that describes nutrient dynamics in a lake[22]. This model (and variations of it) has been extensively used to study mechanistically the potential for abrupt transitions[49-52] and early warning indicators[17, 53, 54] to eutrophication due to phosphorus recycling in a lake. In the model, the lake can switch between two distinct states: an oligotrophic state of low phosphorus concentration, and a eutrophic state where phosphorus concentrations are high. Phosphorus inputs enter the lake from the surrounding watershed. Part of the phosphorus can be lost due to runoff and/or losses to sediment or hypolimnion. At the same time there is recycling of phosphorus back to the water column from the sediment and/or hypolimnion. Recycling is enhanced at lower oxygen levels that are favoured when input loads from the watershed exceed a critical threshold. Above this critical threshold, recycling switches from a low to a high rate that leads to high phosphorus concentration in the lake and thus to eutrophication. In contrast, if phosphorus input is low the lake is trapped in an oligotrophic state due to phosphorus losses to sediment and/or hypolimnion and low

recycling. The model reads:

$$dP = [a - sP + r\frac{P^n}{P^n + 1^n}]dt + \sigma P dW$$

where P is phosphorus concentration, α is the phosphorus input rate (control parameter), r represents maximum recycling rate (=1), s is the phosphorus loss rate (that incorporates losses due to run-off, sedimentation and/or loss to hypolimnion (=1)), n is the exponent that describes the recycling relationship to phosphorus concentration (=8), and white noise is added through a Wiener process dW with intensity σ (=0.25) and scaled to phosphorus concentration P. We did not parameterise the model on the empirical data of lake Erhai, but used values within ranges from earlier contributions[17, 49-54] as our objective was only to explore the behaviour of early warnings in a flickering regime. In addition, our results are independent from the actual parameter setting used, as flickering is mainly driven by the magnitude and nature of stochasticity[17].

We started simulations (Fig. 4) from an oligotrophic regime (low phosphorus concentration at low input rate) and increased input rate α linearly with time in 2000 time steps from 0.1 to 0.7. The transition to eutrophication occurred at α=0.6619 at time step 1848. Due to the strong noise regime we implemented (σ =0.25), dynamics over a range of input rates were characterized by flickering between the alternative basins of attraction. We estimated variance (as standard deviation) and lag-1 autocorrelation in moving windows of half the size of the simulated time series up to the transition after detrending. All simulations and statistical analyses on simulated data were performed in MATLAB (v. 7.0.1). The model was solved using an Euler integration scheme with Ito calculus.

References

28. World Bank. China - *Second Yunnan Urban Environmental Project (Vol. 5 of 5) : Erhai Lake Basin regional environmental assessment*. Report No. E1748, 138 (The World Bank, 2007).

29. Yan, C. Z. *et al*. Ecological protection and sustainable utilization of Erhai Lake, Yunnan. *Environmental Science,* **26**, 38-43 (2005) (in Chinese).

30. Bai, J. K. *Dali Erhai Lake Scientific Research*. 675 (The Ethnic Publishing House, 2003) (in Chinese).

31. Li, P. The analysis of Erhai Lake's feature water level. *Pearl River*, **46-48**, 85 (2008) (in Chinese).

32. Dearing, J. A. *et al.* Using multiple archives to understand past and present climate-human-environment interactions: the lake Erhai catchment, Yunnan Province, China. *J Paleolimnol.* **40,** 3-31 (2008).

33. Appleby, P. G. in *Tracking Environmental Change Using Lake Sediments Vol. 1 Developments in Paleoenvironmental* Research eds Last, W.M. & Smol, J.P. (Springer, Netherlands, 2002).

34. Brooks, S. J., Langdon, P. G., & Heiri, O. *The identification and use of Palaearctic Chironomidae larvae in palaeoecology*. (Quaternary Research Association, 2007).

35. Battarbee ,R.W., Charles, D.F., Bigler, C., Cumming, B.F. & Renberg, I. in *The Diatoms: Applications for the Environmental and Earth Sciences*. 2nd edition. eds Smol, J.P. & Stoermer, E.F. (Cambridge University Press, Cambridge. 2010).

