

局部非线性直接预测方法在检测海洋 混沌和噪音中的应用*

魏恩泊 宋 谦[†] 覃正才[†] 田纪伟[†]

(上海理工大学动力工程学院 上海 200093)

[†](青岛海洋大学物理海洋实验室 青岛 266003)

摘要 利用局部非线性直接预测方法及实测值与预测值的相关系数,提出了混沌时间序列及含有噪音(白噪音)混沌时间序列的检测混沌和噪音特征的一种方法。该方法对 Logistic 映射资料及赤道日 SST 资料进行了混沌及噪音检测。结果表明,该检测方法优于一般谱方法对混沌资料的检测,在资料含噪信比约为 32% 时,实测值与预测值的相关系数达 0.5,在此可信相关系数范围内可检测资料的混沌特征。同时指出了随着资料噪信比的增加,实测值与预测值的相关系数递减,逐渐表现出噪音特性。此外该方法可用于奇怪吸引子关联维数及 Kolmogorov 熵初步估计。

关键词 非线性预测 混沌 噪音 相关系数

学科分类号 P733

近十几年来随着非线性预测理论及方法的迅速发展,非线性预测理论在检测实测资料的混沌及噪音特征方面得到了广泛研究及应用。Sugihara 等(1990)提出了利用单纯形预测方法区分混沌和测量误差,并应用于疾病的混沌特征研究;Wales(1991)从混沌资料出发证明了非线性预测值和实测值的相关系数随预测时间步长呈指数变化关系,随后 Tsonis 等(1992)在此基础上证明了分形噪音的预测值和实际值的相关系数随预测时间步长呈幂级数关系。尽管以上结论在实际应用中有一定的限制,但从根本上指出了混沌与噪音的区别,使人们对混沌及噪音特征有了初步的认识,同时促使人们寻求适合实测资料(含噪音)的混沌特征分析方法。目前含有噪音资料的混沌特征分析已成为人们研究的重要课题,Jimenez 等(1992)提出了局部优化线性重构预测方法检测混沌,并应用于暴雨资料分析;Barahona 等(1996)利用 Volterra-Wiener-Korenberg 级数预测模式及信息节俭原则检测带噪音资料非线性特性取得了良好的效果。

本文提出了局部非线性直接预测方法对海洋实测资料进行混沌及噪音检测,应用该检测方法对赤道日 SST 资料进行了检测。同时利用该方法对赤道日 SST 奇怪吸引子的关联维数及 Kolmogorov 熵进行了初步估计。

* 国家自然科学基金资助项目, 49476254号。魏恩泊,男,出生于1965年1月,博士,E-mail: lab@lib. ouqd. edu.cn

收稿日期: 1997-09-17, 收修改稿日期: 1998-12-11

1 局部非线性直接预测方法

设 $x(i), i = 1, 2, \dots, n$, 为某物理参量的时间序列, 构造一个 m 维状态空间 $\{X(i)\}$, 其中, $X(i) = (x(i), x(i - \tau), \dots, x(i - (m - 1)\tau))$; 延迟时间 τ 为 Δt 的正整数倍, Δt 为采样时间间隔。令 $X(i), X(i + j)$ 分别为系统目前和未来 j 时刻状态, 则在这两个状态之间建立非线性映射关系 (Farmer *et al.*, 1987)

$$X(i + j) = f_j(X(i)) \quad (1)$$

其中, f_j 是非线性预测函数, j 为预测步长。由于奇怪吸引子在相空间中具有局部相似轨道的演变特性。因此, 寻找与 $X(i)$ 邻近的相似状态 $X(nk)$ 及其相应的未来 j 时刻状态 $X(nk + j)$ (其中, $nk + j < i$, 即 $X(i)$ 以前的信息。), 并利用 $X(nk)$ 和 $X(nk + j)$ 估计预测函数 f_j 。然后代入 $X(i)$ 点可求得预测值 $x(i + j)$, 而每向前预测 j 步重新计算一次 f_j , 这样可得预测值 $x(i + j), x(i + 2j), \dots$ 。

本文非线性预测函数 f_j 选取了二次多项式形式, 即: $f_j = a_0 + \sum_{k=1}^m a_k x(i - (k - 1)\tau) +$

$\sum_{k=1}^m \sum_{s=k}^m a_{k,s} x(i - (k - 1)\tau)x(i - (s - 1)\tau)$, 其中, 未知系数个数为 $P = (m + 2)! / m! 2!$, P 是

