

基于主成分回归与灰色神经网络模型的水资源承载力需水量预测

章恒全^a, 何薇^b

(河海大学 a. 商学院 决策与规划研究所; b. 商学院, 南京 211100)

摘要: 我国的水资源利用问题日趋严峻。作为影响水资源承载力的重要因素, 社会经济活动对水资源的影响尤为明显。通过分析影响水资源承载力的社会经济驱动要素, 建立主成分回归模型, 分析得出影响江苏省水资源承载力变化的三个驱动力以及驱动力影响度, 利用三个驱动力中的6个重要驱动因子, 建立灰色神经网络预测模型, 预测出江苏省2012-2013年的年需水量。结果表明: 预测模型精度较高, 最后结合江苏省发展现状提出相关的政策性建议。

关键词: 水资源承载力; 主成分回归模型; 灰色神经网络模型; 需水量预测

中图分类号: TV213.4

文献标识码: A

文章编号: 1672-643X(2014)01-0103-06

Forecast of water requirement of water resources carrying capacity based on regression of principal components and grey neural network model

ZHANG Hengquan^a, HE Wei^b

(a. Institute of Planning & Decision - Making, Business School; b. Business School, Hohai University, Nanjing 211100, China)

Abstract: Our country is facing serious water problems. As the important factor that affect water resources carrying capacity, the influence of social and economic activities on water resources is obvious. By analyzing the key element of socio-economic driver that affects water resources carrying capacity, establishing the principal component regression model, the driving factors which influence the carrying capacity of water resources in Jiangsu were classified into three kinds, the influence degree of the driving factors was gained. Combined with grey neural network theory and six important factors which is coming from the three main factors, the grey neural network model was established to predict the total water uses in Jiangsu Province from 2012 to 2013. The prediction accuracy of the model is higher. At last, the paper gave relative policy recommendations by combining with the actual situation of Jiangsu Province.

Key words: carrying capacity of water resources; principal component regression model; grey neural network model; forecast of water requirement

水是人类维系生命和发展的至关重要的自然资源。在人类发展的历史长河中, 社会经济活动都是围绕着水资源丰富的区域进行的。作为社会经济发展的支撑性资源, 水资源出现了一系列短缺和污染等危机性现象。中国是世界上最大的发展中国家, 水资源总量虽然不少, 但是人均占有量仅有 2 300 m³, 属水资源严重短缺国家。随着我国城市化进程的加快, 社会经济活动超越了自然生态系统对水资源的原始需求量, 对水资源的循环机制产生了多方面的影响, 因此, 进行水资源承载力的社会经济驱动

力分析以及需水量预测, 来优化水资源配置和加强水资源承载力, 对我国社会经济的可持续发展是非常必要的。

关于需水量的预测研究, 国内学者通过一些理论研究而对其预测进行了部分的改良。需水量的预测方法多种多样, 包括时间序列法、回归分析法、灰色模型和神经网络等。杨鹏, 陈金锥采用改进的 GM(1, 1) 模型进行城市年需水量的预测^[1]; 胡彩虹, 吴泽宁等人建立基于主成分分析的需水量预测模型, 探讨需水定额与经济社会的关系^[2]; 冯坤, 张

永丽等人通过主成分因子分析需水量的影响因素,在建立BP神经网络预测模型进行预测^[3];李莉会,陆宝宏等人改进BP神经网络模型中学习率的求解方法,利用改进模型预测郑州市经济社会需水量^[4]。

基于以上学者的研究,本文以江苏省为研究区域,结合主成分分析理论、灰色系统理论与BP神经网络三大理论,首先通过建立主成分回归模型分析影响水资源承载力的社会经济驱动因素和驱动力影响度,其次利用驱动因素建立灰色神经网络的组合预测模型,进行江苏省年需水量预测。