36. Krammer, K. & Lange-Bertalot, H. *Bacillariophyceae. 2 Teil, Bacillariaceae, Epithemiaceae, Surirellaceae.* (Gustav Fischer Verlag Press, Stuttgart, 1988).

37. Krammer, K. & Lange-Bertalot, H. *Bacillariophyceae. 2/1 Teil, Naviculaceae*. (Gustav Fischer Verlag Press, Stuttgart, 1988).

38. Krammer, K. & Lange-Bertalot, H. *Bacillariophyceae. 4 Teil, Achnanthaceae, Kritische Ergänzungen zu Navicula (Lineolatae) und Gomphonema Gesamtliteraturverzeichnis* Teil 1-4. (Gustav Fischer Verlag Press, Stuttgart, 1991).

39. Krammer, K. & Lange-Bertalot, H. *Bacillariophyceae. 3 Teil, Centrales, Fragilariaceae, Eunotiaceae*. (Gustav Fischer Verlag Press, Stuttgart, 1991).

40. Krammer, K. & Lange-Bertalot, H. *Bacillariophyceae*. Part 5, English and French translation of the keys. (Spectrum, 2000).

41. Ryves, D.B., Juggins, S., Fritz, S.C., Battarbee, R.W. Experimental diatom dissolution and the quantification of microfossil preservation in sediments. *Palaeogeogr. Palaeoclimat. Palaeoecol*. **172**, 99-113 (2001).

42. US Environment Protection Agency. *Method 3052: microwave assisted acid digestion of siliceous and organically based matrices*. Report No. SW-846 Ch 3.2, 3052-3051-3052-3020 (EPA Office Solid Waste, Washington, D.C., 1996).

43. ter Braak, C. J. F. & Smilauer, P. CANOCO Reference Manual and CanoDraw for Windows User's Guide: Software for Canonical Community Ordination (version 4.5). (Microcomputer Power (Ithaca), 2002).

44. Hill, M. O. Diversity and evenness: a unifying notation and its consequences. *Ecology* **54**, 427-432 (1973).

45. Bradbury, J.P. A climatic-limnologic model of diatom succession for paleolimnological interpretation of varved sediments at Elk Lake, Minnesota. *J Paleolimnol.* **1**, 115-131(1988).

46. Dearing, J.A. Landscape change and resilience theory: a palaeoenvironmental assessment from Yunnan, SW China, *The Holocene* **18**, 117-127 (2008).

47. Sheather, S. J. and Jones, M. C. A reliable data-based bandwidth selection method for kernel density estimation. *J. Roy. Stat. Soc. B*, **53**, 683-690 (1991).

48. Rodionov, S.N. A sequential algorithm for testing climate regime shifts. *Geophys. Res. Lett.* **31**, L09204 (2004).

49. Carpenter, S R. *Regime Shifts in Lake Ecosystems: Pattern and Variation*. *Excellence in Ecology Series* (Vol. 15, p. 0-in). (Ecology Institute, Oldendorf/Luhe, Germany. 2002).

50. Carpenter, S R. (2005). Eutrophication of aquatic ecosystems: bistability and soil phosphorus. *Proc. Natl. Acad. Sci. USA* **102**, 10002-10005 (2005).

51. Van Nes, E H, & Scheffer, M. Implications of spatial heterogeneity for regime shifts in ecosystems. *Ecology* **86,** 1797–1807 (2005).

52. Carpenter, S. R. & Lathrop, R. C. Probabilistic estimate of a threshold for eutrophication. *Ecosystems* **11**, 601-613 (2008).

53. Dakos, V. *et al*. Spatial correlation as a leading indicator of catastrophic shifts. *Theoret. Ecol*. **3**, 163-174 (2010).

54. van Nes, E.H. &Scheffer, M. Slow recovery from perturbations as a generic indicator of a nearby catastrophic shift. *The American Naturalist* **169**, 738-747 (2007).