寻找与 $X(i)$ 局部相似状态的最小相点数。利用最小二乘法及已知的相点 $X(nk)$ 和 $X(nk + j)$, 求得 f_j 的未知系数。为了保证解的稳定性, 一般情况下邻近点数大约为上 $2p$ 量级。

2 资料

选取 Logistic 迭代模型 $x_{n+1} = 4x_n(1 - x_n)$ (初值 $x_0 = 0.41$) 处于混沌态的 1 024 个点、赤道 ($0^\circ \pm 0.5^\circ S, 165.5^\circ \pm 0.5^\circ E$) 1990—1992 年 El Niño 发生时的日 SST 1 024 点、计算机模拟的取值在 $(-1, +1)$ 上随机噪音 R_n 为 1 024 个点的资料 (即利用计算机产生的随机噪音模拟测量产生的误差)、以及 $0.25(\sin 0.5n + R_n)$ 资料。利用这些资料及它们带有加性噪音 R_n 的资料进行混沌及噪音检测。

令 $x_n, T_n, R_n, S_n (n = 1, 2, 3, \dots, 1024)$ 分别表示 logistic、赤道日 SST、随机噪音时间序列和 $0.25(\sin 0.5n + R_n)$ 资料; $\delta x, \delta T, \delta R$ 分别表示 x_n, T_n, R_n 的均方差; $q\delta R / \delta x, q\delta R / \delta T$ 分别为 x_n, T_n 噪信比 μ , 其中, q 为可调常数。

3 混沌检测应用

3.1 原始资料的混沌检测

由于某物理参量处于混沌态具有有限 Kolmogorov 熵及正 Lyapunov 特征指数, 故可导致系统预测轨迹在有限时间内偏离真实轨道。然而随机噪音本身的混乱特性无规律可寻, 其 Kolmogorov 熵为无穷大, 即在相空间中噪音不存在非线性预测函数, 不具有可预测性, 不随预测时间步长发生变化。此外随机噪音在相空间中具有无穷维, 有限维系统不可能描述噪音, 而混沌系统存在着奇怪吸引子, 奇怪吸引子本身具有局部相似结构, 并且可用有限维相空间刻划。根据上述特性, 利用局部非线性直接预测方法可区分混沌和噪音。

在混沌动力系统中, 某种物理参量的不同模态被激发, 同时隐含着系统物理参量彼此之间的非线性相互作用。然而, 一般噪音 (白噪音) 虽然含有不同随机模态, 但并非存在模态之间相互作用, 这样在谱空间上混沌信号与噪音几乎具有相同几何特征, 一般情况下谱方法较难区分出混沌信号和噪音。为了区分混沌信号和噪音, 必须探讨各模态彼此之间

相关性。利用 FFT 方法对 x_n, T_n, R_n, S_n 作谱分析 (见图 1), 从中很难区分 x_n 及 R_n 的各种模态。图 1a 的噪音谱几乎和图 1b 的 Logistic 谱相同, 混沌和噪音不可区分。这是因为一般谱方法没有真实反映混沌系统的奇怪吸引子局部相似特性及 Kolmogorov 熵和正 Lyapunov 指数特征, 从而没能充分把握混沌特征的实质。由于局部非线性直接预测方法恰好反映这一特性, 可借助预测值 $x(i)$ 及实际值 $y(i)$ 的相关系数 ρ 随预测时间步长衰减变化检测混沌 (Wales, 1991)。

