

中国春季气温异常变率的非线性特征

徐 同, 郭品文, 居丽丽

(南京信息工程大学大气科学学院, 南京 210044)

摘 要: 运用一种基于神经网络的非线性主成分分析法(NLPCA)对中国1951—2003年53 a春季气温距平场进行分析,结果表明,春季气温异常变率具有显著的非线性特征,当非线性主元 u 分别取最大值和最小值时,对应的空间气温异常分布型具有显著的非对称性。

关键词: 非线性主元;气温;方差

中图分类号: P468.0⁺21

文献标识码: A

文章编号: 1673-7148(2007)02-0009-04

引 言

非线性主成分分析法(NLPCA)是由 Kramer^[1]于1991年提出的一种由人工神经网络实现的非线性气候统计方法。Monahan^[2]于2000年将NLPCA应用到Lorenz数据集中,结果表明,NLPCA能得到比PCA更好的一维和二维近似。NLPCA一维和二维近似分别解释了Lorenz数据集总方差贡献的76%和99.5%,而PCA仅分别解释60%和95%。之后,Monahan^[3]运用NLPCA分析了热带太平洋海表温度和海平面气压,结果表明,热带太平洋海表温度的NLPCA一维近似要优于PCA一维近似。海表温度场存在着较显著的非线性,而海平面气压场则不明显。Heish^[4]将NLPCA用于加拿大四季气温的非线性特征研究,发现加拿大气温的非线性特征具有季节性差异,即秋、冬两季较强,春、夏两季较弱。本文采用Heish使用的NLPCA非线性统计模式对中国春季气温场异常变率进行研究,目的是探讨中国春季气温变化的非线性特征。

1 资料和方法

1.1 资料来源及预处理

所用的全国160站逐月气温资料来自国家气候中心资料集,资料长度从1951年1月到2003年12月共53 a,将3—5月气温的平均值定义为春季气温。

1.2 非线性主成分分析法

非线性主成分分析法是一种用人工神经网络提

取多变量资料集主要特征的方法。该方法由5层前向型BP神经网络^[5]实现(见图1),它由一系列平行层组成,第一层为输入层,第二层为编码层(包含

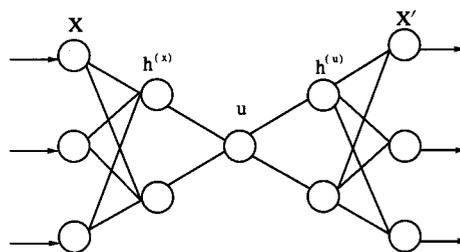


图1 5层前向型BP神经网络

m 个神经元),接着是瓶颈层(包含1个神经元),第四层为译码层,最后为输出层。数据压缩由瓶颈层实现,用 u 表示瓶颈层神经元即非线性主元。本文所用的NLPCA有4个输入和输出, m 取2和3,为了防止网络过适应,在剩余函数中加入权重系数 P 。网络的非线性最优化由MATLAB自带函数“fminu”完成。本文中网络迭代次数为50,随机选取权重和偏移系数,同时随机选取20%的数据用于网络测试,余下的用于网络训练。当测试数据的均方误差比训练数据大时,网络停止运行以避免得到过适应的解。从0~0.18每隔0.006取一惩罚系数值 P ,隐层神经元数重复地取2、3进行运行,选择均方误差(即整个资料集上未被解释的方差百分比)最小的解为所用解。未被解释的百分比 FUV (Fraction of Unexplained Variance)计算公式为

$$FUV = 1 - \frac{\sum_{i=1}^m Var(x'_i)}{\sum_{i=1}^m Var(x_i)} \quad (1)$$

2 EOF 分析

在 NLPCA 分析之前,先对春季的温度距平场进行 EOF 分析(即 PCA 分析),这样做利用了 EOF 在数据压缩降维方面的优势。EOF 第一模态解释了总方差的 42.7%,前 4 个模态共解释了总方差的 80.3%。图 2 为春季 EOF 前两个模态的空间型及与之对应的时间序列。