Supplementary Tables

Table S1 Details of ARIMA (2,1,1) and (2,0,1) models for DCA and HDI respectively showing model types (AR and MA), coefficients, standard errors of coefficients, Student's t-test statistics and probability levels.

DCA_ARIMA(2, 1, 1)

Type	Coefficient	SE Coef	T	P
AR (1)	1.0706	0.1142	9.37	0.000
AR (2)	-0.67	0.0766	-8.75	0.000
MA (1)	0.3909	0.1483	2.63	0.010

HDI_ARIMA(2, 0, 1)

Type	Coefficient	SE Coef	T	P
AR (1)	1.0747	0.1246	8.62	0.000
AR (2)	-0.3014	0.122	-2.47	0.015
MA (1)	-0.5642	0.1089	-5.18	0.000

Table S2 Simple and multiple regression analysis of potential external drivers of the lake ecosystem (crop yields, lake water level, annual precipitation, annual temperature) and algal responses (DCA and HDI) for the period 1994-2007 (using IBM SPSS 19).

a) Multiple regression model

Standardized coefficients

	DCA	HDI
Crop Yield	-0.305	-0.036
Water level	-0.595*	-0.573
Precipitation	0.009	0.139
Temperature	-0.134	-0.222
Model R square	0.596	0.573

All insignificant at p≤0.05, except * p=0.045 (2-tailed test)

b) Simple regression models

R values

	DCA	HDI
Crop Yield	-0.446	-0.147
Water level	-0.709 *	-0.701*
Precipitation	0.298	0.506
Temperature	-0.274	-0.433

All insignificant at p≤0.05, except * p=0.005 (2-tailed test)

Supplementary Figures

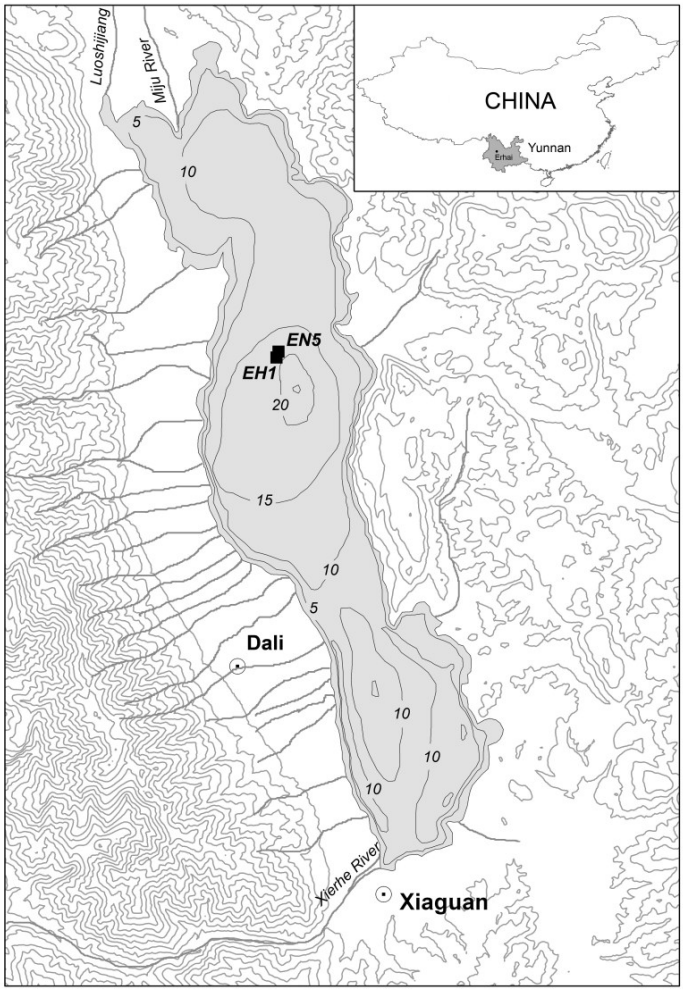

Figure S1 Erhai lake-catchment system showing (inset) location in Yunnan Province, China, underwater and topographic contours (5 m intervals), lake core sampling sites (EH1/EH2 and EN5), locations of Dali and Xiaguan cities, largest inflowing river (Miju River) and outflow (Xierhe River).

