Article

自适应Elman神经网络径流预报方法

Chenming Li, Lei Zhu, Zhiyao He, Hongmin Gao \*, Yao Yang, Dan Yao and Xiaoyu Qu

Department of Information and Telecommunication Engineering, College of Computer and Information,

Hohai University, Nanjing 210000, China; lcm@hhu.edu.cn (C.L.); zhuleihohai@hhu.edu.cn (L.Z.);

181607010036@hhu.edu.cn (Z.H.); rcyyang@hhu.edu.cn (Y.Y.); y.dan@hhu.edu.cn (D.Y.);

qu-xiaoyu@hhu.edu.cn (X.Q.)

**\*** Correspondence: gaohongmin@hhu.edu.cn

Received: 27 March 2019; Accepted: 22 May 2019; Published: 28 May 2019

摘要： 中长期径流预报对水资源的科学利用具有重要意义。从众多的时间序列因子中自动筛选出影响显著的因子，构建一个基于数据驱动的自适应神经网络预报模型，是一项具有吸引力和复杂的任务。 目前，常用的筛选因子方法主要是相关分析法，这种方法选出了很多相关性高的因子，相互之间存在复共线性。如果直接把这些因子作为模型的输入，会使模型参数增加，且过多的冗余和噪声会对模型预报结果产生不利影响。本文在对国内外中长期径流预报方法研究的基础上，提出了互信息和核主成分分析驱动的Elman神经网络径流预报方法。该方法利用互信息筛选因子，使用核主成分分析提取筛选的因子的特征，然后将提取的特征输入Elman神经网络，进行训练。以较少的特征作为Elman神经网络模型的输入，可以减少Elman神经网络模型的参数，并且能够有效的缓解Elman神经网络模型的过拟合问题。利用中国成都的锦屏一级水电站2007年到2011年的年平均径流数据对该方法进行评估，多次预报的最大相对误差都在16%以内，预报效果良好。对多次预报结果取平均值，预报的精度进一步提高。

关键词：互信息；核主成分分析；循环神经网络；径流预报；锦屏一级水电站

1. 引言

径流预报，特别是中长期径流预报，在水资源的综合开发、利用、科学管理和优化中起着重要的作用[1-4]。极端洪水，近年来似乎更频繁发生（由于气候变化），每年给全世界人类带来巨大的痛苦，并造成巨大的经济损失。因此，有必要在洪水事件之前准确预测峰值流量的时间和大小[5]。准确的中长期径流预报是水资源综合开发利用、科学管理和优化调度的重要前提。在过去的几十年中，大量的径流预测方法和应用研究在国内外进行，这些方法，它们大致可以分为： 数据驱动模型和过程驱动模型。数据驱动模型是指，基于历史数据，建立预测对象（如年平均径流）和预测变量（如环流）之间

的最优数学关系，这些数学关系可用于预测未来的水文变量，而不用考虑具体的水文过程。传统的用于建立数学关系的方法包括线性回归、逐步回归[7]、局部回归，人工神经网络[8-10]，以及支持向量机[11-13]。然而，过程驱动模型需要一个能够反映径流特征的水文模型，需要考虑径流形成的物理机制[14]。美国学者Day [15]提出的集合预测方法（ESP）是一种过程驱动的模型，研究人员已经用这种方法对许多流域的中长期径流预测进行了预测。然而，由于水文过程机制尚未完全阐明，该模型的适用性有限[16-20]。因此，数据驱动模型，特别是基于神经网络的径流预测模型，已成为反向传播（BP）神经网络应用于中长期水文预报的一个重点课题。在[26-28]中，研究了小波神经网络在径流预测中的应用。在[29]中，研究了基于BP网络模型的灰色自记忆模型在径流预测中的应用。然而，这些神经网络模型有两个缺点：容易陷入局部最优和收敛缓慢[30]。肖跃红等人进一步比较和分析了Elman神经网络和水文模型在径流预报中的应用。

目前，中长期径流预测常用的方法是以统计方法为基础的，即通过寻找预测对象与预测因子之间的统计关系来实现预测。

将统计学的方法用于中长期径流预报的关键问题包括以下三方面：首先，水文过程复杂，预报因子与预报对象之间除了线性关系外还存在一定的非线性关系。用于预报因子初选的线性相关分析法只能描述变量间的线性关系，不能反映变量间的非线性关系。其次，用于初选因子降噪和去冗余的PCA, 实质是一种线性映射方法，得到的主成分是由线性映射生成的。这种方法忽略了数据之间高于2阶的相互关系，所以抽取的主成分并不是最优的。最后，用于建立预报对象与预报因子之间最优数学关系的模型，常用的多元回归实际上也是一种线性拟合，不能反映预报对象与预报因子之间的非线性关系。与其他模型相比，人工神经网络因鲁棒性好、非线性映射和自学习能力强，在中长期径流预报中得到了较为广泛的应用，但神经网络模型参数的不确定性会对预报的准确性造成一定的影响，每次预报的结果之间会存在一定幅度的差异。

1990年，Elman提出了Elman神经网络，并用它来解决语音处理问题[32]。Elman 网络是一个循环神经网络，具有适应时变特性的能力。与正反馈神经网络不同，它有来自隐藏层神经元的输出到其输入层的反馈连接。其神经元的状态不仅取决于当前的输入信号，还取决于神经元的前一个状态[33]。因此，Elman 神经网络可以保持某种状态，允许它执行序列预测等任务。然而，中长期径流预测领域的相关研究十分有限。

