

人工神经网络方法在沙尘暴短期预报中的应用

赵翠光

(北京大学地球物理系,北京 100871)

提 要

使用人工神经网络方法建立了我国沙尘暴短期预报模型,该神经网络模型的输入因子是几个物理量场 REOF 展开的一些时间系数,输出为我国有无沙尘暴。结果表明 REOF 展开技术和人工神经网络方法两种方法的结合对于预测沙尘暴是可行的。

关键词: 人工神经网络 REOF 展开 沙尘暴预报

引 言

沙尘暴天气是我国西北地区和华北北部地区出现的强灾害性天气。它能加剧土地沙漠化,对大气环境造成严重污染,对生态环境造成巨大破坏,对交通和供电线路产生重要影响,给人民生命财产造成严重损失。我国每年由沙尘暴产生的土壤细颗粒物流失严重,并且绝大部分粒径有增加的趋势,加之土地资源超载的局面短期内难以改善,随着全球气候的变暖,水资源短缺的矛盾将日趋尖锐,因此,沙尘暴对人类的危害也将随之增大。对沙尘暴作出较为准确的预报对防灾减灾具有十分重要的实际意义。我国气象工作者针对沙尘暴已有很多研究工作^[1~5],大多数是侧重于对产生沙尘暴天气的天气气候学特征进行研究和诊断分析,总结归纳其大气环流特征,或者是通过对个别沙尘暴个例的特殊成因分析找出沙尘暴预报指标。客观

的、定量的、自动化的预报方法还很少,特别是针对我国大范围的预报就更少。本文希望通过此工作,能够对沙尘暴的预报方面有所贡献。

产生沙尘暴天气的因素是多方面的,沙源是其物质基础,除此之外,强风和不稳定层结是产生沙尘天气的基本条件。春季,我国西北、内蒙古等地,由于地表裸露,浮土层厚,在有利气象条件下极容易形成沙尘天气。因此本文从有利产生沙尘天气的环流形势、天气系统及一些物理量着手,尝试运用人工神经网络方法客观地制作产生沙尘天气的潜势预报。人工神经网络是近十几年得到迅速发展的一门非线性科学。它以抽象的人脑构造基本单元组成,模拟人脑的部分思维过程,例如人工神经网络所能进行的感知、学习、自适应、联想、模式识别等功能^[6]。人工神经网络有多种模式,这里采用目前应用

广泛的反向传播(back-propagation,简称BP)人工神经网络,利用选取的预报因子,经过样本训练,建立我国北方沙尘暴预报潜势模型。

1 资料处理

为了预报大范围的沙尘暴,需要大量的格点资料。如何从大量的格点变量中选出具有天气学意义和数学意义的因子是预报过程中的重点。黄嘉佑等人提出用转动主分量分析方法提取因子做短期天气的动力统计预报。EOF方法提供了压缩空间场资料成为少数一些主要模态的方法。为了揭示场中不同地理区域变化的特征,使用REOF方法对神经网络模型的输入因子作前处理。同时对输入因子的处理也使得神经网络模型的输入节点数大大减少,如进一步限制隐层节点数,则神经网络模型的自由参数可望得到相对于样本长度来说较有限的网络自由参数,从而使神经网络模型的非线性不稳定得到明显抑制。

1.1 预报因子选取

利用NCEP再分析资料,范围取 $15^{\circ}\sim 70^{\circ}\text{N}$ 、 $50^{\circ}\sim 150^{\circ}\text{E}$,格距为 $2.5^{\circ}\times 2.5^{\circ}$ 。由于资料的限制,时间段取为1988~1994年的2月11日~6月10日。

以往绝大多数工作是将单个变量场作为研究对象,从理论上讲,REOF分解所用的样本观测矩阵(原始场)不受任何条件的约束,所以可以有各种各样的形式,原始场也可以包括几个变量场,由此得到的特征场包含几个要素的分场,它们不仅反映了各个变量的空间分布特征,而且表示了变量场间的相关结构。因此选大尺度背景场500hPa高度场、700hPa东西风场、700hPa南北风场、850hPa温度场做联合REOF分解,4个格点场的空间点数共为 $m=41\times 23\times 4=3772$,时间序列长度为 $n=121\times 7=847$ 。根据过去的研究,海平面气压、500hPa与700hPa之间的温度递减率、 $\partial\theta/\partial p$ 和风垂直切变,700hPa的温度平流与沙尘暴有密切关系。为了选择更全面的因子,对以上物理量分别进行REOF分解,每个格点场的空间点数为 $m=41\times 23=943$,时间序列长度为 $n=121\times 7=847$ 。考虑到沙尘暴是小概率事件,代表沙尘暴的

类型不一定集中在前几个特征向量,因此在因子的选取上,取与沙尘暴日站数相关较好的时间系数作为预报因子。

1.2 预报对象确定

预报对象为相应时段的全国681个站每日沙尘暴资料,本文将5个站以上出现沙尘暴定为1,小于5个站定为0。

2 模型建立和结果分析

本文采用三层BP人工神经网络^[7]。神经网络建模过程是个反复训练和测试的过程,目的是为了选择好的模型以达到最好的预报效果。建模过程先选定一种训练集合和测试集合,确定一种网络拓扑、迭代次数和训练参数作为初始模型。训练完成后,用测试样本集合测试模型的预报效果,若结果不好,重新选定网络拓扑、迭代次数和训练参数,再训练、测试,反复试探,直到测试效果满意为止。沙尘暴预报模型建模框架如图1。

