

日长变化预报中 BP 神经网络 拓扑结构的选择*

王琪洁^{1,2} 廖德春¹ 周永宏¹ 廖新浩¹

(1. 中国科学院上海天文台, 上海 200030; 2. 中南大学测绘与国土信息工程系, 湖南长沙 410083)

提 要

日长变化的预报具有重要的科学意义和实际应用价值。非线性的人工神经网络技术中的反向传播模型(BP 网络)可用于预报日长变化。BP 网络的拓扑结构决定了神经网络解决问题的能力, 针对不同的问题需要采用不同的网络结构。该文分析了神经网络的拓扑结构算法, 选用最小均方误差法确定网络的拓扑结构, 并将此应用于日长变化预报。结果表明, 该方法是可靠和有效的。

主题词: 日长变化—神经网络—BP 网络—最小均方误差法

分类号: P127.1⁺5, O157.5

1 引 言

日长变化的预报具有重要的科学意义和实际应用价值, 尤其是现代空间导航和深空探测等的发展对日长变化的实时监测和预报的需求日益增长。传统的预报日长变化的方法是线性时间序列分析方法, 它是通过对日长变化的时间序列进行分析以确定日长变化的统计特性, 然后据此对其未来的变化趋势作出外推或预测。然而由于地球自转变化复杂的时变特性, 这种分析方法往往难以取得良好的预报效果^[1~3]。随着近 20 多年来人工智能科学的蓬勃发展, 人们发现人工神经网络技术特别适用于分析和预报可能包含非线性作用的动力学过程, 而这是经典线性分析理论和方法所无法企及的, 因此它在各种动力学系统的模式识别、系统控制、时间序列分析和预报等领域中有着广泛的应用^[4,5]。

目前神经网络已发展了 40 多种模型。反向传播模型(BP 网络)主要用于时间序列的预报^[4,5]。Egger(1992)首次将人工神经网络用于地球自转参数(日长和极移)的预报工作^[6], 发现了神经网络在预测准周期和不规则过程中蕴涵有巨大的潜力。Schuh 等人(2002)运用斯图加特神经网络模拟软件对地球自转参数做了短期与中长期预测^[7]。结果表明, 应用神经网络技术进行地球自转参数的预报是较好的选择途径。近年来, 我们应用神经网络技术, 基于空间技术测定的日长变化值和大气角动量(AAM)的重新分析值进行日长变化的预报试验, 大量的模拟试验结果显示了大气角动量函数在日长变化非线性预报中的重要作用; 利用日长变化测定值和大气角动量的实时预报值, 应用神经网络技术能够实现日长变化 1~5 日的实时快速预

收稿日期:2007-07-13; 修回日期:2007-07-25

* 国家自然科学基金(No. 10673025, No. 10633030)项目资助课题。

报^[8-10]。

BP 网络的应用非常广泛。在用它来解决实际问题(例如应用于日长变化预报)时,关键在于 BP 网络拓扑结构(即网络的层数和每层应含有多少个神经元,也称 BP 网络结构)的选择和训练算法的确定,其中网络拓扑结构的选择至关重要。如果网络的结构未选好,即使对训练算法作了很多改进,网络性能也难有大的提高。BP 网络的拓扑结构决定了神经网络解决问题的能力。针对不同的问题需要采用不同的网络结构。本文主要详细介绍日长变化预报中神经网络拓扑结构的选择。