与以往的方法相比，本文的主要贡献和解决的问题如下：

第一，由于实验数据之间的非线性关系，我们采用了基于NMI的预测因子初选方法，该方法不仅能反映变量之间的线性关系，也反映了变量之间的非线性关系。NMI克服了传统线性相关分析的缺陷。第二，核主成分分析法（KPCA）是主成分分析法（PCA）的非线性扩展，即通过映射函数把原始向量映射到高维特征空间，在高维空间上进行PCA分析。原始空间中的线性不可分的数据在高维特征空间中几乎都能线性可分，此时在高维空间做PCA，提取的主成分更都具有代表性。因此基于KPCA的特征提取方法大大提高了非线性数据的处理能力，与传统的基于PCA的特征提取方法相比，更有优势。此外，经KPAC提取的主成分之间相互正交，且数据经过了降噪去冗余，能够很好的防止神经网络的过拟合，提高网络的泛化能力。第三，人工神经网络鲁棒性好、非线性映射和自学习能力强，能够很好的挖掘预报因子与预报对象之间的内在联系。本发明选用的Elman神经网络，是一种典型的动态回归网络，与常用的前向神经网络（例如BP神经网络）相比，多增加了承接层。承接层能够记录上一次网络迭代的信息并作为当前迭代的输入，这使得Elman网络更适合时间序列数据的预测。此外，神经网络存在参数的不确定性问题，为了减少预报的不确定性，采用了多模型集合预报的方法，提高预报精度。

总之，本文使用的NMI、KPCA和Elman神经网络具有处理线性数据以外的非线性数据的能力，此外，处理后的数据去除了噪声和冗余，可以防止神经网络的过度拟合，提高网络的泛化能力。三种方法组合在一起，能够克服传统方法的局限性，提高预报的稳定性与准确性。本文的主要目的是构建一个自适应的数据驱动的径流预报模型[36-37]，该模型通过标准互信息的方法来初步筛选因子，然后使用 KPCA 方法从所选因子中提取特征，最后，在此基础上，构建了径流预测的循环神经网络模型。通过对实验结果的分析和比较，单模型预报的精度有改进，多模型预报的精度有进一步提升，为中长期径流预报提供了一定的参考。

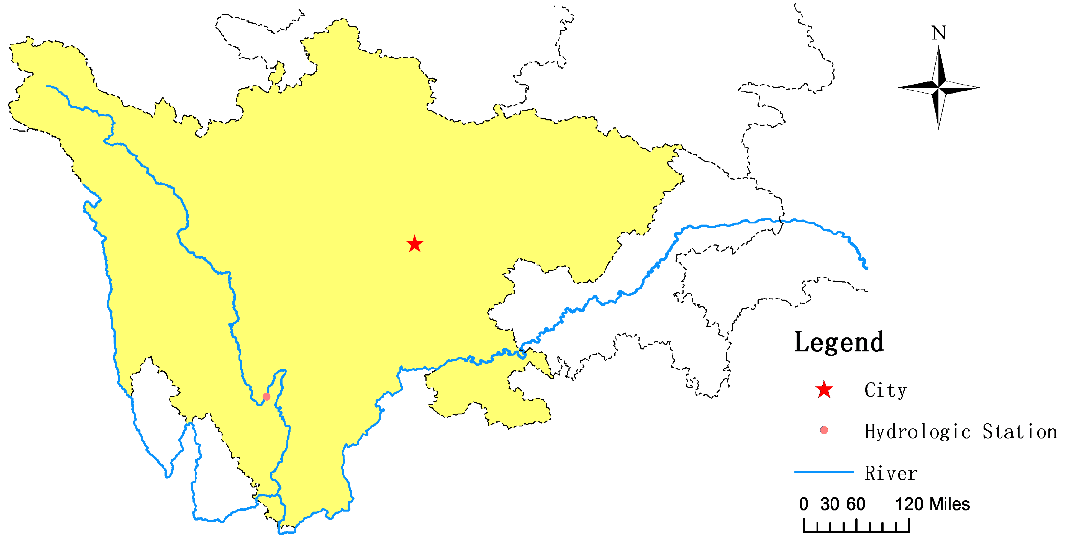
2. 数据与方法

2.1. 研究区域和数据

这项研究的地点是位于中国四川省雅龙江的锦屏I级水电站，其中的雅龙江是金沙江最大的支流，金沙江是长江的上游。水电站主要用于发电、蓄水和防洪。此外，水库的周围是复杂交错的山脉和河流。因此，利用该水库的数据进行研究，实验数据变得更加真实和具有代表性。准确预测该区径流，有利于该区水资源的综合开发利用，实验模型也可以经过简单处理后，应用于其他地区。锦屏I级水电站的位置如图1所示。该模型的实验数据使用了锦屏I级水电站1960年至2011年的平均径流数据（由中国水利水电科学研究院提供）和1959年至2010年的74个大气环流指数（由中国国家气候中心提供）。其中，1960-2006 年的数据用于训练模型，2007-2011年的数据用于验证模型。

坝址年平均流量为1220立方米/秒，6月至10月洪水季节年平均流量为2230立方米/秒，11月至5月干燥季节年平均流量为493立方米/秒，年平均径流量为385亿立方米。

Figure 1. Location of study area.



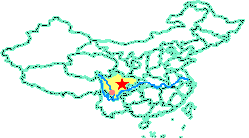
Chengdu

Jinping I Hydropower Station

Chin-sha River

Yangtze River

Ya-lung River



China

2.2. 方法

本文提出的径流预报方法主要包含三个部分：基于标准互信息的预报因子的选择，基于核主成分分析的预报因子主成分的提取，基于循环神经网络的径流的预报。下文将详细介绍这三个部分。

2.2.1. *基于标准互信息的因子初选*

在概率论与信息论中，互信息是两个变量之间相互依存的一种度量[1]。本文通过计算因子时间序列与径流时间序列的互信息，选出标准互信息大于某一阈值（一般取0.9）的因子作为预报因子。基于互信息的预报因子的选择方法，不仅能够反映因子与径流之间的线性关系，还能反应他们之间的非线性关系。而传统的基于线性相关分析（皮尔逊相关，斯皮尔曼相关等）筛选因子的方法，只能反应因子与径流之间的线性关系。因此，基于互信息的预报因子选择方法选出来的因子更具有代表性。径流时间序列与因子时间序列之间的互信息的计算公式如下：