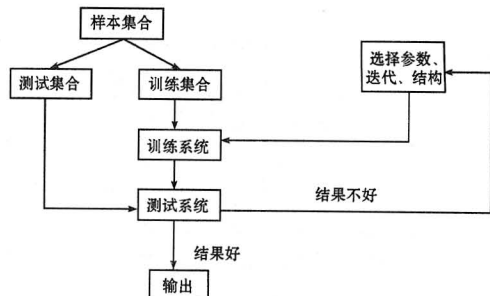


图1 建模过程流程图

为了达到较好的预报效果,本文从网络的拓扑、迭代次数对模式进行选优。输入节点数为56个,隐层1层。输出节点2个,其中第一个节点当预报对象为0时设为1,为1时设为0,第二个节点预报对象为其本身。训练参数:学习率取0.4,惯性系数取0.02。训练样本集合:1988~1993年样本集,共有样本726个,其中沙尘暴日162个,占样本数的22%,非沙尘暴日564个。测试样本集合:1994年,共有121个样本,沙尘暴日24个,占测试集的20%。我们变换隐层隐节点数及迭代次数进行多次试报,预报的成功界限指数定义为:

$$CSI = \frac{c_f}{c_f + w_f} \times 100\%$$

其中 c_f 为正确报出沙尘暴的日数, w_f 为试

漏报和空报之和。

表1给出部分实验结果,从表中可以看到隐层节点数不是越多越好,因为如果隐层节点数过多,虽有利于精确地逼近每一样本点,使网络识别率高且记忆性好,但其振荡剧烈使网络推广能力差。而过少的隐层节点数,将导致网络缺少灵活性和分类能力。所以我们需要在实际应用中,经多次试验后才能确定网络的隐层节点数。迭代次数也不是越高越好,这可能是过于迁就某些不典型、不精确的样本,虽然拟合较好但偏离总体规律造成的。但是,如果网络训练次数过少,识别率太低,同样会造成推广能力较差。表中模

表1 实验结果

模型	迭代次数	隐层节点数	正确报无次数	正确报有次数	空报次数	漏报次数	CSI (%)	准确率 (%)
模型1	100	10	89	13	8	11	41	84
模型2	300	10	84	16	13	8	43	83
模型3	500	10	83	15	14	9	39	81
模型4	300	6	84	9	13	15	24	77
模型5	300	20	83	13	14	11	34	79

本实验在建模及结果分析过程中用的是分析场的资料,在实时预报中代入预报场即可。

3 结 语

用 REOF 展开技术从大范围的大气要素场中提取主要信号,使神经网络模型的输入节点数大大减少。由于人工神经网络的非线性处理能力较强,对于沙尘暴这样非连续量用此方法是可行的,从试报结果来看有一定的预报能力。从上述工作的效果和客观性来看,它有望成为沙尘暴短期业务预报方法。另外输入因子的选取也是非常重要的,进一步引入其它有物理意义的因子从理论上可改善该方法的预报效果。

型2预报效果较好,24个沙尘暴日报出了16次,预报成功界限指数为43%,准确率为83%。但仍有8次漏报和13次空报。这是因为所选的因子没有考虑下垫面的情况,在一定的下垫面下即使高空形势特征不明显也有可能产生沙尘暴,这样上面的客观预报方法就无法将沙尘暴报出。此外所选因子的特征出现了也未必出现沙尘暴,因为除上面选择的因子外,影响沙尘暴预测的因子还有很多,大网格的模式资料还不能包含足够的沙尘暴预报信息。因此进一步引入一些其它有物理意义的因子,包括反映下垫面的因子,从理论上可改善BP方法的预报效果。

参考文献

- 1 王式功,杨德保,孟梅芝等. 甘肃河西“5.5”黑风天气系统结构特征及其成因分析. 中国沙尘暴研究. 北京:气象出版社,1997:62~64.
- 2 徐建芬,牛志敏,陈伟民. 我国西北地区4·5沙尘暴天气研究. 中国沙漠,1996,16(3):281~286.
- 3 项续康,江吉喜. 1996西北地区强沙尘暴成因的中尺度分析. 高原气象,1996,15(4):448~455.
- 4 赵光平,王凡,杨勤. 宁夏区域特强沙尘暴天气成因及其预报方法研究. 中国沙尘暴研究. 北京:气象出版社,1997,52~58.
- 5 保广裕,高顺年,戴升. 西宁地区沙尘暴天气的环流特征及其预报. 气象,2002,28(5):27~31.
- 6 黄崇福,王家鼎. 模糊信息油画处理技术及其应用. 北京:北京航空航天大学出版社,1995:175~181.
- 7 陈守余,周梅春. 人工神经网络模拟实现与应用. 北京:中国地质大学出版社,2000.

Application of Artificial Neural Network to Sand-storm Forecast in North China

Zhao Cuiguang

(Department of Geophysics, Peking University, Beijing 100871)

Abstract

A forecast model of sand-storm in China is established by BP artificial neural network. The inputs of the model are the time coefficients of REOF in a few of physical fields, and the outputs are sand-storm or no sand-storm in China. The result shows that combination of these two methods has better effect on forecasting sand-storm.

Key Words: artificial neural network REOF approach sand-storm forecasting