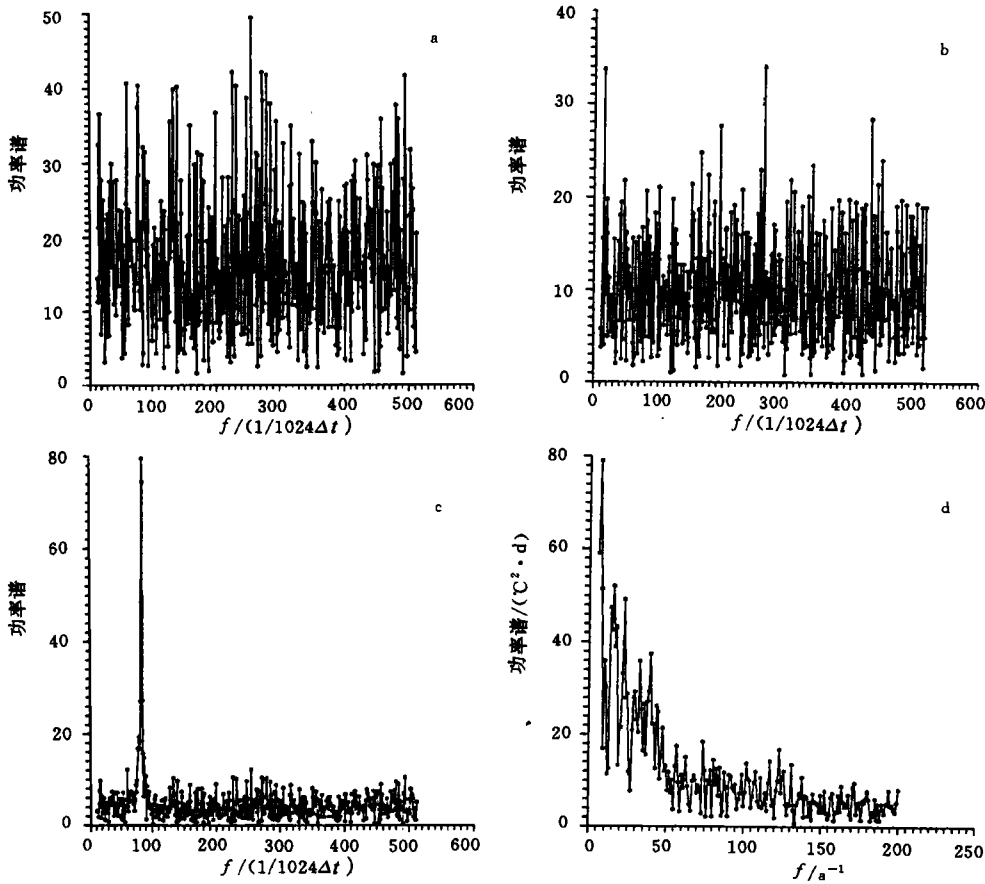


图1 原始时间序列资料的谱分析

Fig.1 Spectral analysis of the original time series

a. R_n ; b. x_n ; c. S_n ; d. T_n

$$\rho = \frac{\sum_i (x(i) - \bar{x})(y(i) - \bar{y})}{\sqrt{\sum_i (x(i) - \bar{x})^2} \sqrt{\sum_i (y(i) - \bar{y})^2}} \quad (2)$$

其中, ρ 为第 j 预测步长的相关系数; \bar{x}, \bar{y} 分别为预测值及实际值的均值。当 $y(i)$ 和每预测 j 步的预测值序列 $x(i)$ ($i = 1, 2, 3, \dots, n$) 的一阶及二阶矩相等, 且 $\Delta_i = x(i) - y(i)$ 服从正态分布, 则

$$1 - \rho(j) = 0.5Ae^{2Kj} \quad (3)$$

其中, A 为正常数, K 为 Kolmogorov 熵, 表示来源某物理参量时间序列信息的平均丢失率。在此条件下相关系数随 j 的增加呈现指数衰减。

由于噪音的不可预测性, 其实际值与预测值的相关系数很小 (对纯噪音其相关系数几乎为 0), 且不随嵌入维数的增加而迅速衰减。混沌系统存在有限维嵌入维数, 在某最小优化嵌入维数 m^* 上, 其相关系数大于其它嵌入维数的相关系数, 从而确定 m^* 。分别利用 x_n 、 T_n 、 R_n 的 1 000 个点的前 500 个做基本预测公式 (1), 随着 j 的增加而相应增加预测模式的所用点数。取 $j = 1$, 预测后 500 个点, 求相关系数 ρ 。 ρ 随嵌入维数的变化见图 2。虽然图 2 的曲线 a 出现极大值, 但是其最大 ρ 仅为 0.1, 正是由于噪音资料不存在有限嵌入维数及局部相似特性, 因而 R_n 的预测值与实际值的 ρ 很小 (ρ 仅在区间 $[-0.1, 0.1]$ 上波动), 可认为不相关, 即该最大 ρ 相对应的嵌入维数无意义。由图 2 中的曲线 b, c 可知 Logistic 资料嵌入维数为 2—4 时 ρ ($\rho = 0.99$) 最大。因而, 取 $m^* = 2$, 由 $m^* = [d_2] + 1$ (其中, d_2 为奇怪吸引子的关联维数) 可知 d_2 约为 1。赤道日 SST 在嵌入维数为 6 时相关系数最大 ($\rho = 0.82$), 可得 $m^* = 6$, d_2 约为 5。

