第 52 卷 第 6 期：2526-2536 高电压技术 Vol.52, No.6: 2526-2536
2026 年 6 月 30 日 High Voltage Engineering June 30, 2026
**DOI**: 10.13336/j.1003-6520.hve.20251542

# 基于 SA-GRU 模型的变电站 GIS 设备 LCC 数据预测

唐　波[1,2]，尚智宇[1,2]，吴遂宜[3]，张龙斌[1]，王　帅[1]
（1. 三峡大学电气与新能源学院，宜昌 443002；2. 湖北省输电线路工程技术研究中心，宜昌 443002；
3. 国网安徽省电力有限公司庐江县供电公司，庐江 231500）

**摘　要**：实现变电站气体绝缘开关(gas insulated switchgear, GIS)设备全寿命周期成本(life cycle cost，LCC)数据的准确预测，是解决 GIS 设备经济比选问题的关键。针对现有变电站 GIS 设备 LCC 数据样本少和数据异常的问题，引入相似变电站 GIS 设备 LCC 数据并结合自注意力机制(self attention, SA)算法以扩充样本量，采用门控循环单元(gated recurrent unit, GRU)模型预测异常时序数据，提出了一种基于 SA-GRU 模型的变电站 GIS 设备 LCC 数据预测方法。该方法考虑到相似变电站 GIS 设备的 LCC 数据与已有样本的数据特征存在差异，利用自注意力机制算法捕捉数据内部规律的能力以强化数据特征，并依此量化相似数据与样本数据间的影响程度权重，使得二者的数据特征保持一致；再利用 GRU 模型的时间连续特性减小异常数据对整体 LCC 数据变化趋势的影响，同时解决训练中的梯度消失爆炸问题，最终预测出全寿命周期内的 LCC 数据。以中部省某变电站 GIS 设备 LCC 数据为例，所提方法在均方根误差的准确度上相较于 Attention-GRU 模型提高了 31.11%，验证了本文所提方法的准确性。
**关键词**：变电站 GIS 设备；LCC 数据；少样本；预测模型；自注意力机制；GRU 模型

## LCC Data Prediction for Substation GIS Equipment Based on SA-GRU Model

TANG Bo[1,2], SHANG Zhiyu[1,2], WU Suiyi[3], ZHANG Longbin[1], WANG Shuai[1]
(1. College of Electrical Engineering and New Energy, China Three Gorges University, Yichang 443002, China;
2. Hubei Provincial Engineering Technology Research Center for Power Transmission Line, Yichang 443002, China;
3. Lujiang County Power Supply Company, State Grid Anhui Electric Power Company, Lujiang 231500, China)

**Abstract**: The key to solve the problem of economic comparison and selection of gas insulated switchgear (GIS) equipment is to realize the accurate prediction of life cycle cost (LCC) data of substation GIS equipment. In accordance with the problem of few LCC data samples and abnormal data of existing substation GIS equipment, LCC data of similar substation GIS equipment is introduced and combined with self-attention (SA) mechanism to expand the sample scale, and then GRU model is used to predict abnormal time series data. A LCC data prediction method for substation GIS equipment based on SA-GRU model is proposed. Considering the differences of data characteristics between the similar data and the sample data, the self-attention mechanism algorithm is used to strengthen the data characteristics, and the weight of the influence degree between the similar data and the sample data is quantified, so that the data characteristics of the two are consistent. Then the GRU model is used to solve the problem of gradient disappearance explosion in training, and finally the LCC data in the whole life cycle are predicted. The LCC data of GIS equipment in a substation in a certain province in central China are taken as an example to verify the method. The proposed method demonstrates an improvement of 31.11% in accuracy based on the root mean square error compared to the Attention-GRU model, validating the accuracy of the method presented in this paper.
**Key words**: substation GIS equipment; LCC data; fewer samples; prediction model; self-attention mechanism; GRU model

## 0　引言

作为变电站气体绝缘开关(gas insulated switchgear, GIS)设备经济性比选的核心指标，全寿命周期成本(life cycle cost, LCC)直接决定了变电站 GIS 设备的选型[1-2]。但实际工程中，通常要求变电站 GIS 设备在全寿命周期结束前，基于已有 LCC 数据计算其结果。然而，当前我国新建、技改变电站 GIS 设备 LCC 数据的收集工作开展较晚，收集

**基金资助项目**：国家电网有限公司科技项目(SGHAYJ00SJJS2200061)。
Project supported by Science and Technology Project of SGCC (SGHAYJ00SJJS2200061).

到的 LCC 数据通常存在数据样本少和数据异常的问题，导致无法实现 LCC 数据的准确预测[3-4]。因此，寻求一种适用于变电站 GIS 设备 LCC 数据的预测方法，是现有条件下解决 GIS 设备经济比选问题的关键[5]。