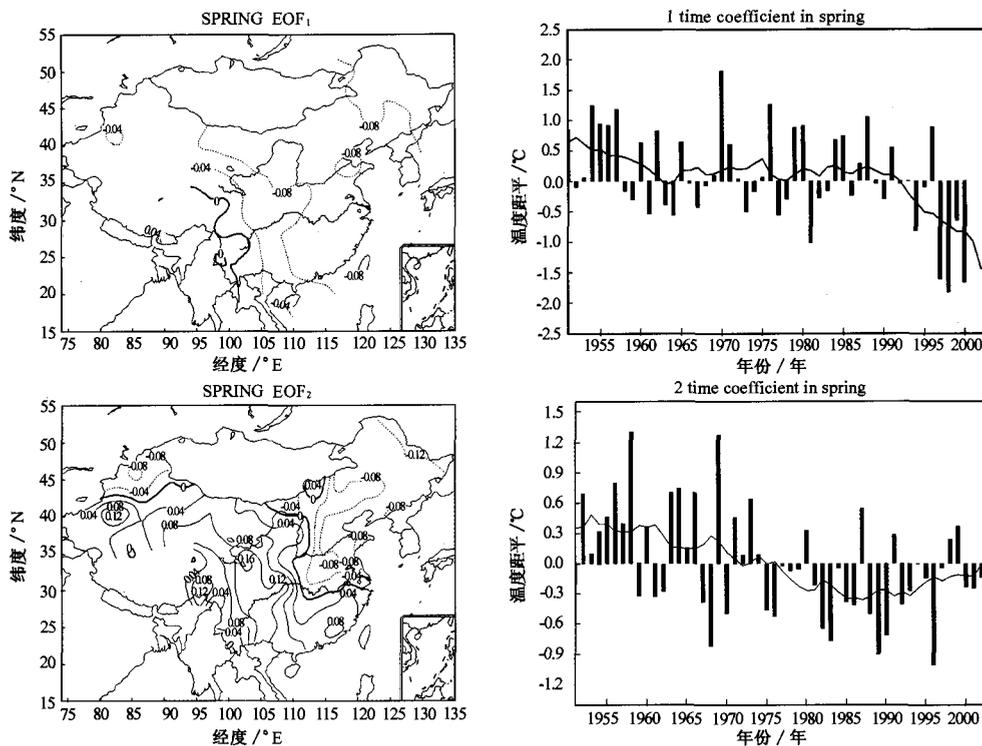


图 2 春季气温 EOF 前两个模态的空间型及与之对应的时间序列

3 NLPCA 第一模态

将前 4 个主成分(PC)作为非线性主成分分析的对象,得到春季 SAT 距平的 NLPCA 第一模态,为了更直观地了解和分析非线性特征,采用三维空间图(见图 3)。从图 3(d)能看到 $PC_1-PC_2-PC_3$ 三维空间中的 NLPCA 曲线,该曲线由一系列重叠的空心方框构成(PCA 模态表现为一直线),同时从图 3(a、b、c)能看到其在 PC_1-PC_2 , PC_1-PC_3 和 PC_2-PC_3 平面上的投影。与 PCA 相比较,可以在 PC_1-PC_2 平面上看到一条明显的曲线(PCA 表现为直线),表现出较强的非线性。衡量非线性强弱的一个重要参数就

如图 2 所示,春季的 EOF 第一模态空间图上,全国气温异常变化一致为负,与空间分布相应的时间序列 PC_1 有着显著的年际变化特征(直方条)和年代际变化特征(实线)。年代际变化主要表现为:20 世纪 50 年代初至 80 年代末为一冷期,而从 90 年代开始,气温呈不断上升趋势。EOF 第二模态空间图上,春季的气温空间分布特征主要为东西向分布,其中,负距平区域主要位于我国长江以北及 110°E 以东地区,负距平中心主要位于华北和东北地区。对应的时间序列从 20 世纪 50 年代初到 70 年代中期之间,空间分布特征基本为正位相,70 年代中期以后基本为负位相。

是均方误差(MSE)的比率,也就是 NLPCA 第一模态未解释的方差百分比与 PCA 第一模态未解释的方差百分比的比值,该比值越小,说明非线性越强,反之则越弱,若两者比率为 1,则表明只有线性解。春季的 MSE 比率为 0.738,说明非线性表现较为明显。与 PCA 第一模态提取的特征相比,NLPCA 第一模态解释了总方差的 55.2%,比 EOF 第一模态高出近 13 个百分点。

