

短期气候(月、季、年) 逐月降水预测系统研究

盛永宽

(国家气象中心,北京 100081)

提 要

描述了短期气候逐月降水预测系统的概念及其实现。首先讨论了把气候模式与人工神经网络结合起来开发一个气候预测支持系统的必要性和可能性。然后,给出了系统设计的框图,并对系统结构作了详细的讨论。最后,提出了有关预测结果分析和进一步完善的看法。

关键词: 气候模式 神经网络 降水预测系统

引 言

人们已经看到,月季年短期气候的预测问题已显露端倪,它正处于技术上的可行性与社会上的迫切需要性的交汇点上。技术上可行,指的是我们已跨入可以用海气耦合模式模拟气候变化这样一个新时期的门槛,可以验证表达时间尺度长达一年的大自然变化的模式性能^[1]。社会迫切需要,指的是现代化国民经济每年防灾减灾的预案已不能基于年度降水偏多偏少这样一些定性的粗略估计,而需要工程化定量预测系统的支持。时代的变迁必然引起科学技术的有关变革。月季年短期气候预测系统的研究,就是在这一背景下进行的探索。

从近代科技的发展来看,研究解决短期气候预测问题的技术路线应从三个方面^[2]着手:一是选用适当的气候模式,对全球大气环流作确定性的计算,这是基础;二是运用多信息集结技术和人工智能技术,求解大气环流场与气候要素场(重点是降水)之间复杂的非线性关系,这是关键;三是研究高效、灵活、开放的交互式气候预测集成系统,这是整体综

合判断所必不可少的决策支撑工具。本项研究就是在这三个主导方向上所作的努力。

当代气候学在推理定量研究气候系统方面所取得的进展表明,必须用比较完善的有坚实物理基础的气候模式才能定量地考虑气候系统各成员间相互作用的气候变化过程,才能准确理解气候变化的机制,从而使气候预测成为可能。但是,气候系统属于复杂巨系统。当人们寻求定量方法处理复杂行为时,不能只注重数学模型的精确处理,而忽视经验的历史的判断或解释,需要重视寻求应用人工智能技术的突破口,这是我们系统建模的指导思想。

根据国情和实际的需要,系统应该是开放的和工程化的。也就是说,系统必须是可操作的,规模可调的。既可以在单台工作站上运行,也可在松耦合并行处理的工作站簇上运行,还可以在巨型机网络上运行。随着模式的更新,能易于更换模块而综合集成。本系统就是试图设计成这样一种工程系统:气候预测支持系统 CPSS (Climate Prediction Support System) 的原型。

1 气候模式与神经网络的结合

近十多年来,用大气环流模式逐日积分,制作月平均环流场的预测,已是普遍采用的方法。而且大家认为,在半年至年度的预测中,大气环流的初始状态仍然起很大的作用。因此,必须有一个海洋动热力过程,冰雪演变过程以及它们和大气过程相互作用并能同时计算的,即海-气耦合的气候模式。

我们选用的气候模式,是基于美国 Oregon State University 气候研究所发展起来的并经 Schlesinger 和 Zhao^[3]改进的全球大气环流(二层)与海洋混合层(60M)耦合模式(OSU AGCM/MLOM)。在1990年国际政府间气候变化专门委员会(IPCC)的“气候变化科学评估报告”中,该模式列为目前世界上的(十五个)主要气候模式之一。

模式的大气部分,采用 Gates^[4]的二层大气环流模式,水平分辨率为 $4^{\circ} \times 5^{\circ}$,垂直方向用 σ 坐标并考虑了地表边界层处理,摩擦和非绝热过程,模式预报大气运动速度、温度、地面气压、地面温度、水汽、土壤水份与云量。

模式的海洋部分,采用 Pollard^[5]的海洋模型简化方案(version 1),不考虑卷人/出和深海层的影响,只具有纯热力的计算。海洋模式预报海洋混合层的温度和海冰厚度。

大气和海洋的耦合方式是,大气环流模式考虑了地表边界条件的强迫影响,按地表型态进行相应的参数计算。对于陆面和海面有不同的热力计算,水面和冰面也赋以不同的反照率,能按海洋模式的预报值,随着每个时间步长而不断变化,这就在一定程度上模拟了海-气相互作用,反映了气候系统演变的物理过程,提供了制作月、季、年短期气候预测的物理基础。