1 水资源承载力驱动因素提取

1.1 承载力社会经济影响要素分析

联合国教科文组织(UNESCO)和世界气象组织(WMO)共同制定的《水资源评价活动—国家评价手册》中,将水资源定义为:“可以利用有可能被利用的水源,具有足够数量和可用的质量,并能在某一地点为满足某种用途而可被利用。^[5]”水资源承载力是在“社会经济—水资源—生态负荷大系统”有机运转下,达到“生态系统良性循环”目标时的“最大社会经济发展规模”^[6]。在“水资源承载力”的概念定义如图1中可以看出,作为其概念的主体,水资源承载着人类的生存,也是人类生存的社会经济系统和环境系统的资源依托,水资源系统与社会经济系统、生态系统可谓是相互依赖、相互影响的循环体系,在地球这个生态系统中,人类的社会经济活动影响度日趋成为水资源系统的至关重要组成部分,在研究水资源承载力大小时,必须要了解水资源与社会经济发展的相互关系,社会经济所带来的工业革命,人口骤增等问题对水资源承载力的影响,因此,要通过对水资源承载力驱动因素的分析,来寻求水资源利用与保护的途径。

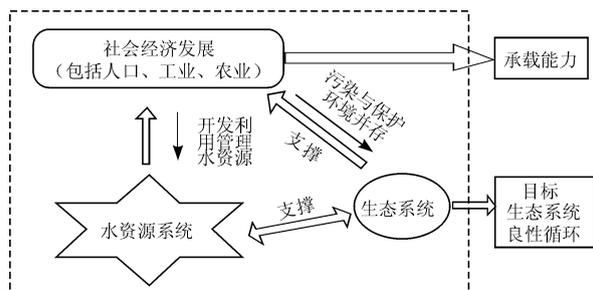


图1 水资源承载力概念模型

通过对水资源承载力以及水资源循环系统内涵的分析,可以将影响水资源承载力的社会经济

驱动要素可归纳为3个方面的量化指标,分别是水资源系统本身指标、社会发展指标、经济状况指标。首先,水资源系统是水资源承载能力的主功能系统,其中水资源承载能力的主要衡量指标量则是水资源可利用总量,水资源生态量是水资源承载能力大小的先决条件。因此,选取地表水资源量、地下水资源量、年降水量作为代表水资源系统的水资源承载能力驱动要素。其次,在社会发展中,人口的发展是社会的根本,人口的发展变化往往推动社会中城市规模、产业经济等方面的变化,近年来,人口的增长,城市的扩张带来的是水资源的需求剧增,人口增长率和人均日生活用水量作为人口变化指标及人口耗水指标是水资源承载能力驱动因素的重要组成部分,而城市化率和污水处理率、水资源综合生产能力则是水资源承载力驱动因素中城市规模发展和水资源处理技术应用能力的代表性指标。最后,水资源是人类经济活动的基础性支撑条件,经济活动的发展又对水资源产生一定的压力,因此,分别选取人均GDP指数、固定资产投资额、地区生产总值、二三产业结构比作为经济发展和产业结构的代表性指标,耕地灌溉率和生产用水量作为经济发展活动对水资源需求大小的代表性指标,以上指标选取见图2。

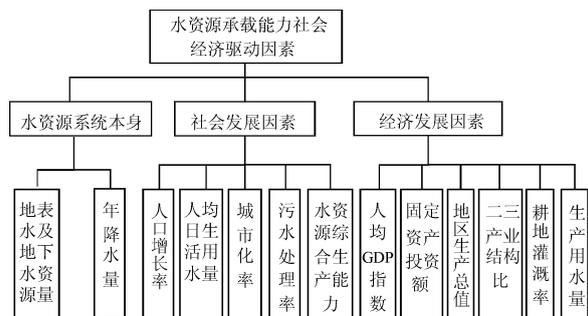


图2 水资源承载力社会经济驱动因素指标体系图

1.2 主成分回归模型建立

1.2.1 驱动主成分提取 主成分分析是利用降维的思想,在损失很少信息的前提下把多个指标转化为几个综合指标的多元统计方法,从而简化问题的复杂性并抓住问题的主要矛盾^[7]。水资源承载力社会经济驱动因素分析是利用主成分分析法,将影响其承载力的驱动因素进行降维分析,力求将众多的驱动因素简化为几个相互独立的公共因子,这些公共因子可以将原驱动因素的主要信息充分反映出来。本文选取的指标数据根据《江苏统计年鉴》1990—2011年共22年13项指标的数据,其中涉及产值3项指标固定资产投资额、人均GDP指数、地区生产总值数据均以