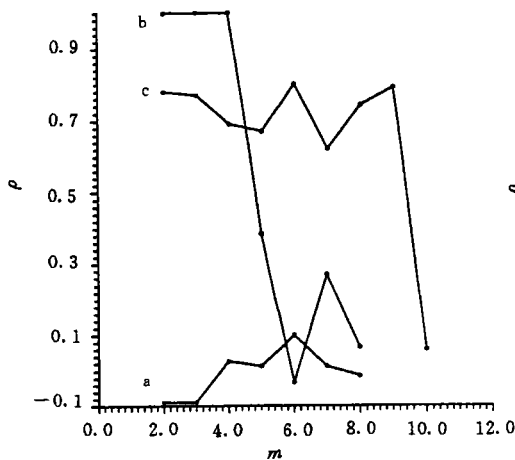


图 2 相关系数与嵌入维数 (m) 的变化关系

Fig.2 Relationships between the correlation coefficients and the embedding dimensions

a. R_n ; b. x_n ; c. T_n

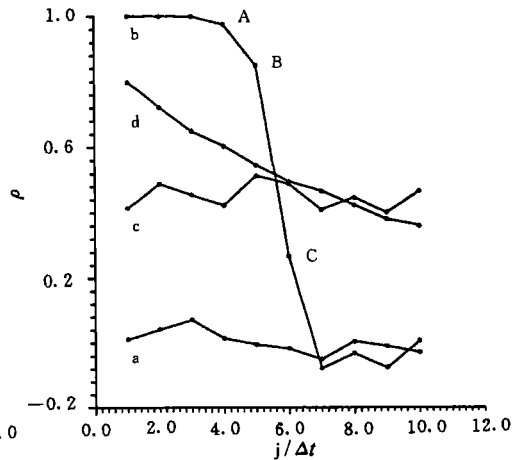


图 3 相关系数与预测步长 (j) 的变化关系

Fig.3 Relationships between the correlation coefficients and the predicting steps

a. R_n , $m = 5$; b. x_n , $m = 2$;

c. S_n , $m = 5$; d. T_n , $m = 6$

由公式 (2) 计算 ρ 随 j 的变化见图 3, 由图 3 的曲线 b 和 d 可知 x_n 、 T_n 的相关系数随 j 增加而呈指数衰减, 从而表现为混沌特征。这是由于 Kolmogorov 熵和正 Lyapunov 指数存在, 使预测轨道偏离真实轨道, 即随着 j 增加, 其预测误差积累现象更为明显。而 S_n 和 R_n 的相关系数并不随 j 增大而下降, 因为该资料没有正 Lyapunov 指数存在, 预测轨道和真实轨道基本保持等距关系, 即保持某相对稳定的相关系数, 并不随着预测步长的增加而衰减。计算 x_n 、 T_n 及其后 500 个预测值的一阶和二阶矩见表 1 (实际值 x_n 、 T_n 的一阶和二阶矩分别为 0.50、0.37、29.89、893.71), 可知实际值与预测值的前 $j = 6$ 个预测步长的一阶、二阶几乎

相等。可用第一个突变下降斜坡(BC)由公式(3)初步估计 Logistic 映射的 Kolmogorov 熵 $K_x = 0.70(1/\Delta t)$ (真实值 $\ln 2 = 0.693\dots$), 第一个缓变斜坡 AB 的 $K_x = 0.86$, 斜坡 AB 和 BC 的平均 $K_p = 0.78$ 。 T_n 的前两个斜坡平均 $K_T = 0.14(1/d)$ 。

表1 Logistic映射和赤道SST预测值的一阶矩和二阶矩

Tab.1 The first and the second moments of the predicted Logistic map and equator SST values

预测步长	Logistic映射		赤道日SST	
	一阶矩	二阶矩	一阶矩	二阶矩
1	0.50	0.37	29.89	893.17
2	0.50	0.37	29.88	892.64
3	0.50	0.37	29.87	892.38
4	0.49	0.36	29.87	891.99
5	0.48	0.35	29.86	891.42
6	0.50	0.33	29.86	891.42
7	0.51	0.30	29.85	891.24
8	0.48	0.26	29.85	891.00
9	0.49	0.27	29.85	890.87
10	0.49	0.27	29.84	890.74