这个气候模式,与迄今为止大多数全球海气耦合模式类似,都是在粗网格分辨率 5×5 的量级上,均存在类似的系统误差,或称模式气候漂移。与实测值比较,热带海洋太冷,高纬度南大洋太暖,副热带高压太强。至

于模式模拟的降水量则差距更大,不能直接作为降水总量的预测。但是,这个气候模式却有相当水平的模拟能力。如何运用模式的模拟能力,是预测系统的关键。

一个气候模式的模拟能力,能够比较真实地反映气候的月季年变化,这说明即使它的计算值不是接近实际的预测值,但它隐含着接近实际的有某种物理意义的信息。虽然我们对隐含的物理信息还缺乏知识(这种知识的不完备性,在一定意义上也就是模式物理结构的不完备性),然而我们可以把气候模式输出的模拟值与历史的观测值之间建立某种映射关系,利用这种映射可以调整气候模式的系统误差,进而作出比模式模拟更加逼近客观实际的预测。

这里,我们看到在气候预测中有两类不同性质的问题,一类是强结构化问题(Well-structured problem),即有较充分的知识,较完善的理论体系和形式化方法,人们称之为强方法,通常能够得到唯一的精确解。求解气候模式属于此类。另一类是弱结构化问题(ill-structured problem),即缺乏足够的知识,在复杂的相互作用下具有很大的模糊性和不确定性,求解的目标不可能是唯一解,而只能是满意解。气候模式系统误差的调整属于此类。现在人们越来越重视弱方法,人工神经网络 ANN(Artificial Neural Network)是弱方法中强有力的工具,基于它的特有功能,已在不确定性系统的智能控制中得到广泛的应用。

神经网络是由众多的神经元连接而成的,每个神经元的结构和功能并不复杂,但网络的动态行为却是极为复杂的,因此它能表达实际物理世界的复杂现象。在现有多种神经网络中,BP(Back-Propagation)反向传播网较为成功。

N. Robert 已经证明^[6],任何一个平方可积函数,均存在一个三层 BP 网络,可在均方差意义下逼近。这显示了 BP 网络实现函数

映射的能力,并为应用于系统控制奠定了理论基础。

气候模式的系统误差正是气候系统不确定性的客观反映。提高短期气候的预测能力,无疑要用强方法,依靠模式物理过程的逐步完善。但是,调整和减少气候模式的系统误差,则必须依靠弱方法。运用神经网络来调整系统误差,把气候模式的输出作为神经网络的输入,用神经网络的输出作预测,是一个有生命力的发展方向。

气候模式系统误差的调整问题,可以看成是从动力学系统中获得信息用来跟踪期望输出量的控制问题。也可以说,把误差调整作为模式识别来处理,使其完成一个模式识别控制器的映射功能,即:输入变化信息 X (模式状态),得到相应的控制信息 U (模式类别)。BP 网络具有这种函数映射或者说函数逼近的功能,实际上它是一个自适应模式识别系统,能够通过学习训练,自组织实现样本集所给定的输入—输出映射关系。

我们的动力学系统是气候模式,误差调整控制器的输入就是模式气候 $X(t) =$

$$\begin{bmatrix} X_1(t) \\ \dots \\ X_n(t) \end{bmatrix}, \text{ 而输出就是实现追踪目标的控制}$$

信息,即需要逼近真实气候的预测值 $U(t) =$

$$\begin{bmatrix} U_1(t) \\ \dots \\ U_n(t) \end{bmatrix}. \text{ BP 网络控制器运行之前的训练过}$$

程,就是通过时间序列样本集 (X_i, U_i) 不断调整权重,使 BP 网络的误差达到足够小的程度,从而学会(记忆)控制气候模式能够得到逼近真实气候的输出值,然后将它应用到实际的气候预测工作中去。

用神经网络来处理定量预测问题的突出优点是它属于一种黑箱体系,使用时不必完全弄清楚处理问题的机理,也不必探讨是否适合所研究的对象,而只强调联想记忆。但这不等于应用时可带随意性。在选择建立怎样