1990 年为基期年进行了换算。

本文水资源承载能力社会经济驱动机制分析选取共 13 个指标,假设分别用 x_1, x_2, \dots, x_{13} 表示,通过标准化得到的变量矩阵:

$$T = \begin{bmatrix} t_{11} & t_{12} & \dots & t_{1p} \\ t_{21} & t_{22} & \dots & t_{2p} \\ \vdots & \vdots & & \vdots \\ t_{q1} & t_{q2} & \dots & t_{qp} \end{bmatrix}$$

$$(p = 1, 2, \dots, 13, q = 1, 2, \dots, 22)。$$

将标准化的变量矩阵建立主成分模型:

$$F_i = a_{1,i}t_1 + a_{2,i}t_2 + \dots + a_{13,i}t_{13} \quad (i = 1, 2, \dots, 13) \quad (1)$$

式中: $F_1, F_2, \dots, F_m (m = 1, 2, \dots, i)$ 为主成分, $a_{1,i}, a_{2,i}, \dots, a_{13,i}$ 为原始指标数据的相关系数矩阵的特征值所对应的特征向量,且满足 $a_{1,i}^2 + a_{2,i}^2 + \dots + a_{13,i}^2 = 1$ 。通过变量矩阵 T 的协方差矩阵 S 推导相关矩阵 R :

$$S = \frac{1}{q-1} \sum_{k=1}^q (t_{ki} - \bar{t}_i)(t_{ki} - \bar{t}_i)' \\ = \frac{1}{21} \sum_{k=1}^{22} (t_{ki} - \bar{t}_i)(t_{ki} - \bar{t}_i)', \bar{t}_i = \frac{1}{q} \sum_{k=1}^q t_{ki}, \quad (i = 1, 2, \dots, 13) \quad (2)$$

$$R = (r_{ij})_{13 \times 13} = (S_{ij} / \sqrt{S_{ij}S_{ij}})_{13 \times 13}, \quad (r_{ij} = S_{ij} / \sqrt{S_{ij}S_{ij}}) \quad (3)$$

求出相关矩阵 R 的特征值,根据特征方程 $|R - \lambda I| = 0$,求得 R 的特征根 $0 \leq \lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_{n-1} \leq \lambda_n$;计算贡献率和累计贡献率,并依据累计贡献率和特征根的大小来选取主成分及其个数,并依此建立主成分方程,其中,第 m 项贡献率为^[8]:

$$a_m = \frac{\lambda_m}{\sum_{i=1}^n \lambda_i} \quad (m = 1, 2, \dots, n)$$

采用 SPSS 软件对数据进行统计分析,该变量指标矩阵的相关矩阵 3 个特征值 $\lambda_1, \lambda_2, \lambda_3$ 合计占比约为 84.62%,见表 1,因此,3 个主成分已经可以在一定程度上代表原始变量的信息,主成分的选取要满足的条件是:其特征值要大于 1,其累计贡献率大于 80.00%。根据分析结果可以看出,前 3 个主成分满足条件。因此 F_1, F_2, F_3 可以用来对水资源承载能力社会经济驱动力进行解释。

通过对主成分荷载矩阵的分析,在第一个主成分荷载较大的变量有二、三产业结构比、城市化率、地区生产总值等指标,其荷载均超过 0.95,与本文进行指标选择时的产业结构因素以及城市发展规模

因素相对应,因此第一主成分可以称之为“社会经济规模发展主成分”;在第二个主成分荷载较大的变量有人均日生活用水量、耕地灌溉率、生产用水量等指标,其荷载均超过 0.65,对应于指标与人类社会、经济活动用水量相关,因此第二主成分定义为“人口与产业发展净耗主成分”;同时,在第三个主成分荷载较大的变量有年降水量、地表及地下水资源量指标等,因此第三个主成分可以概括为“水资源生态量主成分”。

表 1 主成分指标权值及贡献率

成分	初始特征值			提取平方和载入		
	合计	变量贡献	累计贡献	合计	变量贡献	累计贡献
1	6.891	53.011	53.011	6.891	53.011	53.011
2	2.427	18.668	71.679	2.427	18.668	71.679
3	1.683	12.945	84.624	1.683	12.945	84.624
...