其中，X为径流时间序列，,Y为因子时间序列，,分子为X和Y的联合分布律，和分别为X与Y的边缘分布律。

为了比较的方便，需要将互信息标准化。标准化互信息的值在0与1之间。标准互信息的计算公式如下：

其中，和分别为X和Y的熵，和的计算公式如下：

2.2.2. 基于*核主成分分析的特征提取*

核主成分分析法是主成分分析法非线性扩展，即通过核函数Φ把原始向量映射到高维特征空间F，然后在F上进行主成分分析。原始空间中的线性不可分的数据在高维特征空间中几乎都能线性可分，在高维空间中提取的主成分更都具有代表性。因此，基于核主成分分析的特征提取方法大大提高了非线性数据的处理能力，与传统的基于主成分分析的特征提取方法相比，更有优势。此外，核主成分分析提取的主成分之间相互正交，并且主成分经过了降噪和去冗余，能够很好的缓解神经网络过拟合，提高网络的泛化能力。运用核主成分分析法提取主成分的步骤如下：

步骤1： 将2.2.1中选择的预报因子数据通过z-score标准化，公式如下

公式中，为z-score标准化后的数据，为预报因子数据，为所处的时间序列的均值，为所处的时间序列的标准差。

步骤2：计算预报因子的核矩阵。预报因子数据矩阵为的矩阵，每一行对应一个因子，每一列对应一个样本。为的矩阵，第i行第j列的元素的计算公式如下：

公式中，和列向量，代表着预报因子数据矩阵中的样本，\*表示数据经过了z-score表准化，为径向基核函数（Radial Basis Function），

步骤3: 计算中心化的核矩阵。为的矩阵，计算公式为：

公式中为的矩阵，。

步骤4：计算的特征值与特征向量，并把特征值按照由大到小的顺序排列，特征向量的顺序按照特征值做对应的调整。排序后得到的特征值矩阵，特征向量矩阵。

步骤5：计算归一化的特征向量矩阵,的计算公式如下：

公式中，列向量。

步骤6：计算主成分矩阵。主成分矩阵为 的方阵，一般提取前2到3个主成分作为神经网络的输入。第j个主成分为列向量，计算公式如下：

公式中为步骤3中计算得到的中心化核矩阵的转置。

2.2.3. *Elman 神经网络*

Elman网络是一种典型的回归神经网络，其结构图如图2。Elman网络不仅有前向神经网络的输入层、输出层和隐含层，而且增加了一个承接层。因为承接层能够记录上一次网络迭代的信息并作为当前迭代的输入，所以,与前向神经网络相比，Elman神经网络更适合时间序列数据的预测[3]。



图2 Elman网络结构图

Elman网络模型的计算过程可以简单的表示为：

在t时刻输出层的输出：

在t时刻隐含层的输出:

在t-1时刻承接层的输出：

公式中，、 和分别为各层之间的连接权重系数。为激活函数，本文隐含层的激活函数取Sigmoid函数 ，输出层的激活函数取线性函数。

Elman网络的学习过程可以概括如下：

步骤1：采用随机函数初始化网络各层之间的连接权重系数，并确定成本函数（cost function）允许的误差。本文采用均方误差函数作为成本函数:

其中，为t时刻网络的期望输出，为网络的实际输出（径流量的观察值）。

步骤2：对输入数据归一化，计算E的值，并根据E采用动量梯度下降算法更新网络各层之间的连接权值。归一化输入数据的公式和权值的改变量公式分别如下：

其中，为归一化后的数据，，，，为原始的径流序列或主成分序列中的一项，为所在的序列中的最小值，为所在的序列中的最大值。为第k次更新时Elman神经网络权值的改变量，为第k-1次更新时Elman神经网络权值的改变量，为动量常数，，本文中，为学习速率，本文中。 。

步骤3：当E的值大于时，转到步骤2，否则学习结束，并根据公式(14)、(15)和(16)计算网络输出值。

2.2.4. *评价标准*

为了评价模型的表现，本文采用合格率、均方根误差、平均绝对百分误差和平均绝对误差作为评价指标。

合格率的计算公式如下：

其中，m为合格的预报次数，n为总的预报次数，单次的预报误差小于20%时,预报合格。模型的表现可以采用合格率来评估，评估标准如下：

**Table 1.** **预报精度等级**

|  |  |
| --- | --- |
| **Qualified rate** | **Accuracy Level** |
|  | A |
|  | B |
| % | C |

均方根误差的计算公式如下：

公式中，为预测值，为观察值。

平均绝对百分误差的计算公式如下：

平均绝对误差的计算公式如下：

相对误差计算公式如下：

3. 结果与讨论

3.1.预报案例的实现

为了验证本文提出的模型的有效性，采用本文提出的模型对中国四川的雅砻江流域锦屏一级水电站的年平均径流量进行预报。详细的过程如图3所示：



图3 径流量预报流程图

3.1.1. *确定预报因子集合*

如图 3中所示，首先，需要进行物理机制分析，并采用互信息的方法筛选预报因子。具体步骤如下：

步骤1：收集研究地区的历史径流数据和有能可作为预报因子的气象水文资料。常用的气象水文资料包括大气环流特征、高空气压场以及海表温度等指标。本研究案例采用的数据资料包括锦屏一级水电站水库断面1960～2011年的逐年年平均径流量数据和1959～2010年的逐月74项环流特征量数据。

步骤2：由于是对年平均径流量预报，因此因子不能从同期当年的时间内选择，同时，考虑到气象因子对径流的影响存在滞后性，所以，根据表1，建立锦屏一级水电站逐年（1960～2011年）年平均径流量与前一年（1959～2010年）逐月的74项大气环流指数的一一对应关系。其中某一项大气环流指数时间序列与径流时间序列的对应关系如表2，其他指数与径流量的关系类似。