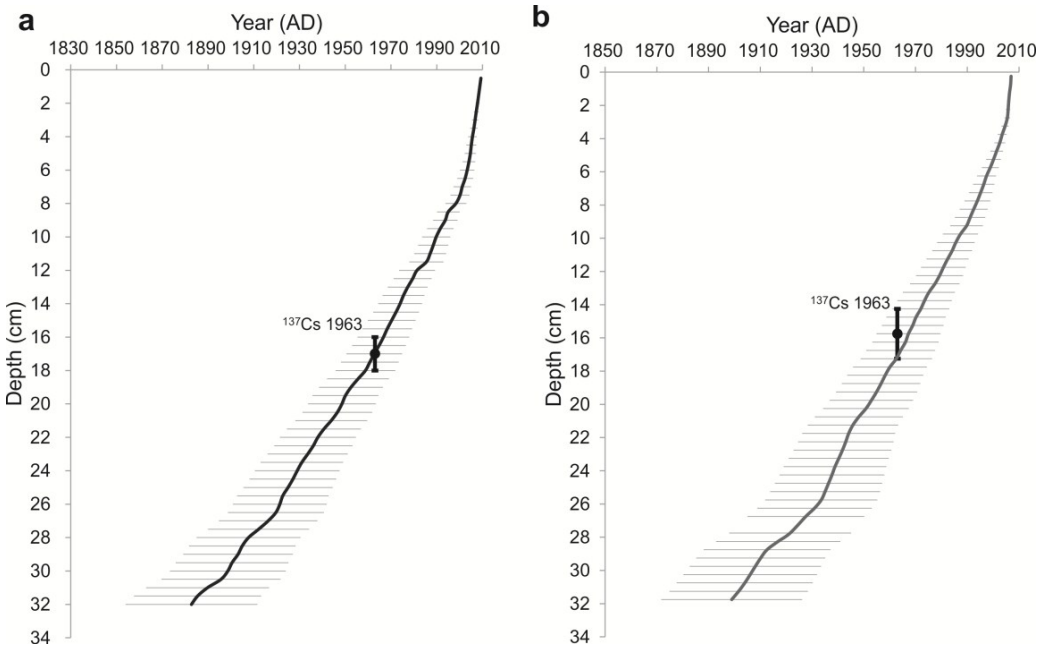

Figure S2 Age-depth curves (solid wavy lines) based on ^{210}Pb analysis for a) EH1 and b) EN5 (Table S1) showing model dating errors (2 standard errors) as horizontal lines. The sediment depth ranges of peak values in ^{137}Cs (are shown (short vertical lines) with mean sediment depth range plotted at 1963.