2 BP 网络拓扑结构优化算法

2.1 BP 网络层数的选择

在 BP 算法被提出以后,人们对多层神经网络的能力进行了较多的研究。Lippmann (1987)指出三层网络可以用来处理凸区域上的模式识别问题^[11];Huang 和 Lippmann (1987)用仿真试验表明了三层网络可以用来处理几种很复杂的模式识别问题^[12];Wieland 和 Leighton (1987)给出了用三层网络将空间划分成凹的子空间的例子^[13];Hecht-Nielsen (1987)证明了带有足够多隐单元的 BP 网络能够用来逼近任何从 $[0,1]^n$ 到 $[0,1]^m$ 的映射或函数^[14];进一步地,Hecht - Nielsen (1989)证明了随着隐单元的增加,三层网络所实现的映射可以一致逼近一连续函数或按 L_2 范数逼近平方可积的函数^[15];Blum E K (1991)也证明了 BP 网络能够以任意精度逼近任意连续函数及平方可积函数,且可精确实现任意有限训练样本集。指出只要隐层神经元数选择合适,只有一个隐层的三层 BP 神经网络就可以实现任意精度的输入层到输出层的非线性映射^[16]。从上述结果可以看出,对于任何在闭区间的一个连续数都可以用包含一个隐层的 BP 网络来逼近,一个三层的 BP 网络可以完成任意的 N 维到 M 维的映射,所以目前对 BP 网络的绝大多数研究都是针对三层网络的。

2.2 选用 BP 网络拓扑结构优化算法

三层 BP 神经网络的输入层神经元数和输出层神经元数都由实际应用问题决定,而其隐层节点(神经元)数目则是可调的。到目前为止,对于如何确定隐层神经元数的问题,还没有一个通用的理论公式。因此,研究 BP 网络的拓扑结构问题,实际上即隐层优化问题(隐层神经元数的确定),而研究具有三层的 BP 网络的隐层优化问题,首先必须考虑隐层神经元数多少的利弊。

BP 网络的拓扑结构设计目前采用的方法主要是经验公式法或直接定型法,其主要思路是针对三层神经网络间接或直接地事先给出隐层神经元的数目,使网络训练之后基本可以反映训练数据的变化趋势。这些方法就是针对特定的实际应用问题,设计者依靠经验和现有的理论指导(如 Kolmogorov 定理),设计并采用一个固定的网络结构;也可设计出多个不同的网络结构,然后分别对之进行比较实验,根据实验结果选用其中较优的一种结构。

作者通过大量试验发现,单独应用上述某种方法确定三层 BP 网络隐层神经元数时,都存在着一定的缺陷,很难达到理想的效果^[8]。以上这些方法只能作为参考,网络的拓扑结构最好还是根据研究的问题和实际的数据确定。通常三层 BP 网络由一个输入层、一个隐含层和一个输出层构成。在我们作日长变化预报时,输入层和隐层神经元都用双曲正切作为传递函数,输出层神经元用线性函数作为传递函数;输出层含一个神经元,输入层和隐层的神经元数目通过实际数据的多次计算后择优确定。为此,选用了最小均方误差法来确定网络的拓扑结

构。均方误差的计算公式如下：

$$E_i = \sqrt{\frac{1}{n} \sum_{j=1}^n (p_i^j - o_i^j)^2} \quad (1)$$

式中, o 为网络的期望输出, p 为网络的实际输出, 即预报值, i 为 预报跨度, j 为 预报点序号, n 为 预报点数。

作为一个例子, 图 1 给出日长变化(LODR) 5 日跨度预报时实际网络的训练误差 E_5 随输入层和隐层神经元个数的变化情况。从图 1 可以看出, 本文是对每一种拓扑结构都进行了计算, 输入层神经元个数从 1 算到 15, 隐含层神经元个数从 1 算到 15。当输入层神经元数目为 13, 隐含层神经元数目为 8 时, 网络的训练误差 E_5 最小, 这样就确定为网络的最佳拓扑结构。

