

基于改进组合神经网络的水资源预测研究

王 坚

(中央财经大学信息学院 北京 100081)

摘 要 我国作为水资源大国,在日益加速的城镇化进程中正面临人口膨胀、环境污染、水质变差等一系列重大的挑战,而科学合理地对水资源需求进行预测成为保护环境、保持可持续发展的关键任务。首先将神经网络应用于水资源需求预测问题背景并比较其算法,同时引入模糊反馈法来改进熵值法以确定组合模型的加权系数,建立组合神经网络预测模型。该算法不仅可以自动根据历史数据推演今后水资源需求的变化趋势,还引入反馈和演化机制,用户可以调整求解精度以控制算法的收敛速度。实验表明,在数据精度不高以及水文数据不全等不利应用背景中,提出的基于组合模型的神经网络在水资源预测中具有较好的性能。

关键词 水资源需求预测,神经网络,组合模型,模糊反馈

中图法分类号 TN912.33 文献标识码 A

Research on Prediction of Water Resource Based on Improved Combination Neural Network

WANG Jian

(School of Information, Central University of Finance and Economics, Beijing 100081, China)

Abstract China is a big country of water resource and in the accelerating process of urbanization, it is facing a series of major challenges such as urban population growth and water pollution, etc, so scientific and rational forecasting to water resource demand becomes a key task to protect environment and maintain sustainable development. This paper summarized various neural network algorithms in the context of water resource demand forecasting, and introduced fuzzy feedback method to improve entropy method to determine the weighting factor of combination forecasting model, to establish neural network forecasting model. The algorithm can not only automatically deduce future change trends of water resources based on historical data, but also introduce feedback and evolution mechanism, so that the users can adjust the solution accuracy to control convergence speed of algorithm. The experiment shows that, the neural network based on combination model proposed in this paper has better performance in application background when data accuracy is not high and hydrological data is incomplete.

Keywords Prediction for water resource requirement, Neural network, Combination model, Fuzzy feedback

1 概述

我国作为人口大国和人均资源相对匮乏的发展中国家,在经济建设和改革开放的过程中一直面临着如何合理分配与使用资源的巨大难题,其中水资源的分配使用又是最为突出和紧迫的。随着社会主义市场经济格局的形成与发展,水资源问题(主要包括过度开发与浪费、水体污染)将长期困扰我国大部分经济发达地区,甚至显示出了向农村和原来水资源相对充裕地区蔓延的趋势。因此,我国一直把水资源的保护与合理使用当作关乎国计民生和经济社会可持续发展的战略性问题,并建设了一大批水利工程调整全国性的水资源分配使用,并对沿江沿河水资源的使用做出了最为严格的限制。在这样的背景中^[1,2],对水资源的变化和使用需求进行准确预测显得十分重要。

比较成熟的数学模型都已被广泛应用到对水资源的预测中,并得出了比较好的预测结果。但是这些数学模型绝大多数都是静态的,它只能将一些水利水文以及领域相关的先验

知识形式化以后作为模型的计算规则包含到预测系统中来,预测系统的灵活性和实用性受到限制。因此,一些具有自学习、自适应功能的方法开始在水资源预测中使用,其中包括了多种人工神经网络的应用等,但是这些方法对数据的精度有一定的要求,而水资源数据中难免有一些噪声甚至错误,这将使预测结果受到不同程度的干扰^[3]。神经网络的预测模式被应用到这一领域,而其有效组合即通过带有模糊反馈机制的熵值法算法确立权重是一种新的尝试。实验证明,本方法在抵抗由于数据噪声等不确定因素引起的预测波动方面有一定作用,预测结果的准确性得到了提高。

2 神经网络及其组合预测模型

2.1 LVQ神经网络

LVQ神经网络^[4]通过统计具有自适应的概率分布特征进行分类和识别。该方法将与领域相关的先验知识中的不变部分作为规则包含到分类器中,而将一些变化的规则以及随机因素作为反馈,在有监督的自学习环境中,通过若干次交互