1.2.2 驱动主成分影响度分析 主成分回归的原理是用主成分分析提取的主成分与因变量回归建模,主成分不但相互之间相关性很小,并且反映相关性指标的综合信息能力较强,所以将主成分作为新的自变量进行回归分析使得回归方程及参数估计更加适合^[9]。本文水资源承载力社会经济驱动机制主成分回归分析具体步骤为:首先,将年用水量 Y 的标准化数据与主成分的得分建立新的标准化数据矩阵;其次,采用普通最小二乘法,做 3 个主成分 F_1, F_2, F_3 与因变量 Y 的多元线性回归,得到主成分回归模型。

$$\hat{Y} = \alpha + \beta F_1 + \gamma F_2 + \omega F_3 + \varepsilon \quad (4)$$

通过主成分分析的线性变换,3 个主成分之间已经不存在相关性,因此,将上述分析出的 3 大主成分作为自变量,房地产开发投资额作为因变量,建立标准化数据矩阵(表)进行回归分析,分析探寻社会经济等驱动因素对我国江苏省水资源承载力变化的影响方向和程度。现在应用 SPSS16.0 对主成分回归模型进行最小二乘法估计,得到式(5)。

$$\hat{Y} = 1.099 \times 10^{-15} + 0.854F_1 + 0.323F_2 - 0.238F_3 \quad (5)$$

系数估计值标准误差 0.078 0.078 0.078
回归系数 T 统计量 (10.919) (4.124) (-3.042)

方程模拟回归结果显示,3 大主成分回归系数的修正样本可决系数为 0.872, F 检验统计量为 48.491, DW 统计量为 1.863,因此其 t 检验均合格,修正样本

可决系数检验和 F 统计量检验来看,回归方程整体上拟合良好,回归系数的影响作用相对显著,主成分 F_1 、 F_2 、 F_3 的回归系数分别为 0.854、0.323、-0.238。因此,3 个主成分对因变量 Y 的影响可以解释为, F_1 、 F_2 、 F_3 每增加一个单位的投入,对因变量 Y 的驱动作用分别为增加 0.854、增加 0.323、减少 0.238 单位。通过主成分回归分析,江苏省水资源承载力的社会经济驱动力可以分为 3 大驱动主成分,分别为社会经济规模发展主成分、人口与产业发展净耗主成分、水资源生态量主成分,其中,社会经济规模发展主成分与产业发展净耗主成分对水资源的承载能力产生了驱动压力,水资源生态量主成分中的指标则反映了水资源总量的变动,其年降水量的增长对水资源承载压力进行了缓解,其驱动力的增长加强了水资源面对人类社会经济活动的负荷能力。

2 基于灰色预测模型的江苏省需水量预测

2.1 灰色预测模型 GM(1, N) 模型

灰色系统理论是以“部分信息已知、部分信息未知”的“小样本”、“贫信息”不确定性系统为研究对象,主要通过对“部分”已知信息的生成、开发,提取有价值的信息,实现对系统运行行为、演化规律的正确描述和有效监控^[10]。灰色 GM(1, N) 模型是灰色系统理论的主要内容之一,该模型是利用离散的数列建立近似连续的微分方程,它反映的是 $N-1$ 个变量对某一变量的一阶导数的影响^[11]。

设 $x^{(0)}$ 为原始序列, $x^{(0)} = \{x_{(1)}^{(0)}, x_{(2)}^{(0)}, \dots, x_{(n)}^{(0)}\}$, $\forall x_{(n)}^{(0)} \in R^+, n \in N$; 记 $x^{(1)}$ 为生成序列 $x^{(1)} = \{x_{(1)}^{(1)}, x_{(2)}^{(1)}, \dots, x_{(n)}^{(1)}\}$, 其中 $x_{(k)}^{(1)} = \sum_{i=1}^k x_{(i)}^{(0)}, k = 1, 2, \dots, n$; 设 $X_1^{(0)} = (x_1^{(0)}(1), x_1^{(0)}(2), \dots, x_1^{(0)}(n))$ 为系统特征数据序列, 而:

$$X_2^{(0)} = (x_2^{(0)}(1), x_2^{(0)}(2), \dots, x_2^{(0)}(n))$$

$$X_3^{(0)} = (x_3^{(0)}(1), x_3^{(0)}(2), \dots, x_3^{(0)}(n))$$

.....