从图 3(a)可以看出, $PC_1(t_n)$ 、 $PC_2(t_n)$ 的非线性主成分近似 $X_1^1(t_n)$ 、 $X_2^1(t_n)$ 的样本点(空心圆点)在该图上聚集为一条曲线。当 $X_1^1(t_n)$ 在 $[-5, +5]$ 内取绝对值近似的正、负值时, $X_2^1(t_n)$ 均为正值;当 $X_1^1(t_n)$

在 $[-15, -5]$ 和 $[5, +15]$ 区域内取绝对值近似的正、负值时, $X_2^1(t_n)$ 均为负值。 PC_1 和 PC_2 之间的这种非线性关系结合 EOF_1 、 EOF_2 空间型就能反映气温异常时空变率的不对称性。 EOF_1 图上, 东北地区为负距平区, 因此 PC_1 取强的正、负值, 分别表示气温负异常和正异常, EOF_2 图上, 东北地区为负距平区, 根据 PCA 原理, 由于前两个模态解释的方差贡献较大,

因此可用 PC_1 乘 EOF_1 加 PC_2 乘 EOF_2 近似的表示原始 SATA 场。由于 PC_1 和 PC_2 之间的非线性关系主要表现为当 PC_1 取强的正负值时, PC_2 都倾向于取强负值, 于是这两个成分合成表示出的东北地区的 SATA 特征为, 气温偏高时的正异常值大于气温偏低时的负异常值, 振荡在强度上的这种不对称性正是通过 PC_1 和 PC_2 之间的这种非线性关系反映出来的。

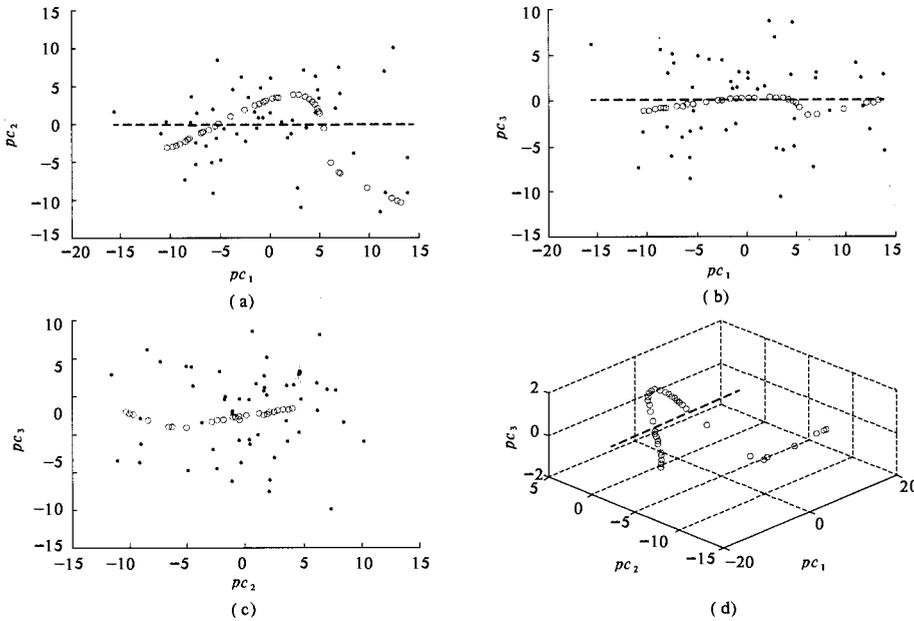


图3 春季气温距平场 NLPCA 第一模态在三维空间的图像

非线性主成分分析法的另一个优势在于能够表现出不同空间振荡型的结构。将神经网络瓶颈层神经元输出的时间序列称为非线性主元 u , 对于一个特定的非线性主元 u , 利用神经网络将其从瓶颈层投影到输出层的 X^1 。 u 可被看作是 NLPCA 曲线的独立曲线调整系统, 因此, 每个 u 都对应于前 4 个 PC 模态的一个空间型。图 4 为非线性主元 u 分别取最大值和最小值时对应的空间型, 气温正负异常

的区域和异常中心都没有明显的对称关系。图 4 (b) 中, 正异常区域主要位于新疆南部和华北、东北大部分地区, 而图 4 (a) 中, 新疆西部的负异常区强度较弱, 并且没有出现相应的异常中心, 而华北、东北区域也没有出现相应的强负异常区。这表明新疆南部和华北、东北地区气温异常偏高的程度要强于异常偏低的程度。这与通过 PC_1 、 PC_2 之间的非线性关系讨论所得的结果是一致的。

