

基于 NAR 神经网络的月径流量预测

贺依韬

陕西省土地工程建设集团有限责任公司渭北分公司

DOI:10.32629/hwr.v4i6.3111

[摘要] 径流预测对于区域水资源的综合开发利用、合理配置、高效管理具有重要指导作用。月径流序列是典型的非平稳时间序列,NAR神经网络是一种基于时间序列的动态神经网络预测模型,在解决一些非线性,或者人们对其本质了解还不够清楚的复杂问题上表现良好,表现出了良好的智能特性。本文在某水文站50年历史月径流量的基础上,建立NAR神经网络模型对该水文站未来24个月的径流量进行预测,预测结果的不确定性系数为0.8367,预测方案的有效性为乙等,对于该区域的水文工作的开展具有重要现实意义。

[关键词] 水文与水资源; 径流预测; NAR神经网络

引言

人工神经网络(Artificial Neural Network)是上世纪80年代以来各个领域的研究热点,它的基本原理是模拟人脑中神经元的的信息处理机制而建立的一种黑箱运算模型,因为其拓扑结构呈网状,所以称其为神经网络。人工神经网络在解决一些非线性,或者人们对其本质了解还不够清楚的复杂问题上表现良好,解决了许多当前理论难以解决的实际问题。月径流序列是典型的非平稳时间序列,本文在历年逐月径流序列的基础上,建立NAR神经网络模型,以期对未来一段时间的月径流量进行预测。

1 研究方法

NAR(Nonlinear-Auto-Regression)神经网络即非线性自回归神经网络,是一种基于时间序列的动态神经网络预测模型,其网络拓扑结构如图1-1所示。

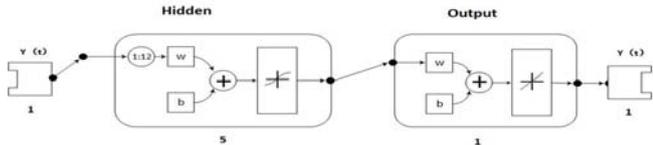


图1-1 NAR神经网络拓扑结构

其中, $y(t)$ 表示神经网络的输入, 1:12表示网络输入延时阶数为12, w 为链接权值, b 为阈值。NAR神经网络的模型可以表示为:

$$Y(t) = f(y(t-1), y(t-2), \dots, y(t-n)) \tag{1-1}$$

以 $y_{t-i}(i = 1, 2, \dots, n)$ 表示网络的输入信号, 输入数据与隐含层数据之间的连接权值为 ω_{ij} , 阈值为 b_j , 基于激活函数 f 计算出各神经元输出 H_j 。

$$H_j = f\left(\sum_{i=1}^n \omega_{ij} y_{t-i} + b_j\right) \tag{1-2}$$

式中, i 为输入数据个数; $j = 1, 2, \dots, l$; l 为隐含层神经元数目; f 为隐含层的激活函数; y_{t-i} 为第 i 个输入数据; ω_{ij} 为第 i 个输入信号和隐含层第 j 个神经元之间的连接权值; b_j 为第 j 个隐含层神经元阈值。输出层根据隐含层的输出 H_j , 做线性计算, 计算出网络的输出 o 。

$$o = f\left(\sum_{j=1}^l H_j \omega_j + b\right) \tag{1-3}$$

式中, ω_j 为隐含层第 j 个神经元与输出层神经元之间的连接权值; b 为输出层神经元阈值。

模型建立完成后, 首先需要将历史数据导入, 将历史数据按一定的比例随机分为训练数据, 验证数据, 测试数据三部分, 训练数据用于网络训练

以不断调整网络的权值和阈值以取得尽量好的拟合效果, 验证数据用于防止训练出现过拟合现象, 测试数据用来检验网络的预测效果。

其次, 对输入的数据需要进行归一化处理, 以缩小数据范围, 以便更加便捷快速地处理数据, 缩短网络地训练时间。一般使用线性函数归一法 (Min-Max scaling) 将原始数据转换到 (0, 1) 范围内。

$$x_{norm} = \frac{x - x_{min}}{x_{max} - x_{min}} \tag{1-4}$$

式中, x_{norm} 为归一化后的数据, x_{min} 为序列最小值, x_{max} 为序列最大值。

然后, 设置网络的输入时滞与隐含层神经元的个数。时间滞时的确定可根据序列自相关图初步判断; 隐含层神经元个数一般采用如下经验公式初步确定, 并经过反复试算最终确定。

$$n_1 = \sqrt{n + m} + a \tag{1-5}$$

$$n_1 = \log_2 n \tag{1-6}$$

式中, n_1 为隐含层节点数, n 为输入层节点数, m 为输出层节点数, a 为常数, a 的取值范围为 $1 \sim 10$ 。

设置好参数后要选择合适激活函数和算法对网络进行训练, 一般隐含层选用非线性sigmoid型激活函数, 以增加网络的非线性映射能力, 输出层选用线性激活函数, 将输出放大使网络可以取任意值。