3.2 带噪资料的混沌及噪音检测

采用原始资料带有加性噪音即 $x_n + \mu R_n$, $T_n + \mu R$ 检测其混沌特性。由于噪音的存在使混沌系统产生的奇怪吸引子结构发生变化, 同时使奇怪吸引子在相空间中表现出虚假的局部相似性, 因而真实的奇怪吸引子发生变形, 其统计特性相应发生变化 (Kennel *et al.*, 1992)。随着噪信比的增大, 渐渐表现出噪音特性, 促使奇怪吸引子的局部相似特性逐渐消失。尽管如此, 在较小的噪信比情况下, 奇怪吸引子虚假局部相似的统计特性并不占优, 而主要表现真实奇怪吸引子的几何特性, 因此在这种情况下可用局部非线性直接预测方法检测其混沌和噪音特性。

在选取 $\mu = 20\%$, 嵌入维数 $m = 2$, 利用非线性直接预测方法计算带噪 Logistic 资料的 ρ 与 j 的变化关系见图 4a。此时带噪 Logistic 资料在较短的预测步长范围内并没有较大相关系数的平台出现而是直接下降 (在无噪音或噪信比充分小的情况下, 在较短的预测步长范围内具有较好的可预测性, 其相关系数接近 1), 说明了由于噪音的存在使奇怪吸引子结构的局部相似性虚假化。图 4a 与图 3 中的赤道日 SST 变化具有相似的趋势, 表明了赤道日 SST 原始资料含有部分噪音。尽管如此, 总体上具有较大的相关系数, 仍具有混沌特征并非完全表现为噪音 (噪音特性见图 3a)。

为了验证该方法在噪信比增大时混沌特征的有效性检测, 分别计算了 Logistic 和赤道日 SST 的相关系数 ρ 与 μ 的关系见图 5 (其中去掉了赤道日表层水温预测绝对误差大于 2 的点数不超过总点数的 5%)。随着 μ 的增大, 混沌系统奇怪吸引子的几何结构虚假化增强, 使原混沌系统的奇怪吸引子局部相似特性逐步失真, 从而在相空间中其预测函数失去预测能力, 故其相关系数随 μ 的增大而减小。该现象反映了在噪信比增大时系统的混沌特性逐渐消失而表现出噪音特征, 即其相关数变小。一般认为相关系数在大于 0.5 的可信范围

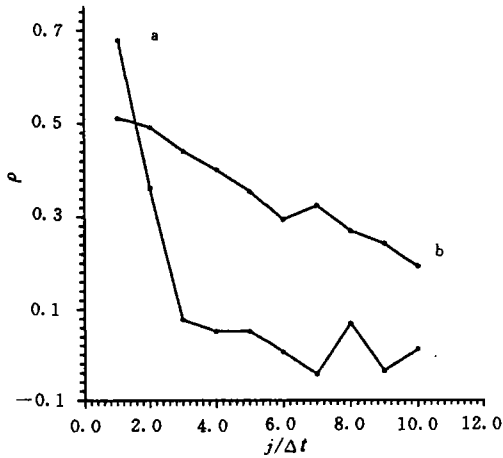


图4 相关系数与预测步长(j)的关系

Fig.4 Relationships between the correlation coefficients and the predicting steps

a. x_n , $m = 2$, $\mu = 20\%$; b. T_n , $m = 6$, $\mu = 32\%$

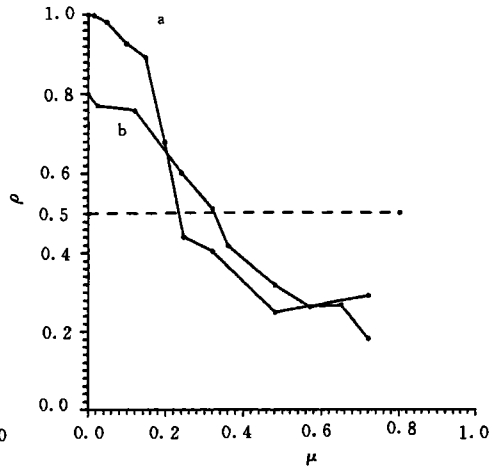


图5 相关系数与噪声比(μ)的关系

Fig.5 Relationships between the correlation coefficients and the noise-signal ratio