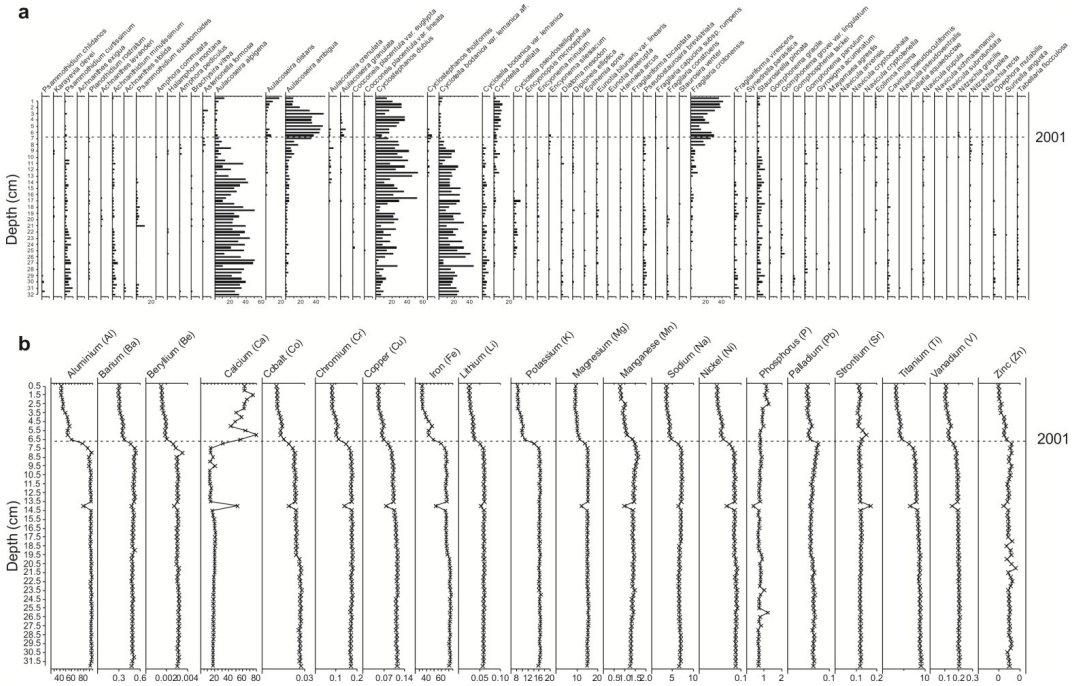

Figure S3 Diatoms and geochemical analyses for EH1 core. a, Diatom communities (percentages of total counts); b, Total elements (mg g^{-1}).