模型预测过程如图1-2所示。

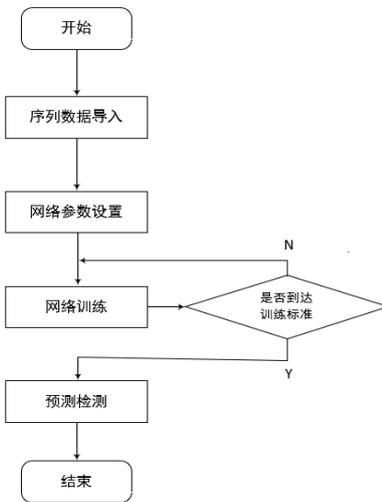


图1-2 NAR神经网络预测过程

2 实例分析

2.1 拟合与预测

本文根据水文站径流量非线性特点, 结合神经网络的非线性映射能力, 建立基于时间序列分析的动态NAR神经网络去映射出径流量变化规律。由于径流序列存在年际变化周期, 所以选择时间滞时为12, 根据经验公式和反复试算, 确定隐含层神经元个数为4, 根据建立的NAR模型, 对历史径流数据进行拟合, 结果如图2-1所示。

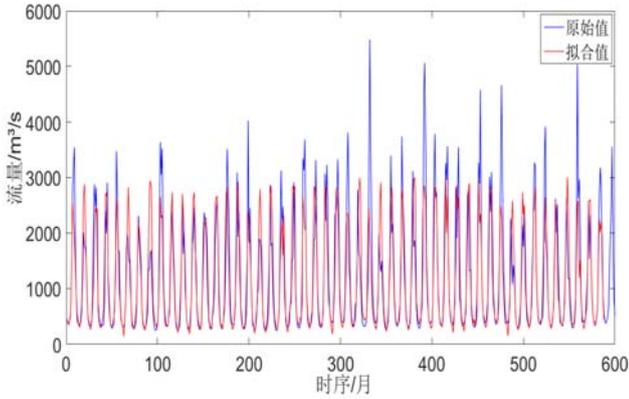


图2-1 NAR神经网络拟合图

观察拟合误差自相关图2-2, 除过零阶处显著不为零, 其它各阶误差自相关系数均在置信区间内, 说明模型拟合效果较好。

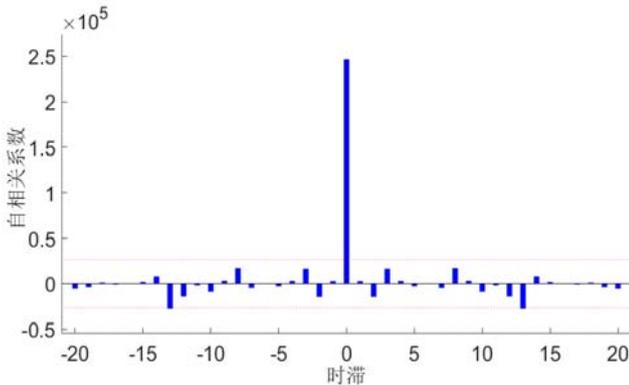


图2-2 拟合误差自相关图

采用闭环模式对未来24个月径流量进行预测, 结果如图2-3所示。

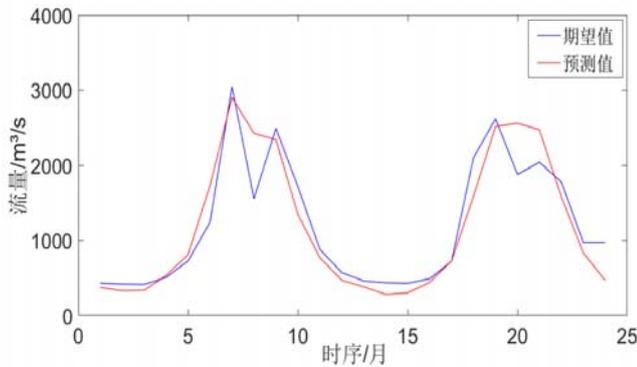


图2-3 NAR模型预测结果

NAR模型预测结果如表2-1所示。

表2-1 NAR模型预测结果

| 时序 | 实测值 | 预测值 | 误差值 | 相对误差 |
|----|---------|---------|---------|---------|
| 1 | 432.36 | 375.19 | 57.17 | 13.22% |
| 2 | 419.45 | 330.21 | 89.24 | 21.28% |
| 3 | 415.45 | 336.14 | 79.31 | 19.09% |
| 4 | 507.23 | 533.96 | -26.73 | -5.27% |
| 5 | 728.13 | 811.30 | -83.17 | -11.42% |
| 6 | 1241.93 | 1730.00 | -488.07 | -39.30% |
| 7 | 3047.03 | 2903.70 | 143.33 | 4.70% |
| 8 | 1554.42 | 2424.16 | -869.74 | -55.95% |
| 9 | 2491.80 | 2341.38 | 150.42 | 6.04% |
| 10 | 1714.16 | 1345.75 | 368.41 | 21.49% |
| 11 | 880.73 | 769.69 | 111.04 | 12.61% |
| 12 | 571.29 | 470.18 | 101.11 | 17.70% |
| 13 | 459.90 | 381.86 | 78.04 | 16.97% |
| 14 | 438.11 | 281.34 | 156.77 | 35.78% |
| 15 | 430.61 | 302.19 | 128.42 | 29.82% |
| 16 | 489.60 | 443.01 | 46.59 | 9.52% |
| 17 | 722.84 | 727.16 | -4.32 | -0.60% |
| 18 | 2101.27 | 1588.91 | 512.36 | 24.38% |