$$X_N^{(0)} = (x_N^{(0)}(1), x_N^{(0)}(2), \dots, x_N^{(0)}(n))$$

为相关因素序列, $X_i^{(1)}$ 为 $X_i^{(0)}$ 的 1-AGO 序列, $i = 1, 2, \dots, N$, $Z_1^{(1)}$ 为 $X_1^{(1)}$ 的紧邻均值生成序列, 则称:

$$x_1^{(0)}(k) + aZ_1^{(1)} = \sum_{i=2}^N b_i x_i^{(1)}(k) \quad (6)$$

为 GM(1, N) 模型。其中: $Z_1^{(1)} = \frac{1}{2}[X_1^{(1)}(k) +$

$X_1^{(0)}(k-1)], k = 2, 3, \dots, n$ 。灰色预测模型的建立如下:

① 对 $X_i^{(1)}$ 建立白化形式微分方程:

$$\frac{dx_1^{(1)}}{dt} + ax_1^{(1)} = b_2 x_2^{(1)} + b_3 x_3^{(1)} + \dots + b_N x_N^{(1)} \quad (7)$$

记方程的参数列为 $\hat{a} = [a, b_1, b_2, \dots, b_N]^T$, 在模型中, $-a$ 成为系统发展系数, $b_i x_i^{(1)}(k)$ 成为驱动向, b_i 成为驱动系数, 成为参数列。

② 利用最小二乘法求得参数列。设

$$B = \begin{bmatrix} -Z_1^{(1)}(2) & x_2^{(1)}(2) & \dots & x_N^{(1)}(2) \\ -Z_1^{(1)}(3) & x_2^{(1)}(3) & \dots & x_N^{(1)}(3) \\ \vdots & \vdots & & \vdots \\ -Z_1^{(1)}(n) & x_2^{(1)}(n) & \dots & x_N^{(1)}(n) \end{bmatrix} \quad (8)$$

$$Y = [x_1^{(0)}(2), x_1^{(0)}(3), \dots, x_1^{(0)}(n)]^T \quad (9)$$

$$\begin{aligned} \text{参数列判别: } \hat{a} &= [a, b_1, b_2, \dots, b_N]^T \\ &= (B^T B)^{-1} B^T Y \end{aligned}$$

③ 通过以上分析,得到 GM(1, N) 的近似时间响应方程为:

$$\hat{x}_1^{(1)}(k+1) = \left[x_1^{(1)}(0) - \frac{1}{a} \sum_{i=2}^N b_{i-1} x_i^{(1)}(k+1) e^{-ak} + \frac{1}{a} \sum_{i=2}^N b_{i-1} x_i^{(1)}(k+1) \right] \quad (10)$$

$$\hat{x}_1^{(0)}(k+1) = \hat{x}_1^{(1)}(k+1) - \hat{x}_1^{(1)}(k) \quad (11)$$

式中: $x_1^{(1)}(0)$ 取为 $x_1^{(0)}(1)$ 。其中,式(10)是数据序列预测值的累加,还要通过 $\hat{x}_1^{(0)}(k+1) = \hat{x}_1^{(1)}(k+1) - \hat{x}_1^{(1)}(k)$ 进行累减运算将其还原成原始数据序列的预测值。