的映射关系时,必须遵循所研究的物理机制。

降水系统是气候系统中最为活跃,相互作用最为复杂的子系统之一。降水的分布是多因素形成的。因此,预测降水需要用多种方法,适当选择和建立能够表征影响降水的多种映射关系来制作组合预测。假定降水预测有 n 种预测方法,则可用映射

$$x \in X \subset R \quad y \in Y \subset R'$$

表示第 i 种预测模型所确定的预测方法,其中 X 表示历史数据域, Y 表示预测数据域。组合预测是对于不同的预测的非线性组合函数:

$$y_i = \Phi(t) = \Phi(\varphi_1(t), \varphi_2(t), \dots, \varphi_n(t))$$

其中 $\Phi(t)$ 是非线性映射,在某种测度的度量下,其误差满足

$$\|\Phi(t) - T\| \leq \min_{i \in \{1, n\}} (|\varphi_i(t) - T|)$$

T 是实际观测值,而 BP 神经网络的信息综合功能可以实现非线性映射 $\Phi(t)$ [7]。

在本系统中, φ 主要考虑以下几项:

a. 大气环流场(北半球 500hPa)是影响我国降水的主导因素。我们用模式输出场 GH' 与实测降水场 SR 建立映射关系:

$$\varphi_1 = BP(GH' - SR)$$

b. 模式输出的序列较短,还需要用长序列实测环流场 GH 与实测降水场 SR 建立映射关系:

$$\varphi_2 = BP(GH - SR)$$

c. 海温场(赤道太平洋)是影响我国降水的强信息区,需建立模式输出海温 SST' 与实测降水 SR 之间的映射关系:

$$\varphi_3 = BP(SST' - SR)$$

d. 模式输出降水场 RR' 与实测降水场 SR 之间的映射关系:

$$\varphi_4 = BP(RR' - SR)$$

为消除不同季节的影响, φ 的各 BP 网络模型采取分月建立。同时,本年度的逐月预

测模型也与下年度即年际的逐月预测模型分开,因为月际预测和年际预测的气候模式性能有很大差别。

2 系统结构及其实现

根据上节所述,本系统的总体框图如下:

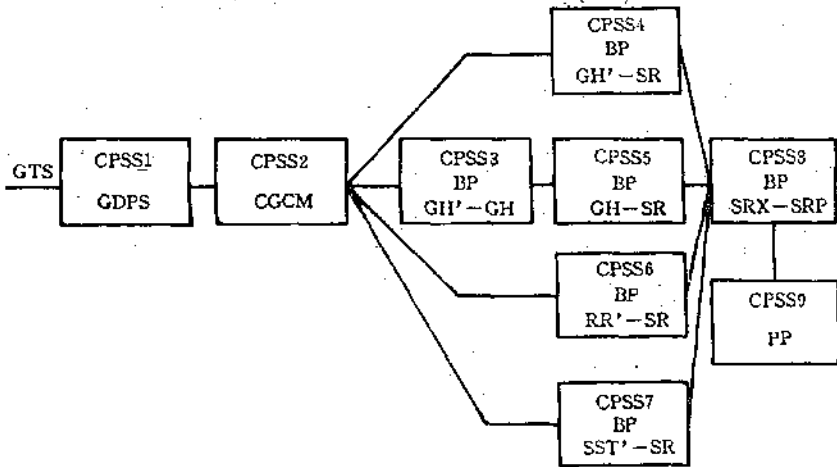


图1 系统框图

2.1 初始场的预处理

月季年时间尺度的气候预测,要求气候模式的计算区域必须是全球范围。因此,系统的第一组成部分是全球资料处理系统(GDPS)的初始场获取。

初值对动力学模式计算至关重要,所以初始场要尽可能用实测数据。

本系统初始场的建立主要输入下列资料:

(1)从全球通信系统GTS上取得2月28日实测值:风(u, v),气温(T),比湿(Q),地面气压(PS),地面温度(TS),海面温度(SST)。模式输入400、800hPa分别由300、500hPa和700、850hPa内插而得。