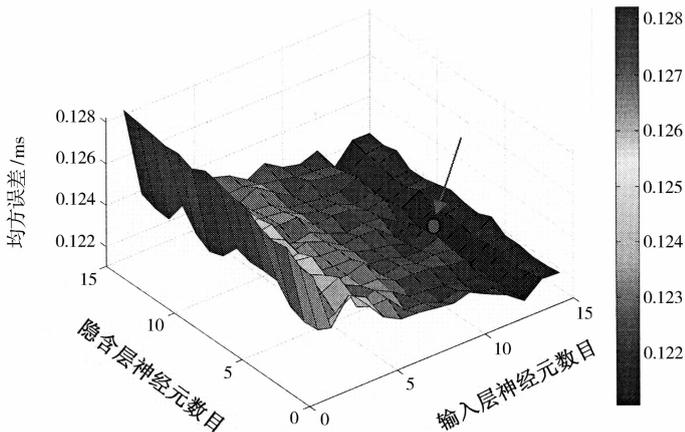


图 1 网络拓扑结构的确定

Fig. 1 Variations of RMS errors of neural network training with respect to the neuron numbers in the input and hidden layers

还需要考虑的是, 由于神经网络的初始化参数是随机赋予的, 为了减小初始化参数的随机波动对拓扑结构和预报结果的影响, 本文对每一种网络模型的训练和预报都重复了多次, 然后对这些预报结果取平均作为最终的预报值。起初的计算中, 本文对每一种拓扑结构的训练和预报都计算 100 次, 期望把网络初始化参数的随机误差的影响大大减小。但是我们发现这样做工作量太大, 耗时太长; 于是我们尝试把每一种拓扑结构的重复训练次数减少到 25 次甚至 5 次, 这样所用的时间大大缩短了。最后, 为了既能减小网络初始化参数的随机误差的影响, 提高预报结果的精度, 又能缩短计算时间, 本文采用的方法是先在较大的拓扑结构范围内, 对每一种网络模型的训练都重复计算 5 次, 得到预报误差 E_{RMS} 随较大范围的网络拓扑结构的分布情况, 再根据这一分布结果确定出一个较小的“网络拓扑结构”的范围, 然后再对这个较小的“拓扑结构”范围的每个模型重复训练 25 次, 找到最佳的拓扑结构。下面是本文在 LODR 与大气的联合预报中, 对跨度为 365 天的预报所作的试验结果, 模型的拓扑结构如图 2 所示, 每个模型的重复训练次数分别为 5 次, 10 次和 25 次。