2.2 灰色神经网络预测模型

灰色神经网络模型是一种组合预测模型,预测系统结合了灰色系统理论和人工神经网络,目的是把灰色理论弱化数据波动性和随机性的优点同神经网络非线性特点有机结合,取长补短,从而提高神经网络的融合精度^[11]。模型预测机理是首先利用灰色预测模型先进行预测,然后通过神经网络模型对灰色预测模型预测产生的残差进行调整,再结合灰色预测模型的预测数据进行综合,得到最终预测结果。灰色神经网络预测模型的方法如下:设有时间序列 $\{x_{(i)}^{(0)}\} (i = 1, 2, \dots, n)$, 利用 GM(1, N) 模型所得计算数据还原值 $\hat{x}_{(i)}^{(0)} (i = 1, 2, \dots, n)$ 。

时刻 L 还原值 $x_{(L)}^{(0)}$ 与 GM(1, 1) 模型模拟值 $\hat{x}_{(L)}^{(0)}$ 之差,成为时刻 L 的残差,记为 $e_{(L)}^{(0)}$, 即

$$e_{(L)}^{(0)} = x_{(L)}^{(0)} - \hat{x}_{(L)}^{(0)} \quad (12)$$

(1) 建立残差序列 $\{e_{(L)}^{(0)}\}$ 的 BP 网络模型

设 $e_{(L)}^{(0)}$ 为残差序列, $i = 1, 2, \dots, n$, 若预测阶数为 S , 即用 $e_{(i-1)}^{(0)}, e_{(i-2)}^{(0)}, \dots, e_{(i-s)}^{(0)}$ 作为 BP 神经网络的训练样本, 利用它们进行将来时刻值的预测。首先将 $e_{(i-1)}^{(0)}, e_{(i-2)}^{(0)}, \dots, e_{(i-s)}^{(0)}$ 作为网络训练的输入样本, 然后将 $e_{(i)}^{(0)}$ 的值作为训练的预测输出值。通过采用 BP 神经网络的“模式顺传播”与“误差逆传播”的往返更替进行网络的案例训练, 利用适当数量的残差数据序列案例进行学习收敛, 使输入样本得到经实验检验值。通过以上步骤保证了 BP 神经网络内部权重系数值、阈值经过学习收敛达到正确表示以方便进行残差序列的预测。

(2) 确定 $\{e_{(L)}^{(0)}\}$ 的新预测值。设对 $\{e_{(L)}^{(0)}\}$ 用 BP 神经网络训练模型预测出的残差序列为 $\{\hat{e}_{(L)}^{(0)}\}$, 在此基础上构造新的预测序列 $\{\hat{x}_{(i,1)}^{(0)}\}$

$$\hat{x}_{(i,1)}^{(0)} = \hat{x}_{(i)}^{(0)} + \hat{e}_{(1)}^{(0)} \quad (13)$$

则 $\hat{x}_{(i,1)}^{(0)}$ 是灰色人工神经网络组合模型的预测序列。

2.3 实验仿真

通过对江苏省水资源承载力驱动因素的分析, 建立了驱动主成分模型, 利用此模型分析出的 3 大驱动主成分选取了与需水量密切相关的 6 个相关因素, 分别为二、三产业结构比、城市化率、人均日生活用水量, 生产用水量、地表及地下水水资源量和年降水量。以年用水量为需水量的往年指标数据, 通过 1990 - 2011 年的 6 个相关因素数据, 以 MATLAB 为分析工具进行 2012 和 2013 年的江苏省水资源需求量的预测。数据仿真实验流程见图 3。

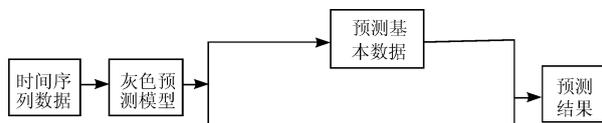


图 3 灰色神经网络预测模型预测流程图

2.3.1 基于灰色预测模型的水资源需水量预测

第一步, 首先通过灰色预测模型, 建立以江苏省年用水量为系统特征数据序列, 以二三产业结构比、城市化率、人均日生活用水量, 生产用水量、地表及地下水水资源量和年降水量 6 个因素为相关因素列的原始数据序列, 再进行累加生成序列, 根据累加生成序列, 利用分析软件进行数据拟合运算, 求出江苏省水资源需求量的 GM(1, 7) 灰色预测模型公式:

$$\hat{x}_1^{(1)}(k+1) = \left[x_1^{(1)}(0) - \frac{1}{a} \sum_{i=2}^7 b_{i-1} x_i^{(1)}(k+1) e^{-ak} + \frac{1}{a} \sum_{i=2}^7 b_{i-1} x_i^{(1)}(k+1) \right] \quad (14)$$