(2)气候平均值:地面反照率,土壤湿度,地表温度,积雪量,海冰等。

2.2 公式计算

系统的第二组成是海—气耦合模式(CGCM)。模式计算从每年2月的最后一天(28日)为初始场开始逐日积分,到第31天求31天平均(总量)作为3月份的预测,以此类推,直到第二年的9月,年度全程共计19个月。

2.3 神经网络

本系统重点考虑我国主汛期降水预测。因此,神经网络分别按5—8月的6个模块(CPSS3—8)综合组成,每个月均含有BP($GH'-SR$),BP($GH-SR$),BP($RR'-SR$),BP($SST'-SR$)4组网络,最后由BP($SRX-SRP$)组合网络模型作出逐月降水预测(SRP)。模块CPSS5是长序列(1951—1992年)环流场—降水场BP网络,它的输入需要将模式输出 GH' 经系统误差调整后的格点值,即BP($GH'-GH$)的输出,才能作为BP($GH-SR$)的输入。

2.4 后处理

系统的后处理(PP Post-Processing)主要包括图形处理、可视化和统计检验。

业务上检验模式性能常用“相关量度”,即两组数据间空间位相(phase)差的量度。两组数据是模式预测值和真实观测值,相关是距平间的相关,故称距平相关AC(Anomaly Correlation)。

$$AC = \frac{\sum (X_p - \bar{X}_p)(X_o - \bar{X}_o)}{[\sum (X_p - \bar{X}_p)^2 \sum (X_o - \bar{X}_o)^2]^{1/2}}$$

其中, p 和 o 分别表示模式预测值和真实大气观测值, X 为真实大气(1951—1992年)的平均值。

2.5 系统实现

整个系统由 Fortran 和 C 语言编程, 在 UNIX 环境下运行。系统可以全部装载在 Sun SPARC-1 工作站上, 运行一个模式月大约 5 小时, 训练一个神经网络平均 96 小时。也可将系统装载在 Sun670 服务器上(双 GPU), 通过网络, 图形显示在工作站上, 运行时间可以比工作站缩短 1/3—1/2。

3 预测检验与分析

我们用本系统作了 1990—1994 年 5—8 月共 20 个月我国 160 点逐月降水量的预测。同时也作了年际的, 即 1991—1995 年 5—8 月逐月降水量预测。因为各 BP 网络均以 1990—1992 年为训练集, 1993—1994 年为预测集。年际的则以 1991—1993 年为训练集, 1994—1995 年为预测集。因此, 只有 1993—1994 年(年际 1994—1995 年)是完全的预测。现将其全部 AC 检验值列于附表。

附表: AC 检验值

年份	5月		6月		7月		8月	
1990	0.9638		0.9633		0.9874		0.9895	
1991	0.9826	0.9805	0.9845	0.9825	0.9945	0.9894	0.9843	0.9875
1992	0.9814	0.9749	0.9653	0.9632	0.9589	0.9747	0.9907	0.9870
1993	-0.1301	0.9798	-0.2469	0.9627	-0.0042	0.9713	0.0438	0.9949
1994	0.1349	-0.0189	-0.1003	-0.0177	0.1090	0.0761	0.0242	-0.1233

从 AC 检验值看预测准确率不高, 这主要是样本数过少所致。神经网络的预测能力与样本、结构、算法等许多因素有关, 当然首先要有足够比例的样本分布, 否则将直接导

致网络预测能力的下降。但有意义的是, 尽管样本不足, 夏季主要雨带(>200mm)的基本特征已能预测出来。以 1994 年 5—8 月为例, 从图 2 可以看出, 5—6 月主要雨带稳定在华