本文受中央财经大学重点学科建设项目,北京高等学校青年英才计划项目(YETP0988)资助。

王 坚(1975—),男,博士,讲师,主要研究方向为模式识别、数据挖掘, E-mail: wanderingful@126.com。

后化为规则包含到系统中。由于学习矢量化算法的思想中融入了演化算法中物竞天择的思想,试图将包含过多噪声的数据被正常的数据所淘汰掉,这将大大加快预测系统的收敛速度。

2.2 ELMAN 神经网络

ELMAN 神经网络^[5]作为一种典型的反馈神经网络,一般是由 4 层组织架构组成:输入层、隐含层、承接层和输出层。其中,输入层、隐含层、输出层的连接类似于前馈式网络,而承接层的作用是作为一步延时算子,起到记忆作用,从而帮助系统拥有自适应时变特性,并能够直接动态反映系统特性。

在实际源预测过程中,基于输入 n 维向量 x ,其反馈状态向量 y 通过 $y(k)=y(k-1)$ 实现由隐含层到输入层的承接反馈,从而实现历史数据的有效建模,尽可能地反映待预测目标的运作规律。

2.3 SOM 神经网络

SOM^[6]神经网络通过构建全连接神经元阵列来实现无监督的自学习。依靠不同神经元对输入的响应差异及其之间的关联性,有效地近似输入向量的拓拓结构。按照获胜单元优先更新的策略,其临近神经元也会按照某种函数曲线实现同步。依据排序规则,实现输入层到竞争层之间的非线性映射。

在实际预测过程中,SOM 网络有效模拟和抽取输入信号的模式特征,在由神经元构成的处理单元阵列上描述出某一类输入权值分布。这种方法可以找到数据之间的相似度,提升预测准确度。

水资源预测需要综合衡量水文数据,仅依赖一种神经网络在预测效果上往往难以达到较为满意的结果。因此,本文应用组合预测统一框架,实现预测能力的提升。

2.4 组合模型预测及改进权重计算

组合模型(Combination Model,CM)^[7,8]预测通过对加权系数进行综合考量,实现基于多种方法对统一预测对象的评价。每种评价方式,可以依靠权重系统的调节,发挥它在整体结果中的重要性。在一定的约束准则条件下,目标函数即预测的最佳结果能更好地实现拟合,有较高的预测精度。

设对某一预测对象采用 m 种预测方法,观测 n 个时刻,其中在 $t(t=1,2,\dots,n)$ 时刻得到 m 种模型的预测值 $f_{it}(i=1,2,\dots,m;t=1,2,\dots,n)$,并选取一组适当的权重系数 $W(w_1,w_2,\dots,w_m),w_i(i=1,2,\dots,m)$ 为第 i 种预测方法子组合预测模型汇总的权系数,且满足 $\sum_i w_i=1$,则 t 时刻组合预测模型的预测值可以表示为 $f_{it}=\sum_i w_i f_{it}$ 。

利用熵值法确定组合预测加权系数,具体步骤可以参考文献^[9]。通过其预测过程可以看出,系统的学习没有有效的反馈,即单纯的预测结果不能构成有监督的学习过程,因此,需要对一些样本的分类预测不断地进行动态的“激励-抑制”,对正确的分类预测结果进行激励,强化学习效果,增加变量权重的置信度,反之进行抑制,降低变量权重的置信度。

对样本空间的每个权重值设置一个距离影响因子 $n(t)$,使得距离较远的权重值也会相互影响。主要步骤不发生显著变化,则其激励函数为:

$$w_i(t+1)=w_i(t)+\sum n(t) \star (x_i(t)-w_i(t))$$

抑制函数变为:

$$w_i(t+1)=w_i(t)-\sum n(t) \star (x_i(t)-w_i(t))$$

这里, $\sum$ 表示在定义域以内所有距离不大于规定阈值的权重点都按照一定的比例影响下次训练过程。为了计算方便,一般将影响因子反比例于其距离值。此外,由于各种因素变量的取值范围差异极大,其权重值需要在计算精度范围内做规划化处理,这里采用了高斯归一化方法,即认为权重值的变化符合高斯正态分布,其变化范围在以均值为中心的 3 倍的方差内。对于落到了范围之外的权重值样本只做简单的截断处理,就认为它们的取值在邻域边界上。