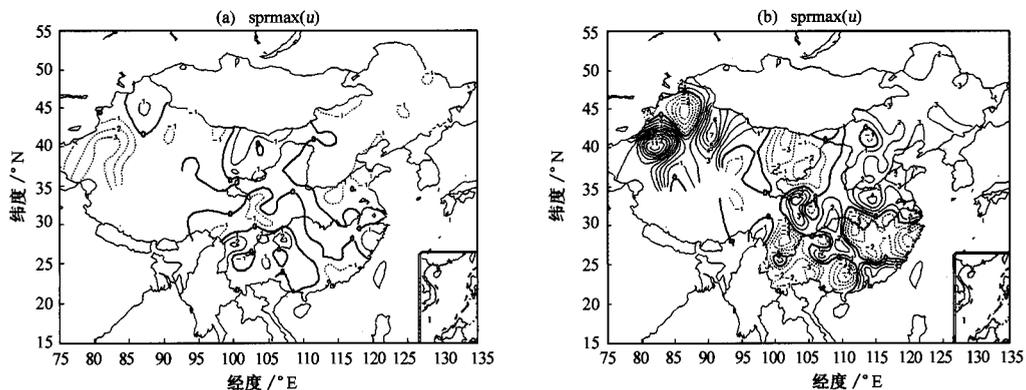


图4 非线性主元 u 分别取最大值和最小值时对应的空间分布型

4 结 论

①NLPCA 第一模态结果表明,春季气温异常变率具有显著的非线性特征,且这种非线性特征主要表现为前两个主分量之间的关系,即 PC_1 的近似 $X_1^1(t_n)$ 取较小的正、负值时, PC_2 的近似 $X_2^1(t_n)$ 均为正值,相反当 PC_1 的近似 $X_1^1(t_n)$ 取较大的正、负值时, PC_2 的近似 $X_2^1(t_n)$ 均为负值。 PC_1 和 PC_2 之间的这种非线性关系结合 EOF_1 、 EOF_2 空间型就能反映出气温异常时空变率的不对称性。

②非线性主元 u 取最大值和最小值时对应的气温异常空间型没有反对称的关系, u 取最小值时,正异常区域主要位于新疆南部和华北、东北地区; u 取最大值时,新疆西部的负异常区强度较弱,而华北、东北区域也没有出现相应的强负异常区。这也从另一个角度证明了春季气温异常的非线性特征非常显著。

③通过 NLPCA 方法可以得到传统线性统计方法不能得到的结论,但也正是由于它的复杂性,不能

期望这种非线性统计方法可以完全代替一些传统的统计方法,比如 EOF、一点相关图等,但是应该相信,该统计方法在大气科学多元统计分析中有很好的应用前景。

参考文献

- [1] Kramer M A. Nonlinear principal component analysis using autoassociative neural networks[J]. AIChE Journal, 1991, 37:233 - 243.
- [2] Monahan A H. Nonlinear principal component analysis by neural networks: Theory and application to the Lorenz system[J]. Clim,2000, 15(2):821 - 835.
- [3] Monahan A H. Nonlinear principal component analysis :Tropical Indo - Pacific Sea Surface temperature and Sea Level Pressure [J]. Clim,2001,14(1):219 - 233.
- [4] Wu A , Heish W W. Nonlinear characteristics of the surface air temperature over Canada [J]. Journal of Geophysical Research, 2002, 107(21):1 - 15.
- [5] 吴洪宝,吴蕾. 气候变率诊断和预测方法[M]. 北京:气象出版社,2005:55 - 75.

Non-linear Characteristic of Spring Temperature Anomalous Changes over China

Xu Tong, Guo Pinwen, Ju Lili

(School of Atmospheric Sciences, NUIST, Nanjing 210044, China)

Abstract: Chinese spring surface air temperature anomaly (SAT) during 1951–2003 was investigated by applying a neural-network-based non-linear principal component analysis (NLPCA) method. The results showed there was a distinct non-linear characteristic in China spring temperature anomaly changes, and the space distribution was obviously unsymmetrical when the u was selected as its maximum value and minimum value.

Key words: non-linear principal component; temperature; variance