式中: $a = 1.3686$, $b_2 = 1445.3$, $b_3 = 3273$, $b_4 = 188.1$, $b_5 = 0.9$, $b_6 = -127.8$, $b_7 = 42.5$ 。

第二步, 通过预测模型中的年用水量累加值, 进行模型回测, 求出残差序列, 包括 22 年的年用水量预测残差序列。

第三步, 进行灰色预测模型的精度检验, 根据灰色系统理论判断模型的精度检验方法, 本文选择后验差检验方式, 其中, 原始序列 $x^{(0)}$ 的均值、方差分别为 $\bar{x} = \frac{2}{n} \sum_{k=1}^n x_{(k)}^{(0)}$, $S_1^2 = \frac{1}{n} \sum_{k=1}^n (x_{(k)}^{(0)} - \bar{x})^2$; 残差的

均值、方差分别为 $\bar{\varepsilon} = \frac{1}{n} \sum_{k=1}^n \varepsilon_{(k)}^{(0)}$, $S_1^2 = \frac{1}{n} \sum_{k=1}^n (\varepsilon_{(k)}^{(0)} - \bar{\varepsilon})^2$, $C = \frac{S_2}{S_1}$ 称为均方差比值, 对于给定的 $C_0 > 0$, 当 $C < C_0$ 时, 成模型为均方差比合格模型; $P = P(|\varepsilon_{(k)}^{(0)} - \bar{\varepsilon}| < 0.6745S_1)$ 称为小误差概率。对于给定的 $P_0 > 0$, 当 $P > P_0$ 时, 称模型为小误差概率合格模型, 其中, 平均模拟相对误差 $\bar{\Delta}$ 及均方差比值 C 要求越小越好, 小误差概率 P 要求越大越好^[12]。

通过残差序列的计算, 得出灰色预测模型的小误差概率为 0.8636 ($0.95 > C > 0.80$); 均方差比为 0.4430 ($0.35 < P < 0.5$), 根据精度检验等级参照表 (表 2) 得出, 该模型的小误差概率不和标准差比均在合格的范围内。

通过残差序列的计算, 得出灰色预测模型的小误差概率为 0.8636 ($0.95 > C > 0.80$); 均方差比为 0.4430 ($0.35 < P < 0.5$), 根据精度检验等级参照表 (表 2) 得出, 该模型的小误差概率不和标准差比均在合格的范围内。

表 2 精度检验等级参照表

精度等级	均方差比值 C_0	小误差概率 P_0
一级 (好)	< 0.35	> 0.95
二级 (合格)	< 0.50	> 0.80
三级 (勉强)	< 0.65	> 0.70
四级 (不合格)	≥ 0.65	≤ 0.7

2.3.2 基于 BP 神经网络的年用水量预测残差修正

神经网络的预测可以分为两大部分, 分别为神经网络的训练和预测两个部分。

第一步, 选取训练样本数据构造训练样本。首先, 选取 1990 - 2011 年的残差序列数据作为 BP 神经网络的训练样本, 其中, 1990 - 2007 年的数据作为 INPUT, 2008 - 2011 年的数据作为 OUTPUTS。

第二步, 对训练样本数据进行预处理。使用 premnmx() 函数将输入输出数据归一化到 $[-1, 1]$

范围内,尽可能的平滑数据,消除预测结果的噪声。

第三步,构造3层结构BP神经网络预测模型,训练网络。经过多次调试,确定BP网络结构为8-13-1,即输入层5个神经元,两个隐蔽层分别有8、13个神经元,输出层1个神经元,网络训练通过711步后误差达标。

第四步,通过仿真函数 $\text{sim}()$ 进行仿真,以及函数 $\text{postmnmx}()$ 的反归一化处理,得到2010-2013年的修正的残差序列数据。