表2 某大气环流指数时间序列与径流时间序列的对应关系

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 年径流时间序列 | 某一项大气环流指数时间序列 | | | |
| 1960年径流量 | 1959年1月数据 | 1959年2月数据 | … | 1959年12月数据 |
| 1961年径流量 | 1960年1月数据 | 1960年2月数据 | … | 1960年12月数据 |
| 1962年径流量 | 1961年1月数据 | 1961年2月数据 | … | 1961年12月数据 |
| … | … | … | … | … |
| 2009年径流量 | 2008年1月数据 | 2008年2月数据 | … | 2008年12月数据 |
| 2010年径流量 | 2009年1月数据 | 2009年2月数据 | … | 2009年12月数据 |
| 2011年径流量 | 1959年1月数据 | 1959年2月数据 | … | 2010年12月数据 |

步骤3：将大气环流指数时间序列与年平均径流时间序列分成两部分，一部分作为训练样本，另一部分作为检验样本。本实施例用前47年的数据作为训练样本，后5年的数据作为检验样本。

步骤4：计算互信息。对于本实施例，即按照公式（1）计算表2中第1列的年平均径流时间序列与表中剩余各列的大气环流指数时间序列之间的互信息。需要注意的是，只使用训练样本数据计算互信息，检验样本数据不应该加入，确保检验的可靠性。

步骤5：计算标准化互信息，即用公式（2）、（3）和（4）把步骤4计算得到的MI值映射到0与1之间。

步骤6：选出标准互信息大于某一阈值（本实施例取0.9）的指标作为初选的预报因子。本实施例中，标准互信息大于0.9的指标有205个，前20个指标的信息如下：

表3 前20个初选的预报因子

|  |  |  |
| --- | --- | --- |
| 初选的因子 | NMI | MI |
| 8月太阳黑子 | 0.988375 | 5.426929 |
| 4月太阳黑子 | 0.988375 | 5.426929 |
| 7月太阳黑子 | 0.988375 | 5.426929 |
| 10月太阳黑子 | 0.988375 | 5.426929 |
| 12月太阳黑子 | 0.988375 | 5.426929 |
| 2月太阳黑子 | 0.98444 | 5.384376 |
| 9月太阳黑子 | 0.98444 | 5.384376 |
| 11月太阳黑子 | 0.98444 | 5.384376 |
| 1月太阳黑子 | 0.98444 | 5.384376 |
| 3月太阳黑子 | 0.98444 | 5.384376 |
| 5月太阳黑子 | 0.98444 | 5.384376 |
| 8月北半球副高强度指数(5E-360) | 0.980474 | 5.341823 |
| 3月北半球极涡面积指数(5区,0-360) | 0.980474 | 5.341823 |
| 6月北非大西洋北美副高强度指数(110W-60E) | 0.976477 | 5.299270 |
| 6月北半球副高强度指数(5E-360) | 0.976291 | 5.256717 |
| 4月北半球副高强度指数(5E-360) | 0.972448 | 5.256717 |
| 7月北非大西洋北美副高强度指数(110W-60E) | 0.972448 | 5.256717 |
| 9月北非大西洋北美副高强度指数(110W-60E) | 0.972448 | 5.256717 |
| 6月太阳黑子 | 0.972448 | 5.256717 |
| 6月太平洋副高强度指数(110E-115W) | 0.970919 | 5.240655 |

3.1.2. *提取主成分*

如图 3中所示，采用互信息的方法筛选因子后，需要进行核主成分分析，提取主成分作。在第3.1.1节中，本研究案例选出了205个因子，这些因子序列之间往往存在多重共线性，重复的信息和噪声会直接影响Elman神经网络的训练速度和泛化能力，因此需要进行特征提取。本实例按照公式（5）、（6）、（7）与（8）计算主成分，主成分按照方差贡献率由大到小的顺序排列。前5个主成分的方差贡献率如表4，前5个主成分数据如表5。

**Table 4.** 前5个主成分的方差贡献率

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 主成分 | 主成\_1 | 主成分­­\_2 | 主成分\_3 | 主成分\_4 | 主成分\_5 |
| 方差贡献率 | 25.7% | 6.9% | 5.6% | 5.1% | 3.9% |

**Table 5.** KPCA提取的前5个主成分

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 样本序号 | 主成分\_1 | 主成分\_2 | 主成分\_3 | 主成分\_4 | 主成分\_5 |
| 1 | -0.00605 | -0.025852 | 0.007556 | 0.013533 | 0.017703 |
| 2 | -0.02725 | -0.006741 | 0.001084 | 0.017781 | 0.009853 |
| 3 | 0.007912 | 0.0031653 | -0.00755 | 0.007487 | -0.02185 |
| 4 | 0.002089 | 0.0131535 | 0.006936 | -0.00139 | -0.00235 |
|  |  |  |  |  |  |
| 48 | -0.05354 | 0.0081831 | 0.013478 | -0.03407 | 0.018846 |
| 49 | -0.04108 | 0.0216432 | 0.00613 | -0.01603 | -0.02593 |
| 50 | -0.01539 | 0.0190786 | 0.005982 | -0.02812 | 0.013842 |
| 51 | -0.03675 | 0.0243891 | 0.00997 | -0.00845 | 0.013842 |
| 52 | -0.07989 | 0.0283544 | -0.04816 | 0.007058 | -0.00822 |

从表4可以看出，第一个主成分的方差贡献率已经达到了25.7%，包含了所选因子的大部分信息。其他主成分的方差贡献率越来越小，包含所选的因子的信息也越来越少。值得注意的是，为了训练样本和检验样本提取主成分时所用的标准是一致的，需要将训练样本序列和检验样本序列组合在一起进行KPCA。本实施例中，训练样本序列的长度为47，检验样本序列的长度为5，序列样本和检验样本序列组合的长度为52，因此，表4中提取的主成分的序列长度为52。