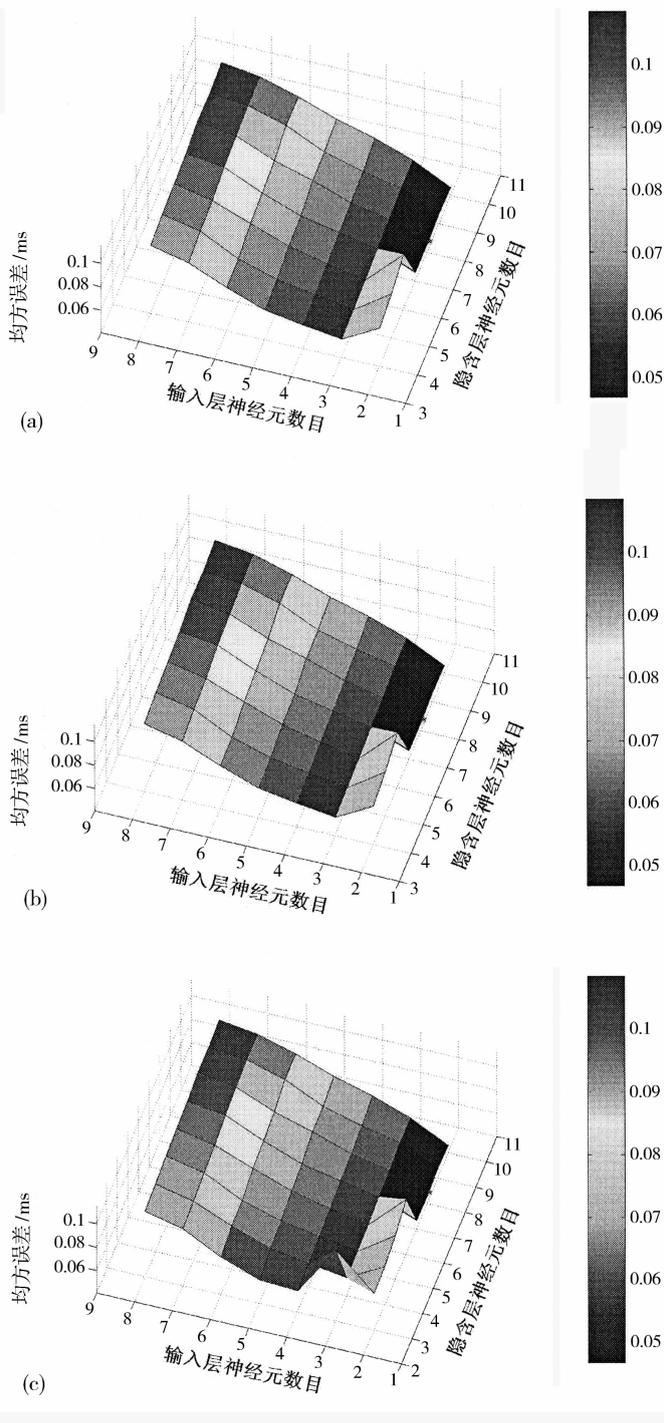


图 2 计算次数对网络拓扑结构的影响

(a) 5 次计算, (b) 10 次计算, (c) 25 次计算

Fig. 2 Influence of computation times on the topology of the network

(a) 5 times, (b) 10 times, (c) 25 times

从图 2 可以看到,对于同样的拓扑结构范围,分别重复训练 5 次,10 次和 25 次,虽然所得到的 E_{RMS} 分布情况不完全一致,但所得到的最佳拓扑模型却完全相同。作者还试验了日长变化预报的多组这样的实例,结果都表明:计算 5 次、10 次和计算 25 次的最佳拓扑范围是一致的。经过这样的试验,在以后的预报中,我们先用 5 次重复训练,大致确定一个较小的拓扑结构的范围,然后再对这个较小的拓扑结构范围做重复训练 25 次的计算,以找到最佳的拓扑结构,得到预报结果,也就是“先 5 次→后 25 次→最佳拓扑”的方法。这样找到的最佳拓扑结构,可以满足一定精度要求,也节省了大量的计算时间。

3 最小均方误差法应用于日长变化预报实例

本节是将最小均方误差法应用于日长变化的预报实例。日长资料序列取自国际地球自转和参考系服务(IERS)的 EOPC04 序列,时间跨度为 1962 - 2005 年,每日一个数据,共约 16000 数据点。日长变化序列包含多种时间尺度的变化,其中由日月引力的潮汐作用所引起的日长周期性变化,可以用理论公式进行精确计算和扣除,扣除了这些潮汐影响后的日长变化,常用 $LODR$ 表示。本文按照 IERS 规范(Convention 2003)对周期为 5d ~ 18.6a 的 62 个固体地球带谐潮汐项对日长变化的影响作了计算和扣除。下文中所指的对日长变化的线性和非线性预报,均不包括这些潮汐成分。日长变化中的另外一些周期性变化,如季节性变化的周年、半年项等可以用线性模型进行拟合和预报。为了降低网络模型的复杂程度和网络训练时的迭代次数,本工作所用的线性模型除了包括周年项和半年项外,还包括周期长度等于 1, 1/2 和 1/3 数据序列长度的项。用最小二乘法对 $LODR$ 序列拟合上述线性模型以后,得到线性模型的参数、模型拟合值序列和残差序列。运用所得到的模型参数可以对 $LODR$ 进行线性模型预报。