a. x_n , $m = 2$, $j = 1$; b. T_n , $m = 6$, $j = 1$

内检测的混沌有效, Logistic 映射资料及赤道日 SST 分别可检测含噪信比达 27%、32%。

3.3 讨论

混沌系统和噪音系统的几何结构本质上是不同的,混沌系统存在着正 Lyapunov 特征指数及有限 Kolmogorov 熵,其奇怪吸引子可用有限维相空间来描述,并且具有局部相似性,在较短的演化过程中具有可预测性。描述噪音需无穷多个自由度,噪音(白噪音)本身不存在奇怪吸引子及其局部相似性,无论演化时间长短都不具可预测性。正是由于这些性质,利用相关系数随预测时间步长及嵌入维数的变化可检测混沌及初步估计关联维数和 Kolmogorov 熵。依据混沌系统的局部相似特性的非线性直接预测方法及其相关系数来检测混沌总体上是有效的,它优于一般谱分析方法。从应用结果分析可知当实测的日 SST 资料低于噪信比为 32% 时,检测其混沌特征是有有效的。

4 结论

由本文分析可知,根据混沌系统的奇怪吸引子局部相似性及其 Kolmogorov 熵和正 Lyapunov 指数性质,利用非线性直接预测方法对混沌特征检测是有有效的。由该方法检测赤道日表层水温资料可知:实测赤道日表层水温具有混沌特征,同时含有噪音(含噪信比的大小没能检测出)。在加性噪信比为 32% 时,该方法仍有效地检测出其混沌特征(即相关系数约为 0.5),由该方法估计原始实测赤道表层水温的 Kolmogorov 熵约为 0.14 (1/d),奇怪吸引子的关联维数约为 5。

虽然该方法在实测资料的混沌检测中有一定的效果,但是最优非线性预测方法检测混沌还有待深入研究。特别是利用非线性预测方法对原始资料噪信比大小的确定需进一步探讨,这有益于人们分析实测资料的噪音含量及其混沌特性的可信程度。

参 考 文 献

- Barahona M, Poon C S, 1996. Detection of nonlinear dynamics in short, noise time series. *Nature*, 381: 215—217
- Farmer J D, Sidorowich J, 1987. Predicting chaotic time series. *Phys Rev Lett*, 59: 845—848
- Jimenez J, Moreno J, Ruggeri G, 1992. Forecasting on chaotic time series: A local optimal linear-reconstruction method. *Phys Rev A*, 45: 3553—3558
- Kennel M, Brown R, Abarbanel H D I, 1992. Determining embedding dimension for phase-space reconstruction using a geometrical construction. *Phys Rev A*, 45: 3403—3411
- Sugihara G, May R M, 1990. Nonlinear forecasting as a way of distinguishing chaos from measurement error in time series. *Nature*, 344: 734—741
- Tsonis A A, Elsner J B, 1992. Nonlinear prediction as a way of distinguishing chaos from random fractal sequences. *Nature*, 358: 217—220
- Wales D J, 1991. Calculating the rate of loss of information from chaotic time series by forecasting. *Nature*, 350: 485—488

APPLICATION OF LOCAL NONLINEAR DIRECT FORECASTING
METHOD TO DETECT OCEAN CHAOS AND NOISEWEI En-bo, SONG Qian[†], QIN Zheng-cai[†], TIAN Ji-wei[†]*(College of Power Engineering, University of Shanghai Science and Technology, Shanghai, 200093)*[†]*(Physical Oceanography Laboratory, Ocean University of Qingdao, Qingdao, 266003)*

Abstract This paper gives a method for detecting chaos and noise in chaotic time series and in time series with white noise by using a local nonlinear direct forecasting method and the coefficient of correlation between predicted and actual values. The method is applied to Logistic map data, the equator day SST data and data with noise. The results of application show that the method is better than general spectral analysis in detecting chaos and noise. When the data is contaminated by additive noise 34% in proportion to that of the signal, the correlation coefficient is about 0.5. Within this creditable correlation coefficient, the chaotic phenomenon of the time series is detected by the method.

When the rates of noise-signal increase, the correlation coefficients decrease and the noise character increases. The correlation dimension and the Kolmogorov entropy can also be determined by the method. Our results showed that the correlation dimension and Kolmogorov entropy of equator day SST data were about 5 and 0.14(1/d), respectively.

Use of the above results yielded the main conclusions below.

- 1) The application of nonlinear direct forecasting to detect the deterministic chaos in natural signals with noise is valid.
- 2) The time series of the equator day SST is chaotic and contaminated by some additive noises.
- 3) The higher the noise-signal ratio is, the smaller the corresponding correlation coefficient is.

Key words Nonlinear forecast Chaos Noise Correlation coefficient

Subject classification number P733