3.1.3. *确定Elman神经网络结构*

如图 3中所示，提取主成分后，需要确定Elman网络的结构。即，训练算法，输入层的节点数、隐含层的节点数 、承接层的节点数和输出层的节点数需要确定。本研究案例采用动量梯度下降算法和反向传播算法作为Elman网络的训练算法；输出层的节点数等于预报对象的个数,本实施例是对年平均径流量进行单值预报，因此输出层节点个数为1；承接层的节点数等于隐含层的节点数，因此，只要确定隐含层的节点数，承接层的节点数就确定了；输入层的节点个数等于被选取的主成分的个数。隐含层节点的数目对网络的泛化性能有重要影响，但是，目前还没有一个系统和标准的方法来确定隐含层节点数。本研究案例采用试错法，即通过输入层的节点数和隐含层的节点数的不同组合，比较Elman神经网络的预报结果，确定最佳的输入层的节点和隐含层的节点数的组合。

本文将前47年的主成分序列和径流时间序列数据作为训练样本，后5年的数据作为检验样本。各个不同结构的Elman神经网络模型的预报结果如下。

**Table 5.** 以第一主成分作为输入时，不同隐含层节点的Elman网络的表现

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Model** | **Arcet-eture** | **Training** | | |  | | **Validation** | | | | |
| **MAPE RMSE MAE** | | | | **MRE MAPE RMSE MAE QR** | | | | | |
| 1 | 1-2-1 | 0.160 | 224.616 | 189.946 | | 0.272 | | 0.133 | 172.002 | 138.249 | 60% |
| 2 | 1-3-1 | 0.161 | 224.755 | 190.386 | | 0.257 | | 0.129 | 165.348 | 133.795 | 60% |
| 3 | 1-4-1 | 0.160 | 224.426 | 189.586 | | 0.255 | | 0.128 | 163.045 | 132.988 | 60% |
| 4 | 1-5-1 | 0.160 | 224.339 | 189.688 | | 0.260 | | 0.129 | 165.309 | 134.334 | 60% |
| 5 | 1-6-1 | 0.156 | 222.070 | 184.753 | | 0.348 | | 0.146 | 191.760 | 152.287 | 60% |
| 6 | 1-7-1 | 0.155 | 221.486 | 183.193 | | 0.371 | | 0.146 | 196.996 | 152.686 | 60% |
| 7 | 1-8-1 | 0.154 | 221.553 | 182.727 | | 0.375 | | 0.149 | 199.493 | 155.148 | 60% |
| 8 | 1-9-1 | 0.153 | 221.309 | 182.060 | | 0.382 | | 0.152 | 203.666 | 158.386 | 60% |
| 9 | 1-10-1 | 0.154 | 221.619 | 183.238 | | 0.362 | | 0.145 | 193.697 | 151.519 | 60% |
| 10 | 1-11-1 | 0.154 | 221.698 | 182.949 | | 0.363 | | 0.147 | 195.395 | 153.525 | 60% |
| 11 | 1-12-1 | 0.156 | 222.047 | 185.180 | | 0.330 | | 0.143 | 185.852 | 148.957 | 60% |
| 12 | 1-13-1 | 0.154 | 221.828 | 183.298 | | 0.357 | | 0.148 | 195.269 | 154.778 | 60% |
| 13 | 1-14-1 | 0.154 | 221.582 | 182.858 | | 0.369 | | 0.149 | 198.017 | 155.001 | 60% |
| 14 | 1-15-1 | 0.162 | 225.553 | 192.005 | | 0.286 | | 0.139 | 179.275 | 143.302 | 60% |

表格5记录了以主成分\_1作为输入，隐含层的节点从2到15变化时，Elman神经网络的表现。从表格5可以看出，以主成分\_1作为Elman神经网络输入，随着隐含层节点数的增加，Elman神经网络在验证期的预报的最大相对误差均大于20%，合格率均只有60%。其中预报效果最好的模型是Model 3，在验证期的预报最大相对误差为25.5%，合格率为60%，预报效果欠佳。

**Table 6.** 以第一、二主成分作为输入时，不同隐含层节点的Elman网络的表现

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Model** | **Arcet-eture** | **Training** | | |  | | **Validation** | | | | |
| **MAPE RMSE MAE** | | | | **MRE MAPE RMSE MAE QR** | | | | | |
| 1 | 2-2-1 | 0.155 | 222.051 | 182.820 | | 0.295 | | 0.138 | 173.111 | 144.462 | 60% |
| 2 | 2-3-1 | 0.148 | 213.096 | 176.010 | | 0.216 | | 0.115 | 143.593 | 122.346 | 80% |
| 3 | 2-4-1 | 0.131 | 187.552 | 153.562 | | 0.289 | | 0.142 | 177.806 | 153.352 | 80% |
| 4 | 2-5-1 | 0.138 | 200.655 | 164.461 | | 0.158 | | 0.078 | 105.308 | 88.466 | 100% |
| 5 | 2-6-1 | 0.132 | 189.035 | 155.098 | | 0.170 | | 0.086 | 125.228 | 97.631 | 100% |
| 6 | 2-7-1 | 0.132 | 188.624 | 154.443 | | 0.178 | | 0.090 | 129.553 | 102.225 | 100% |
| 7 | 2-8-1 | 0.136 | 190.345 | 159.353 | | 0.165 | | 0.080 | 115.598 | 89.780 | 100% |
| 8 | 2-9-1 | 0.135 | 197.805 | 160.718 | | 0.127 | | 0.069 | 95.339 | 78.964 | 100% |
| 9 | 2-10-1 | 0.129 | 188.338 | 152.090 | | 0.155 | | 0.097 | 124.622 | 109.113 | 100% |
| 10 | 2-11-1 | 0.131 | 189.741 | 154.595 | | 0.152 | | 0.086 | 116.728 | 97.551 | 100% |
| 11 | 2-12-1 | 0.140 | 200.470 | 164.854 | | 0.159 | | 0.092 | 112.918 | 102.019 | 100% |
| 12 | 2-13-1 | 0.134 | 192.657 | 157.422 | | 0.156 | | 0.089 | 119.336 | 99.144 | 100% |
| 13 | 2-14-1 | 0.134 | 193.523 | 158.657 | | 0.166 | | 0.097 | 119.267 | 107.923 | 100% |
| 14 | 2-15-1 | 0.134 | 191.521 | 157.844 | | 0.155 | | 0.084 | 114.188 | 95.899 | 100% |