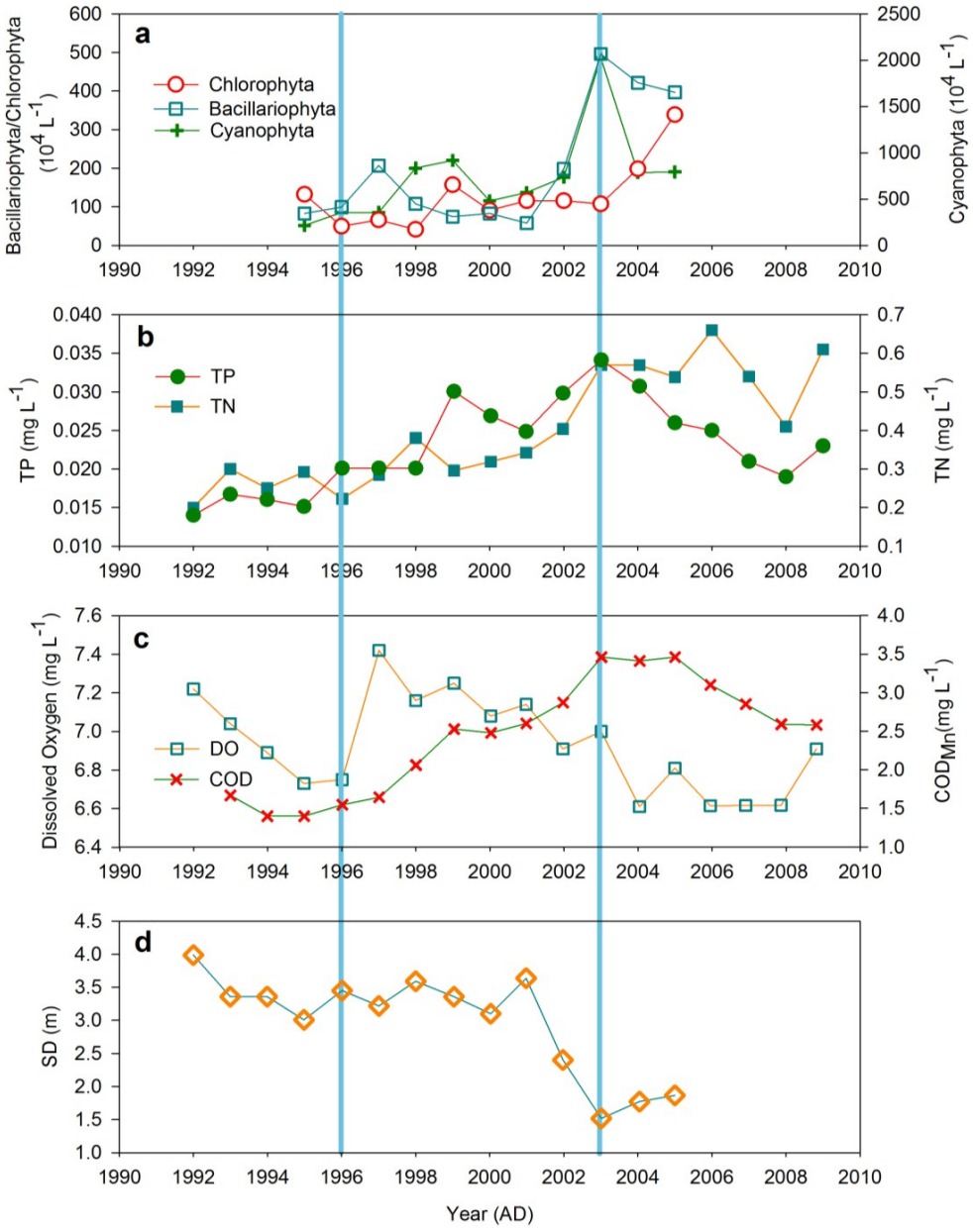

Figure S4 Monitored water quality data 1990s to 2005-2009 (annual means of monthly samples) showing the timing (vertical lines) of the first algal bloom (1996) and the shift to a permanently eutrophic state (2003). a, Algal concentrations (counts L^{-1}) for diatom algae (Bacillariophyta) and blue-green algae (Cyanophyta) show increases of ~3 fold and ~10 fold respectively in the two years before 2003. Green algae (Chlorophyta) show concentrations rising ~3 fold between 2003 and 2005. b, Concentrations (mg L^{-1}) of dissolved total phosphorus (TP) and dissolved total nitrogen (TN) show a rising trend from the

1990s peaking in 2003 (TP) and 2006 (TN) followed by declines to 2008. c, Concentrations (mg L^{-1}) of chemical oxygen demand (taken here to represent dissolved organic compounds) track a similar trend to TP and TN while concentrations of dissolved oxygen (mg L^{-1}) show only a weak link to the algal bloom with a declining trend through the 2003 shift until 2008. d, The secchi disk (SD) depth indicator for water transparency remains fairly constant from the 1990s to 2001, but declines rapidly 2002-2003 from 3.5 m at 2001 to about 1.5 m at 2003 (ref. 29).

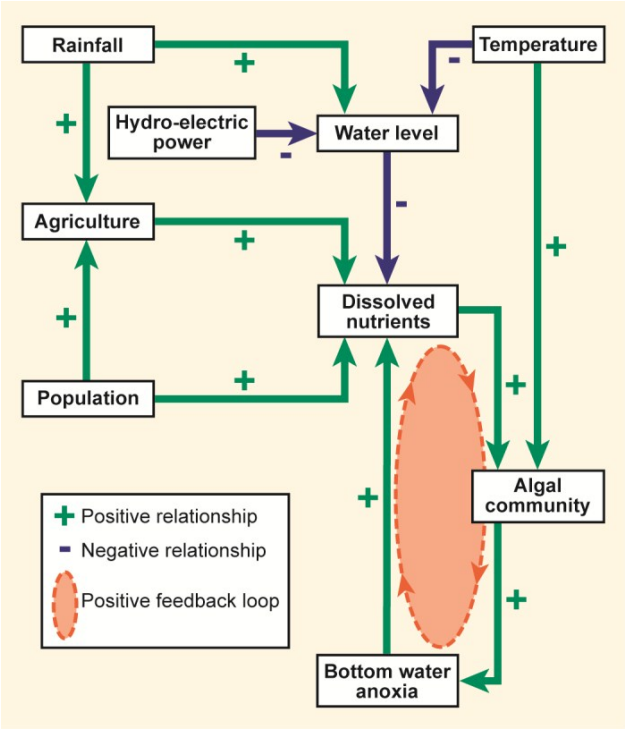

Figure S5 A simplified systems dynamic framework for the Erhai eutrophication process. The figure represents the stage of eutrophication (cf. Fig. S4a) in 2002-2004 when the positive feedback loop (dashed oval) involving recycling of phosphorus from sediments to lake was firmly established. The figure shows the diversity and interconnectedness of major external drivers (bold) of the lake algal community with typical relationships (+ and -) shown for major interactions.