最后灰色神经网络预测模型得到的新的预测值为灰色预测模型的预测结果和BP神经网络的残差修正数据的整合,所得预测值为表3所示。

表3 江苏省年用水量需求预测结果对比表 万t, %

年份	实际值	灰色神经网络预测	
		预测值	误差
2010	402286	402940	0.163
2011	404950	390820	-3.489
2012		401510	
2013		406730	
相对误差均值			1.826

通过对江苏省年用水需求量为实验对象,利用灰色预测模型进行数据模拟分析,得出了实验预测结果,如表3所示。从表3中可以看出,灰色预测模型相对误差平均值为1.826%,鉴于此结果,对于灰色神经网络预测模型来说,结合灰色预测模型的数据需求小,弱化灰数波动性的特点,以及BP神经网络的强大非线性映射和自学习、自适应、容错等能力特性,相较于单一的预测模型,其精度有较大幅度的提高。

3 结 语

根据本文水资源用水需求量的预测值可知,江苏省的水资源压力仍然处于不断上升的趋势。总体来看,江苏省作为我国第一大省,其经济、文化教育等方面的地位在各个省份中都是名列前茅,随着人口的不断增长和城市化进程的加快,三大产业的发展导致了用水量的不断增加,这些都不利于江苏省水资源的可持续利用发展。根据主成分分析得出的驱动成分可以看出,为保证水资源的可持续发展,要从人口、产业结构两个方面进行水资源的保护。

人口数量的增长和人均用水量的增加无疑给水资源的优化利用增加了难度,江苏省人均GDP和地区生产总值均居全国省域第一,另一方面人口压力

和水资源利用率却是排在全国末尾,这种现象形成了城市发展与水资源利用的倒“U”型曲线关系,城市经济的发展以水资源等环境状况的恶化为代价,因此,要保证人口的增长和城市化进行与水资源的可持续发展同步。

从产业发展来看,水资源是3大产业发展的重要因素,尤其以农业、和工业为突出。在所有的第一、二产业中,各个过程都需要有水的参与。因此,产业结构的调整对于水资源承载能力的提高是至关重要的,低耗能、高产出的产业结构是调整、优化产业结构布局的宗旨。利用水资源的先进技术手段,调整农业发展方向和工业发展结构、布局,加强用水管理,提高水的重复利用率,大力发展第三产业低耗水行业,与此同时,保护环境资源。

参考文献:

- [1] 杨鹏,陈金锥.改进的灰色预测模型在城市年需水量预测中的应用[J].水资源与水工程学报,2007,18(2):80-82.
- [2] 胡彩虹,吴泽宁,尹君,等.基于主成分分析的需水量预测模型研究[J].数学的实践与认识,2008,38(21):101-109.
- [3] 冯琨,张永丽,戴沂伽.基于因子分析的BP神经网络对成都市需水量的预测[J].四川环境,2011,30(2):125-130.
- [4] 李莉会,陆宝宏,张杰,等.改进BP神经网络在郑州市需水量预测中的应用[J].三峡大学学报(自然科学版),2013,35(3):26-30.
- [5] 刘汉湖,裴宗平.水资源评价与管理[M].徐州:中国矿业大学出版社,2007.
- [6] 左其亭.城市水资源承载力——理论·方法·应用[M].北京:化学工业出版社,2005.47-51.
- [7] 何晓群.多元统计分析[M].北京:中国人民大学出版社,2008:152-160.
- [8] 张吉刚.基于主成分分析与BP网络的房地产投资风险预测[J].咸宁学院学报,2008,28(3):12-15.
- [9] 李毛侠.安徽省消费需求影响因素的主成分回归分析[J].现代经济(现代物业中旬刊),2010,9(2):48-51.
- [10] 刘思峰,党耀国,等.灰色系统理论及其应用[M].北京:科学出版社,2010.
- [11] 张晓伟,沈冰,黄领梅,等.和田绿洲蒸发能力预测的多因素GM(1,3)模型[J].沈阳农业大学学报,2005,36(3):364-366.
- [12] 孙爱荣,程亚鹏.基于灰色理论和BP神经网络的房地产价格指数预测[J].企业经济,2010(4):124-126.