| 19 | 2618.39 | 2515.38 | 103.01 | 3.93% |
| 20 | 1878.32 | 2558.89 | -680.57 | -36.23% |
| 21 | 2041.03 | 2472.26 | -431.23 | -21.13% |
| 22 | 1776.16 | 1575.26 | 200.90 | 11.31% |
| 23 | 972.20 | 834.35 | 137.85 | 14.18% |
| 24 | 972.20 | 465.70 | 506.50 | 52.10% |

2.2 方案评定

对水文预报方案的有效性评定采用确定性系数 d_y 进行。 d_y 越大, 方案的有效性越高。

$$d_y = 1 - \frac{S_e^2}{\sigma_y^2} \quad (1)$$

$$S_e = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y})^2}{n}} \quad (2)$$

$$\sigma_y = \sqrt{\frac{\sum_{i=1}^n (y_i - \bar{y})^2}{n}} \quad (3)$$

式中, S_e 为预报的均方差; σ_y 为预报要素值的均方差; y_i 为实测值; \hat{y} 为预报值; \bar{y} 为实测系列的均值; n 为实测系列的点数据数。评定方案的有效性时按表2-2中标准进行。预报方案有效性达到甲、乙两个等级时, 方案可被用于作业预报; 方案等级为丙等时, 只可用作参考性预报; 方案有效性低于丙等时不能用于作业预报, 只能做参考性评估。

表2-2 方案有效性评定标准

| 方案的有效性 | 甲等 | 乙等 | 丙等 |
|--------|-------|-----------|----------|
| d_y | >0.90 | 0.70~0.90 | 0.5~0.69 |

计算该预测方案的确定性系数, 得 $d_y=0.8367$, 预报方案的有效性为乙等, 可用于作业预报。

3 结论

本文在对某水文站50年历年逐月径流量序列分析的基础上, 建立了NAR神经网络模型对历史径流量进行拟合并预测了未来24个月的径流量。预测结果表明NAR神经网络模型在该水文站的月径流量预测上有较好的效果, 预测的确定性系数为0.8367, 预报方案的有效性为乙等, 可以用于作业预报。

【参考文献】

[1] 舒服华. 基于NAR神经网络的湖北省社会消费品零售额预测[J]. 青海师范大学学报(自然科学版), 2019, 35(3): 24-30.
 [2] 李新男, 王宁, 梅亚东, 等. NAR神经网络的应用与检验——以城市居民生活需水定额为例[J]. 灌溉排水学报, 2017, 36(11): 122-128.
 [3] 胡宾, 崔广柏, 朱灵芝. BP神经网络预测河流月径流量[J]. 浙江水利科技, 2007, (02): 15-16+19.
 [4] 杭州电子科技大学. 基于卡尔曼滤波和NAR神经网络组合的车位预测方法: CN201810449854.6[P], 2018-10-12.
 [5] 岳新征, 李磊民, 孙飞. 基于NAR动态神经网络的石英挠性加速度表参数建模与预测[J]. 西南科技大学学报, 2016, 31(1): 88-92.