以上先扣除线性部分和已知的周期部分可以认为是对资料的预处理。对预处理后得到的残差序列可运用神经网络进行建模和预报,网络的拓扑结构的选择由上述最小均方误差法给出。图 3 是本文采用的神经网络结构示意图。BP 网络的学习过程是由模式正向传播和误差反向传播组成的。当给定网络一组残差序列时, BP 网络将依次对这组残差序列中的每个输入按如下方式进行学习:把残差序列处理后从输入层传到隐含层神经元,经隐含层神经元逐层处理后,产生一个残差预报序列传至输出层,这一过程称为正向传播;如果经正向传播在输出层没有得到所期望的残差预报序列,则转为误差反向传播过程,即把误差信号沿原连接路径返回,并通过逐层修改与上层神经元的连接权值,使误差信号为最小;重复正向传播和反向传播过程,直至得到所期望的残差预报序列为止。图 4 绘出 $LODR$ 10 日跨度的预报序列与残差序列的比较情况,其中虚线表示本工作的残差预报序列,实线是 $LODR$ 残差序列,(b)图是(a)

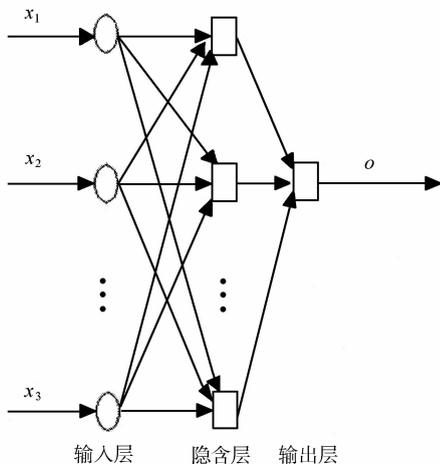


图 3 本文采用的神经网络结构图

Fig. 3 Schematic map of back-propagation neural networks in this study

图 4 绘出 $LODR$ 10 日跨度的预报序列与残差序列的比较情况,其中虚线表示本工作的残差预报序列,实线是 $LODR$ 残差序列,(b)图是(a)

图在 2005 年部分的局部放大。针对 LODR 10 日跨度的预报,当输入层神经元数目为 13,隐含层神经元数目为 6 时,网络的训练误差 E_{RMS} 最小。

从图 4 可以看出,应用最小均方误差法优选神经网络的拓扑结构,预报值和残差序列保持了较好的一致性,二者的相关系数达到 0.93。

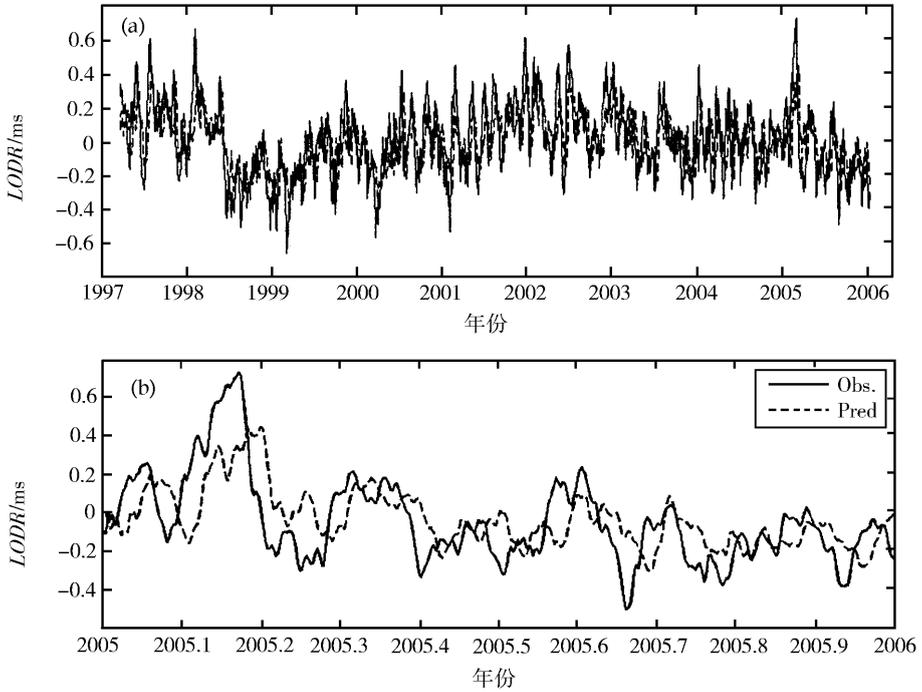


图 4 10 日跨度的 LODR 预报值序列(虚线)与残差序列(实线)的比较

(a)1997-2005 年期间的全部预报,(b)2005 年期间的部分预报

Fig. 4 Comparison of 10 days forward prediction (dashed line) of the neural networks with the residuals (solid) of the linear model