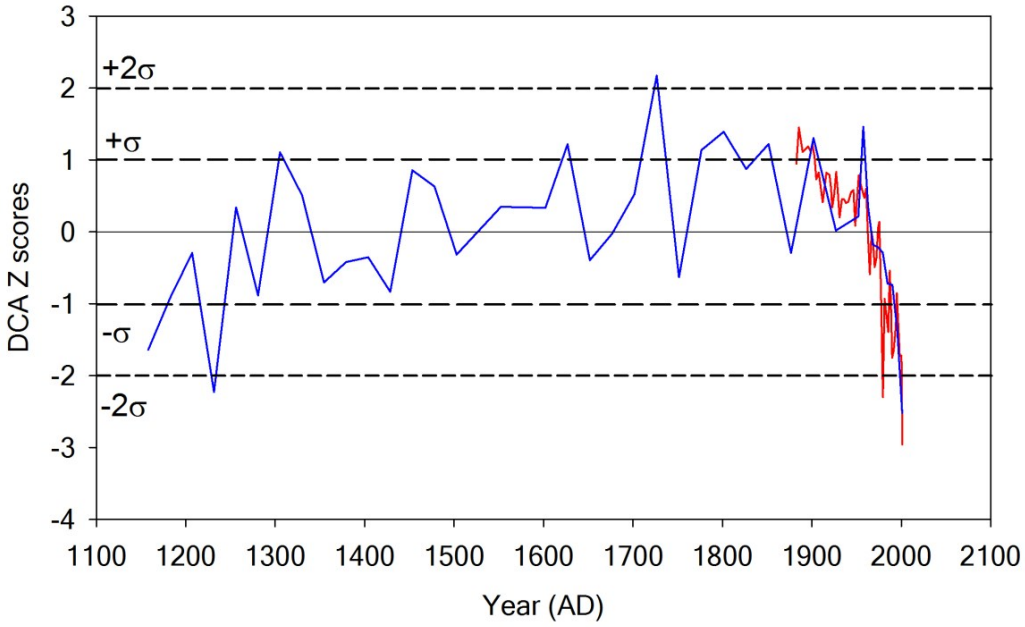

Figure S6 Variability (z scores) of DCA values in long core EH2 1250-2001 and short core EH1 1883-2001. Relatively low, long term variability within ±1 standard deviations up until late 20th century suggests that the aquatic system has only recently shifted to an alternate state.