训练好组合模型的各项变量的权重值,就可以开展预测工作,但是与传统的神经网络不同的是,这些权重值以及它们的变化范围依然会随着预测结果与真实结果的差距进行调整,特别是在两者有较大差异的时候,可以有效调节其相对位置,帮助实现正确结果走向。

3 模型算例分析

水资源预测是一个规模巨大的问题。为了使得预测模型应用背景更加明晰,主要以重庆嘉陵江流域北碚水文站 1939—2004 年的水文流量数据^[10]为例说进行预测。该水文站位于重庆合川与四川平原的交汇处,地形较为复杂,降水充沛,每年的 6—9 月为雨季,流量变化较大。20 世纪的最大洪水水位超过了 210m,而该水文站的保证水位为 202m。表 1 是对 1939—2004 年间的洪水波峰值统计。

表 1 1939—2004 年每 5 年的洪峰值

Year	flux(m ³ /s)
1939	18500
1944	9700
1949	15000
1954	15500
1959	14700
1964	17000
1969	10000
1974	21000
1979	20500
1984	20300
1989	21000
1994	7200
1999	18600
2004	21400

根据水利部的预测规范^[1],以距平值 D 代替方差作为洪水等级评测的标注,并根据重庆嘉陵江流域的水文情况对洪水进行级别划分,如表 2 所列。

表 2 洪水级别的划分

Level of flood	Average Distance value S(%)	Logical threshold
1	Less than -40	Low
2	[-40,-25)	Relatively low
3	[-25,-5)	Medium low
4	[-5,15)	Medium
5	[15,35)	Relatively high
6	[35,50)	High
7	More than 50	Very high

(下转第 532 页)

4 测试

实现上述设计后,进行实际的外场测试,在测试过程中为重点模拟当公网瘫痪时,基于 RDS 技术的应急信息接收,将手机终端设置为飞行模式,详细测试条件如表 4 所列。

表 4 测试条件表

FM 广播频点	98MHz
调频编码发射机输出功率	12W
发射天线型号	GP-2 伞型 FM 天线、阻抗 50Ω
发射天线增益	3dbm
发射天线辐射方向	水平全向
位置因素	距地面高度 20m,周围无明显遮挡覆盖物
接收 FPC 天线中心频点	90MHz~105MHz
测试手机机型	宝捷讯 T01
机型方案版本	联发科 MT 6582

终端实物图如图 8 所示。



图 8 应急信息获取关键代码图

基于上述真实外场测试条件,内置 FPC 接收天线的智能手机终端的有效语音接收距离为 3~4km,RDS 应急文本信息的接收范围为 1~1.5km,即可实现该覆盖范围内的应急语音与文本信息的覆盖。在该有效接收范围内,RDS 应急文本信息的接收时延如表 5 所列。

表 5 RDS 文本接收时延迟

序号	字节数	是否收到	接收时延(s)
1	10 字以内	是	5
2	20 字以内	是	5~10
3	30 字以内	是	5~15
4	40 字以内	是	15~20

结束语 本文阐述了一种应急通信终端。通过对普通智能手机终端的软件 APP 程序与硬件内置天线进行设计,实现了复杂灾难场景下特别是在公网瘫痪时公众对于应急信息的接收,为我国应急通信技术发展提供了一种有效的解决方案。