表格6记录了以主成分\_1和主成分\_2作为输入，隐含层的节点从2到15变化时，Elman神经网络的表现。从表格6可以看出，以主成分\_1和主成分\_2作为Elman神经网络输入，隐含层节点数从2增加到9时，Elman神经网络在验证期的预报的最大相对误差呈现下降趋势。当隐含层的节点数从9增加到15，在验证期的预报最大相对误差呈现先上升后保持稳定的趋势。当隐含层的节点数大于4后，在验证期，预报的合格率都达到了100%。其中，预报效果最好的模型是Model 8，预报最大相对误差为6.9%，合格率为100%，预报效果良好。

**Table 7.**以第一、二和三主成分作为输入时，不同隐含层节点的Elman网络的表现

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Model** | **Arcet-eture** | **Training** | | |  | | **Validation** | | | | |
| **MAPE RMSE MAE** | | | | **MRE MAPE RMSE MAE QR** | | | | | |
| 1 | 3-2-1 | 0.147 | 214.506 | 176.575 | | 0.274 | | 0.119 | 153.358 | 124.235 | 80% |
| 2 | 3-3-1 | 0.147 | 213.411 | 176.485 | | 0.202 | | 0.106 | 132.191 | 113.081 | 80% |
| 3 | 3-4-1 | 0.122 | 179.191 | 143.758 | | 0.438 | | 0.149 | 213.823 | 155.190 | 80% |
| 4 | 3-5-1 | 0.121 | 170.440 | 141.632 | | 1.106 | | 0.319 | 509.077 | 336.476 | 60% |
| 5 | 3-6-1 | 0.125 | 184.424 | 148.838 | | 0.197 | | 0.115 | 136.876 | 128.372 | 100% |
| 6 | 3-7-1 | 0.119 | 172.714 | 141.197 | | 0.471 | | 0.197 | 278.944 | 222.419 | 60% |
| 7 | 3-8-1 | 0.113 | 167.382 | 133.095 | | 0.239 | | 0.136 | 184.009 | 155.512 | 60% |
| 8 | 3-9-1 | 0.136 | 194.311 | 161.320 | | 0.191 | | 0.110 | 133.774 | 121.915 | 100% |
| 9 | 3-10-1 | 0.117 | 173.909 | 138.046 | | 0.280 | | 0.143 | 195.198 | 163.710 | 60% |
| 10 | 3-11-1 | 0.129 | 181.719 | 152.612 | | 0.169 | | 0.114 | 146.366 | 126.855 | 100% |
| 11 | 3-12-1 | 0.126 | 187.950 | 150.310 | | 0.207 | | 0.106 | 131.850 | 118.583 | 80% |
| 12 | 3-13-1 | 0.115 | 169.871 | 136.686 | | 0.379 | | 0.135 | 211.644 | 149.332 | 60% |
| 13 | 3-14-1 | 0.114 | 166.467 | 133.790 | | 1.170 | | 0.294 | 527.855 | 303.349 | 60% |
| 14 | 3-15-1 | 0.128 | 188.502 | 152.513 | | 0.179 | | 0.113 | 148.934 | 126.428 | 100% |

表格7记录了以主成分\_1、主成分\_2和主成分\_3作为输入，隐含层的节点从2到15变化时，Elman神经网络的表现。从表格7可以看出，以主成分\_1、主成分\_2和主成分\_3作为Elman神经网络输入，随着隐含层节点的增加，评价指标值波动大，预报的效果欠佳。其中，预报效果最好的模型是Model 10，验证期的预报最大相对误差为16.9%，合格率为100%。

**图4在验证期各模型的预报最大相对误差**

为了更加直观的比较以不同主成分组合作为输入时，隐含层节点数不同的各网络模型预报效果，图4被绘制了。图4为表1、表2与表3中的最大相对误差。从图4中以可以看出，以主成分\_1和主成分\_2作为Elman神经网络的输入，在验证期各模型的表显明显好于以主成分\_1或者同时以主成分\_1、主成分\_2和主成分\_3作为Elman神经网络输入时的表现。根据以上分析，本研究案例选择主成分\_1和主成分\_2作为Elman神经网络的输入。因此，输入层的节点为2。此外，从表6中可以看出，当隐含层的节点数为9个时，在验证期预报效果的各评价指标达到最好，因此，本研究案例中模型的隐含层的节点数为9。

3.2. 预报表现评估

根据步骤3.1.3确定的网络结构，采用1960～1999年的锦屏一级水电站年平均径流量数据对Elman网络模型进行训练，采用2000～2011年的锦屏一级水电站年平均径流量数据对Elman网络模型进行检验。

在验证期，5次预报的误差如表8所示。预报结果以平均绝对百分误差(MAPE, Mean Absolute Percentage Error)、最大相对误差(MRE, Maximum Relative Error)、合格率（QR, Qualified Rate）为评价指标。