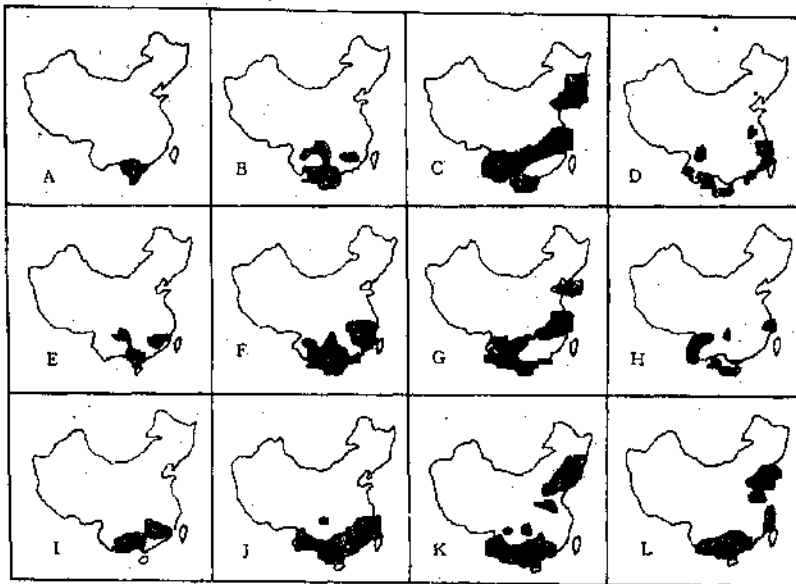


图2 预测与实况(月降水量>200mm 落区)对比

A—D 1994.5—8 预测(月季) E—H 1994.5—8 预测(年际) I—L 1994.5—8 实况

南(广西)一带,7月华南仍有雨带维持,同时主要雨带发展到华东、东北(辽、吉)一带,与实况比较接近,华南(广西)、东北(辽、吉)严重洪涝。一般来说,我国主汛期月降水量 $>200\text{mm}$ 的落区(主雨带)若连续一两个月稳定少动或中心雨量较大,则该地区容易发生洪涝。如果将本系统的降水输出,作为水(灾)情模型的输入,则有预测洪涝的使用价值。

大家知道,模式识别与相关系数分析不同,它不随时间变化,可随序列的延长和新资料的加入而提高识别能力。历史拟合较好,但预测(泛化)能力不够的问题,即训练后的网络对未在训练集中出现过的状态作出正确反应的能力问题,可以通过进一步合理构造网络(如扩展映射关系等)得到改善。至于气候变化的转折,则主要靠气候模式性能的改进,但神经网络与知识发现的结合,也有识别气候转折变化的潜力。因此可以预见,今后能从理论上和实践上表明:随着样本序列的延长,映射关系的扩展,新资料的增加,模式(包括气候模式和人工智能模式)性能的改进,模式气候可以逐步逼近真实气候,使预测准确率

不断提高,在不远的将来达到业务化目标。

致谢:本研究在气候模式和神经网络方面得到中国气象科学研究院赵宗慈同志和国家模式识别开放实验室李滨同志的合作,特此感谢。

参考文献

- 1 John T. Houghton, *The Global Climate*, published by the press Syndicate of the University of Cambridge, New York, U. S. A., 1984, 56—57.
- 2 盛永宽, 关于气候预测的业务化问题, 国家气象中心技术报告, 1989, 8—11.
- 3 Schlisinger, M. E., Zongci Zhao, *Seasonal Climatic Changes Induced by Doubled CO_2 as Simulated by the OSU Atmospheric GCM/Mixed-Layer Ocean Model*. *J. Climate*, May, 1989, 459—461.
- 4 Ghan, S. J., J. W. Lingas, M. E. Schlisinger, R. L. Mobley W. L. Gates, *A documentation of the OSU two-level atmospheric general circulation model*. Report No. 35 Climatic Research Institute, Oregon State University, Corvallis, 1982, 395pp.
- 5 Pollard, M. L. Batteen, Y. J. Han, *Development of a Simple Upper-Ocean and Sea-Ice Model*. *J. Physical Oceanography*, 1983, 754—755.
- 6 H. N. Robert, *Theory of the Back-Propagation Neural Network*, IJCNN, Washington D. C. 1989, 593—604.
- 7 文新辉, 陈开周, 神经网络广义组合预测模型, 中国科学院, 神经网络学术大会论文集, 1993, 1038—1042.

Short-term Climatic Monthly Rainfall Prediction System Study

Sheng Yongkuan

(National Meteorological Centre, Beijing 100081)

Abstract

The concept of short-term climatic monthly rainfall prediction system is presented. It is discussed that the necessity and possibility of combining climate model and artificial neural network to develop a Climate Prediction Support System. Then, a scheme for system design is given and the system structure is also discussed in detail. Finally, some viewpoints about analysis of prediction results and further improvements are proposed.

Key Words: climate model neural network rainfall prediction system