参 考 文 献

- [1] 刘春江,马艳,李玉环,等.基于调频副载波的应急广播消息传输协议探讨[J].电视技术,2014,38(20):37-40
- [2] 李一兵,黄春子.应急通信发展现状和技术手段探析[J].电子游戏软件,2014(13):81-81.
- [3] 陈泽虎,倪林,严杰.一种调频广播数据接收显示系统的设计[J].电视技术,2014,38(15):114-117
- [4] 余尚林,宗瑞朝.副载波技术在应急广播系统的应用[J].电视技术,2014,38(10):42-45
- [5] 代传堤,吴多龙,吴艳杰,等.一种紧凑型多频段手机内置环形天线的设计[J].广东工业大学学报,2014(4):100-103
- [6] 陈去疾,李敬华,郭华磊.JNI 在 Android 硬件开发中的应用[J].电信快报,2014(1):27-29

(上接第 517 页)

以 1939—1979 年的数据作为 LVQ 模型的训练样本集合,对 1979 年以后的数据进行预测的结果如表 3 所列。

表 3 预测结果对比

Year	实际流量	LVQ 预测值/ 误差(%)	Elman 预测值/ 误差(%)	SOM 预测值/ 误差(%)	CM 预测值/ 误差(%)
1984	20300	17700/12.8	17700/12.8	18500/8.9	18900/6.9
1989	21000	18400/12.4	18100/13.8	19100/9.0	20800/9.5
1994	7200	6500/9.7	6600/8.3	6510/9.6	6850/4.8
1999	18600	16900/9.1	17800/4.3	17600/5.4	18200/2.2
2004	21400	19300/9.8	19980/6.6	20400/4.7	20800/2.8
平均误差%		11	9.8	8.2	3.7

可以看出,组合模型预测过程中的误差均值为 3.7%,最小误差为 2.2%,可见在经过训练以后系统的误差不仅得到了控制,还有所缩小,预测结果基本满意。与神经网络单独预测结果相比,性能得到提升。

结束语 在传统的数学模型预测水资源的基础上,发现了多种神经网络模型在预测中的不足之处,提出了基于模糊反馈的神经网络组织预测模型。该模型基本不需要水资源问题领域的专业知识,可以从大量的有噪声、不准确的数据中自动生成模型,权重变量得到控制,同时根据预测结果的准确度进行必要的反馈,以提高预测模型的精度。由于结合了无监督学习以及有监督反馈的特点,既使人为的干预最小化,又使得在关键误差出现时系统可以根据操作对变量权重进行修

改,从而提高了系统的自适应性。实验结果表明,较以往智能模型,结合模糊反馈的神经网络组合预测模型对洪峰最大峰值的预测精度有所提高。

参 考 文 献

- [1] 中华人民共和国水利部.水文情报预报规范 SL2502-2000[S].北京:中国水利水电出版社,2000
- [2] 金兴平,程海云,等.长江流域水资源预测技术[J].人民长江,2005,36(12):18-20
- [3] 冯利华.神经网络在水资源预测中的运用[J].农业系统科学与综合研究,2003,19(1):35-38
- [4] 罗志增,熊静,等.一种基于 WPT 和 LVQ 神经网络的手部动作识别方法[J].模式识别与人工智能,2010,23(5):695-700
- [5] 杨钟亮,陈育苗.基于 GGA-Elman 网络的头部体态语言 sEMG 识别方法[J].智能系统学报,2014(4):385-391
- [6] 王平勋.基于自组织竞争神经网络的抽油井故障诊断系统[J].电子设计工程,2013,21(11):112-115
- [7] 王莎莎,陈安,苏静,等.组合预测模型在中国 GDP 预测中的应用[J].山东大学学报(理学版),2009(2):56-59
- [8] 易丹辉.统计预测方法与应用[M].北京:钟鼓统计出版社,2001
- [9] 张星.基于熵权的福建粮食产量组合预测模型[J].中国农业气象,2008,29(2):194-196
- [10] 王渺林.嘉陵江径流长期变化分析及预报[J].重庆交通学院学报(自然科学版),2006,25(6):135-138