表8各模型在检验期的预报误差

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 评价指标 | | 年份 | 第一次 | 第二次 | 第三次 | 第四次 | 第五次 | 集合预报 | |
| RE | | 2007 | 14.10% | 13.15% | 11.41% | 12.86% | 15.24% | 13.35% | |
| 2008 | -14.85% | -14.20% | -12.93% | -14.76% | -11.26% | -13.60% |
| 2009 | 1.63% | 2.26% | 3.70% | 1.91% | 6.57% | 3.21% | |
| 2010 | 4.63% | 6.80% | 7.97% | 4.96% | 6.78% | 6.23% |
| 2011 | -4.51% | -11.77% | -13.70% | -2.06% | -12.95% | -9.00% | |
| QR | — | | 100% | 100% | 100% | 100% | 100% | 100% | |
| MRE | — | | 14.85% | 14.20% | 13.70% | 14.76% | 15.24% | 13.60% | |
| MAPE | — | | 7.94% | 9.64% | 9.94% | 7.31% | 10.56% | 9.09% | |

从表8中可以看出，在验证期内，单次预报的最大相对误差都在16%以内，根据文献38中对中长期径流预报的精度评价方案，合格率都达到了100%。这说明了基于互信息和核主成分分析驱动的Elman神经网络模型具有良好的泛化能力和预报稳定性。此外，从表8中还可以看出，单个模型对于不同年份的年平均径流量的预报误差和5个模型对于同一年的年平均径流量的预报误差都不相同。各个模型的预报误差主要是由Elman神经网络的参数的不确定性造成的。反向传播算法和动量梯度更新算法实现了对Elman神经网络的参数空间的搜索，通过不断的训练，减小了历史径流数据的实际值和Elman神经网络的预测值之间的误差。但是，误差曲面可能含有多个不同的局部极小值，在对Elman神经网络的参数空间的搜索过程中，可能会停留在局部极小值点，而不一定是全局最小值点。因此，虽然各个Elman神经网络的结构相同，但是参数是不同的，这导致各个Elman神经网络预测结果之间的差异。

为了减少由于模型参数不确定性而导致的预测结果的偏差，本应用实例采用集合预报的方法，即做5次径流量预报，并把这5次预报结果的平均值作为最终的预报结果。从表8可以看出，集合预报最大的预报相对误差小于13.60%，与单次预报的结果相比，预报精度有进一步提高。

4. 结论

由于原始数据的误差和模型参数的不确定性，单个预测模型之间的预测结果可能大不相同。本文在互信息分析的基础上，结合物理原因分析，确定了预测因子，实现了基于Elman神经网络的年平均径流的单模型预报和多模型集合预报，为水库中长期径流预报提供参考。本文中描述的一般方法流程如图 5 所示。

本文详细讨论了预报因子的自动选择、预报因子的自动特征提取以及Elman神经网络的自适应构建，提出了由互信息和核主成分分析驱动的Elman神经网络模型。将该模型应用于中国四川省锦屏一级水电站的年平均径流预报，预报结果表明，基于标准化相互信息和KPCA的因子筛选方法可以有效降低大量预报因子的噪声和冗余。以这些因子作为Elman神经网络的输入，Elman神经网络具有良好的泛化性能，预报的稳定性较高，预报的精度能够满足实际生产的需要。采用多个神经网络的集合预报的方式，能够有效解决Elman神经网络模型的参数不确定性的问题，提高预报的精度。此外，对于基于Elman神经网络的单一预报模型，预报精度可以满足要求，但单预报模型之间的预报结果可能大不相同。多模型集合预报可以减少不确定性的影响，提高预测精度。然而，由于原始数据的误差、模型参数的不确定性和不同地区的环境差异，以及实际水文应用的建模要求存在显著差异，因此，需要开发更加可靠和智能的专家系统来进行实时预报[48]。多模型集合预报能够减少预测偏差，如何确定最佳的集合预测模型还有待进一步研究。



**图5预报流程图**

**Author Contributions:** Conceptualization, C.L. and L.Z.; methodology, Z.H. and H.G.; software, C.L. and L.Z.;validation, Y.Y., D.Y. and X.Q.; formal analysis, L.Z.; investigation, C.L.; resources, H.G.; data curation, L.Z.;Writing—original draft preparation, C.L.; writing—review and editing, H.G.; visualization, Z.H.; supervision, C.L.and H.G.; project administration, C.L.; funding acquisition, H.G. and C.L.

**Funding:** This study is supported by National Natural Science Foundation of China (No. 61701166), National Key R&D Program of China (No. 2018YFC1508106), the China Postdoctoral Science Foundation (No. 2018M632215),the Fundamental Research Funds for the Central Universities (No. 2018B16314), Young Elite Scientists Sponsorship Program by CAST (No. 2017QNRC001), National Science Foundation for Young Scientists of China (No. 51709271),Projects in the National Science & Technology Pillar Program during the Twelfth Five-year Plan Period (No.2015BAB07B01)