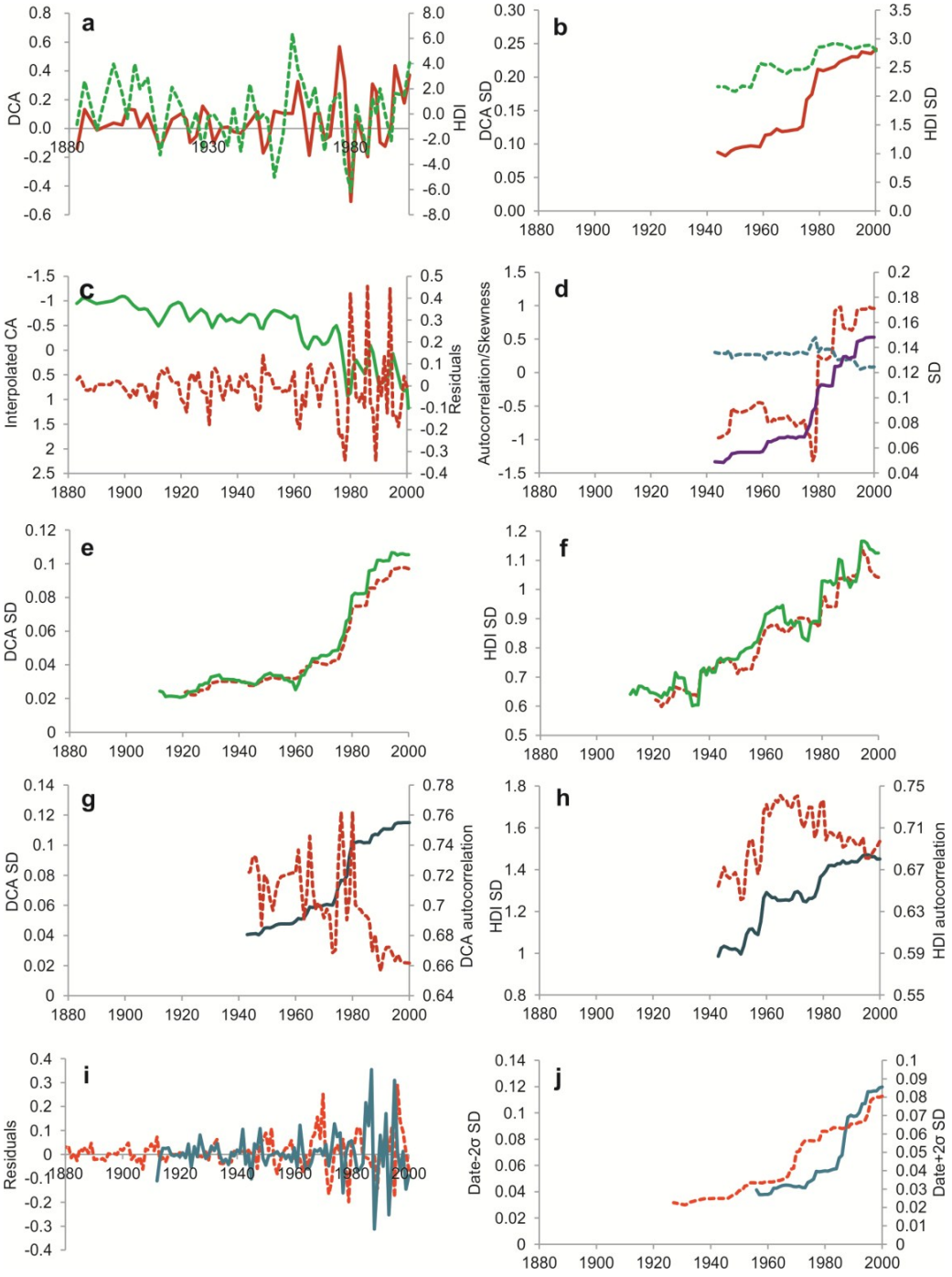

Figure S7 Various early warning signal metrics for diatom DCA and HDI data. a, DCA (solid red line) and HDI (dashed green line) residuals based on real data without interpolation. b, DCA (solid red line) and HDI (dashed green line) variance (SD) based on real data without interpolation. c, Correspondence analysis (CA axis 1) of changes in diatom communities, showing interpolated CA (solid green line) and residuals (red dashed line). Note that the CA axis is

reversed to aid comparison with the community directional changes shown by DCA analysis (see Fig 1). d, CA variance (SD solid purple line), autocorrelation (dashed blue line) and skewness (dotted red line) for interpolated CA data. e, DCA variance (SD) using different moving windows size (30 yr-solid line, 39 yr-dash line). f, HDI variance (SD) using different moving windows size (30 yr-solid line, 39 yr-dash line). g, DCA variance (SD solid line) and autocorrelation (dashed line) using Gaussian kernel smoothing method (with 5 as bandwidth). h, HDI variance (SD solid line) and autocorrelation (dashed line) using Gaussian kernel smoothing method (with 5 as bandwidth). i, DCA residuals calculated with +2 standard errors on dates (solid blue line) and -2 standard errors on dates (dashed red line). j, DCA variances (SD) with +2 standard errors on dates (solid blue line) and -2 standard errors on dates (dashed red line).