(a) all predictions for 1997-2005, (b) parts of predictions for 2005

4 结 论

本文针对日长变化预报详细讨论了 BP 网络拓扑结构的设计和优化方法。根据日长变化预报结果可以看出,本文优选的神经网络的拓扑结构是比较适当的。针对三层 BP 网络,本文以均方误差为指标,先在较大的拓扑结构范围进行少次数的重复训练,确定出较小的拓扑结构范围,然后再进行多次重复训练,以找出最佳拓扑结构。这种最小均方误差选择 BP 神经网络结构的方法具有普适性。将其应用于日长变化预报的结果表明,该方法是可靠和有效的。

参 考 文 献

- [1] McCarthy D D, Luzum B J. Bull Geod., 1991, 65: 18~22
- [2] Kosek W, McCarthy D D, Luzum B J. Journal of Geodesy. 1998, 72: 189~199
- [3] Malkin E, Skurikhina E. Comm IAA, 1996, 93
- [4] Jacek M Z. Introduction to Artificial Neural Systems. New York: West Pub Company, 1992

- [5] 焦李成. 神经网络系统理论, 西安: 西安电子科技大学出版社, 1996
- [6] Egger D. In: Allgemeine Vermessungsnachrichten (AVN), 1992, 11/12, S: 517 ~ 524
- [7] Schuh H, Ulrich M, Egger D *et al.* Journal of Geodesy, 2002, 76: 247 ~ 258
- [8] 王琪洁. 基于神经网络技术的地球自转变化预报, 上海: 中国科学院上海天文台, 博士学位论文, 2007
- [9] WANG Qi-jie, LIAO De-chun, ZHOU Yong-hong. Chinese Science Bulletin. 2007, 52(14), (in press)
- [10] 王琪洁, 廖德春, 周永宏等. 天文学报, 2008, 49(已录用, 待出版)
- [11] Lippmann R P. IEEE ASSP Magazine. 1987: 4 ~ 22
- [12] Huang W, Lippmann R P. proceedings of the IEEE First International Conference on Neural Networks. 1987, 3: 741 ~ 746
- [13] Wieland A, Leighton R. proceedings of the IEEE First International Conference on Neural Networks. 1987, 3: 385 ~ 392
- [14] Hecht-Nielson R. proceedings of the IEEE first International Conference on Neural Networks. 1987, 3: 11 ~ 14
- [15] Hecht-Nielson R. International Joint Conference on Neural Network. Washington D C: Reading pub, 1989
- [16] Blum E K, Li L K. Neural Networks. 1991, 4(4): 511 ~ 515

DETERMINATION OF THE TOPOLOGY OF THE NEURAL NETWORKS IN THE PREDICTION OF LOD

WANG Qi-jie^{1,2} LIAO De-chun¹ ZHOU Yong-hong¹ LIAO Xin-hao¹

(1. Shanghai Astronomical Observatory, Chinese Academy of Sciences, Shanghai 200030;

2. Department of Geomatics, Central South University, Changsha 410083)

Abstract

The prediction of the Length of Day (LOD) is of great scientific and practical importance. This study employs the non-linear artificial neural networks (BP network, i. e. Back-Propagation network) to predict the LOD change. The predicting ability of the BP network is determined by the topology of the network. Different topologies are needed to solve different problems. This study analyzes the algorithms of topology determinations, and chooses the least Root Mean Squared Error (RMSE) as a criterion to determine the topology of the network. Finally, this paper applied the developed method to predict the LOD change. The results show that this method is reliable and effective.

Key words Length of Day — Neural Networks — BP network — the least Root Mean Squared Error (RMSE) method