**Conflicts of Interest:** The authors declare no conflict of interest

参考文献

1. Yaseen, Z.M.; El-Shafie, A.; Jaafar, O.; Afan, H.A.; Sayl, M.N. Artificial intelligence based models for stream-flow forecasting: 2000-2015. J. Hydrol. 2015, 530, 829-844.
2. Duan, Q.Y.; Sorooshian, S.; Gupta, V. Effective and efficient global optimization for conceptual Rainfall-runoff models. Water Resour Res. 1992, 28, 1015-1031.
3. Kneis, D.; Burger, G.; Bronstert, A. Evaluation of medium-range runoff forecasts for a 500 km2 watershed. J. Hydrol. 2012, 414, 341-353.
4. Wang, W.C.; Chau, K.W.; Xu, D.M.; Chen, X.Y. Improving forecasting accuracy of annual runoff.
5. Wang W.; Ma J. Review on some methods for hydrological forecasting. Adv. Sci. Techn. Water Res. 2005, 5, 6-60.
6. Kisi, O.; Nia, A.M.; Goshen, M.G.; Tajabadi, M.R.J.; Ahmadi, A. Intermittent streamflow forecasting by using several data driven techniques. Water Resour. Manag. 2012, 26, 457-474.
7. GE Z.X.; XUE M.; SONG Y.L. Application of multi-factor stepwise regression cycle analysis in medium and long-term hydrological forecast. J. HHU, 2009, 37, 255-257.
8. YAO L.S.; CAI Y.D. Long-range runoff forecast by artificial neural network. Adv. Water Sci. 1995, 6, 61-65.
9. Machado, F.; Mine, M.; Kaviski, E.; Fill, H. Monthly rainfall-runoff modelling using artificial neural networks. Hydrol. Sci. J. 2011, 56, 349-361.
10. Alvaro, L.R.; Vicente, L.F.; David, P.V.; Joaquin, T.P. One-Day-Ahead Streamflow forecasting using artificial neural networks and a meteorological mesoscale model. J. Hydrol. Eng. 2015, 20, 05015001.
11. Kisi O.; Cimen M. A wavelet-support vector machine conjunction model for monthly streamflow forecasting. J. Hydrol. 2011, 399: 132-140.
12. Huang, S.Z.; Chang, J.X.; Huang, Q.; Chen, Y.T. Monthly streamflow prediction using modified EMD-based support vector machine. J. Hydrol. 2014, 511, 764-775.
13. Liu, Z.Y.; Zhou, P.; Chen, G.; Guo, L.D. Evaluating a coupled discrete wavelet transform and support vector regression for daily and monthly streamflow forecasting. J. Hydrol. 2014, 519, 2822-2831.
14. Yang L.; Tian F.Q.; Hu H.P. Modified ESP with information on the atmospheric circulation and teleconnection incorporated and its application. J. Tsinghua Univ. (Sci. &Tech.). 2013, 53, 606-612.
15. Day G.N. Extended streamflow forecasting using NWSRFS. J. Water Resour. Pl. Manag., 1985, 111(2):157-170.
16. Alan F. Hamlet, Dennis P. Lettenmaier. Columbia River streamflow forecasting based on ENSO and PDO climate signals [J]. Journal of Water Resources Planning and Management-ASCE, 1999, 125(6):333-341.
17. K. Lamb, T. Piechota, H. Moradkhani. Improving ensemble streamflow prediction using interdecadal/interannual climate variability [D]. Las Vegas, USA: University of Nevada, 2010.
18. K. Werner, D. Brandon, M. Clark, S. Gangopadhyay. Climate Index Weighting Schemes for NWS ESP-Based Seasonal Volume Forecasts [J].Journal of Hydrometeorology, 2004, 4(26):1076-1090.
19. T. Y. Gan and A. Gobena. Incorporation of seasonal climate forecasts in the ensemble streamflow prediction system [J]. Journal of Hydrology, 2010, 385(1-4):336-352.
20. D. J. Druce. Insights from a history of seasonal inflow forecasting with a conceptual hydrologic model [J].Journal of Hydrology, 2001, 249(1-4):102-112.
21. Hsiang-Te Kung, L. Yu. Lin, S. Malasri. Use of Artificial Neural Networks in Precipitation Forecasting. Proceedings of international symposium, on impact of climatic change global warming on hydrology and water resources, China, 1993, 72-82.
22. M L Zhu, M Fujita, and N. Hashimoto. Application of neural networks to runoff prediction. International conference on stochastic and statistical methods in hydrology and environmental engineering. Ontario. Canada.1993, 205-216.
23. CHEN Shou-yu, WANG Da-gang. Genetic algorithms-based fuzzy optimization BP neural network model and its application [J]. Journal of Hydraulic Engineering, 2003, 5(1):116-120.
24. B.H.M. Sadeghi. A BP-neural network predictor model for plastic injection molding process [J]. Journal of Materials Processing Technology, 2000, 103(3):411-416.
25. M. Shoaib, A.Y. Shamseldin, B.W. Melville, M.M. Khan. A comparison between wavelet based static and dynamic neural network approaches for runoff prediction [J]. Journal of Hydrology, 2016, 535(1):211-225.
26. GUO Rui-li, ZHAO Fen-yao, LI Yan-an. Dynamic modeling of rainfall runoff process in river basin with recurrent wavelet neural network [J]. Journal of Hydroelectric Engineering. 2013, 32(2):55-59.
27. SONG Hai-liang, SONG Yang, ZHANG Deng-hua, ZHONG Wei. Research on runoff forecasting of reservoir by using fuzzy clustering analysis and wavelet neural networks [J]. Journal of Hydroelectric Engineering. 2008, 27(1):6-10.
28. ZHANG Xiao-wei, SHEN Bing, HUANG Ling-mei. Grey self-memory model based on BP neural network for annual runoff prediction [J]. Journal of Hydroelectric Engineering, 2009, 28(1):69-77.
29. Elman, Jeffrey L. (1990). "Finding Structure in Time". Cognitive Science. 14 (2): 179–211.
30. Chiang Y.M.; Chang L.C.; Chang F.J.; Comparison of static-feedforward and dynamic-feedback neural networks for rainfall-runoff modeling. J. Hydrol. 2004, 290, 297-311.
31. Cruse, Holk; Neural Networks as Cybernetic Systems, 2nd and revised edition.
32. WANG Chang-hong, GAO Xiao-zhi, XU Li-xin, ZHANG Xian-yi. A new modified Elman neural network model [J]. Journal of Electronics. 1997, 19(6):793-743.
33. GB/T 22482-2008, Standard for hydrological information and hydrological forecasting [S].Beijing: China Standard Press, 2008:4-10.

H:\documents\layout\new template June 2014\figures\CC-BY logo original v1.wmf© 2019 by the authors. Submitted for possible open access publication under the   
terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).