最早用于 GIS 等电气设备 LCC 数据预测的算法是灰色预测法。如文献[6]试图通过生成拟合函数的方式对主设备的 LCC 数据进行预测，但其预测结果近似于指数增长趋势，与设备 LCC 理论上应该呈现的二次函数曲线趋势不符[7]。后续，文献[8]引入最小二乘算法，通过最小化平方和的方法推导与设备 LCC 数据分布趋势匹配的最佳拟合函数。然而，当自变量间存在多重共线性时，可能产生不可逆矩阵，会导致预测的准确度下降。在此基础上，文献[9]提出了优化的偏最小二乘法，通过提取自变量和因变量间正交潜在变量的方式消除自变量间的共线性。但这种方法基于最大化协方差的方式提取潜在变量，易受缺失异常数据等因素的影响，导致预测产生的数据误差较大[10]。不过，文献[11]在解决变电站整站方案的 LCC 数据预测问题时，提出了一种基于模糊神经网络的数据预测方法，降低了缺失异常数据对预测结果的影响。因此，可以考虑将神经网络算法引入 GIS 设备 LCC 数据的预测问题中。然而，GIS 设备 LCC 数据属于时序数据[12]，直接采用神经网络的数据处理方法，在网络训练更新过程中容易受参数不稳定的影响，出现梯度消失或爆炸问题[13]。需要注意的是，随着新能源的发展，类似于 GIS 设备 LCC 时序数据的风电功率数据预测的研究已较为成熟[14-15]。如文献[16]在神经网络中引入门控循环单元(gated recurrent unit, GRU)模型，通过 GRU 模型中重置门和更新门的控制作用增强时序数据间的依赖关系。但该模型需大量数据进行训练，在数据样本少的应用情况下，无法保证预测准确性[17]。为此，文献[18]结合注意力机制对 GRU 模型进行改进，通过注意力机制量化已有风电场数据特征与相似风电场数据之间的影响程度，再将加权后的相似风电场数据作为训练数据进行 GRU 预测。但该方法获得的数据特征由人为提取，当数据分布规律复杂时，提取难度较大，主观性较强，导致预测结果的准确性依旧不足[19]。

为此，本文针对变电站 GIS 设备 LCC 数据为时序数据的特点，将 GRU 模型应用到 GIS 设备 LCC 数据的预测问题中，同时为解决已有样本数据少导致的数据预测精度低问题，采用自注意力机制(self attention，SA)自动提取已有样本数据特征，根据已有样本数据特征量化已有样本数据与相似数据的影响程度，再对应影响程度赋予这些数据不同的权重，从而提出了一种 SA-GRU 模型的变电站 GIS 设备 LCC 数据预测方法。该方法被成功应用在国内中部某省某 220 kV 变电站 GIS 设备的技改大修工程中，解决了该变电站 110 kV GIS 设备 LCC 数据预测问题。

## 1　变电站 GIS 设备 LCC 数据及预测方法

### 1.1　变电站 GIS 设备的 LCC 数据

GIS 设备的经济性通常根据该设备的 LCC 进行判断。如图 1 所示，GIS 设备的 LCC 由初始投资成本、退役成本、运行成本、检修成本和故障成本这几个不同种类的成本构成。其中，初始投资成本和退役成本在变电站建成时可以基本确定，属于固定成本，但运行、检修和故障成本需要每年进行统计，为年变成本。

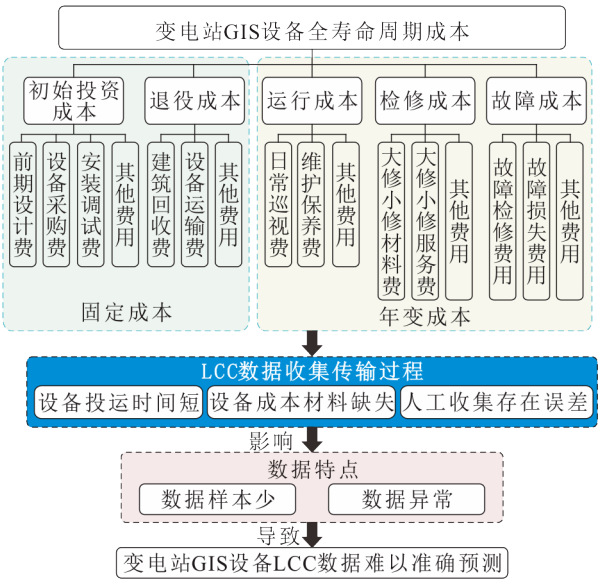



图 1　变电站 GIS 设备的 LCC 数据及问题

Fig.1　Problem and LCC data in substation GIS equipment

从年变成本的实际数据收集情况来看，由于变电站 GIS 设备 21 世纪以来才开始在国内大范围地投入使用，且相关的数据收集工作开展较晚，现收集到不同年份的 LCC 数据数量较少，导致整个 LCC 数据样本偏少。另外，在早期 LCC 数据的收集和传输过程中，通常依赖人工记录的方式，因此会出现成本材料不足、丢失和错误记录等情况，这些情况将导致收集的 LCC 样本数据中存在异常数据。

综上，这些数据问题给 LCC 数据的准确测算带来了极大困扰，需要根据变电站 GIS 设备 LCC 数据的特点找到一种合适的数据预测方法，以满足在扩充数据样本规模的同时，降低异常数据对预测的影响。

### 1.2 现有数据预测方法的不足和改进思路

由 1.1 节可知，为获取准确预测变电站 GIS 设备 LCC 数据，首先应尽量扩充样本数量规模，解决数据样本少的问题；然后，分析数据样本特征，解决数据异常问题。

当前，在样本数据具有一定变化规律的条件下，有效解决少样本数据预测问题的方法主要是，引入与已有样本数据同规模的相似数据从而扩充原有样本数据的规模[20-22]。但考虑已有样本数据与相似数据的数据特征并不能保证完全一致，因此，采用具有聚焦式特性的注意力机制的方法量化已有样本数据与相似数据的影响程度权重，将赋权后的数据作为扩充的样本数据[18]。但注意力机制需要人为给定数据特征的向量，从而围绕该向量进行处理分析。考虑到变电站 GIS 设备 LCC 数据属于成本数据，其产生受设备实际运行情况影响大，数据分布规律复杂，数据特征不明显，人为提取数据特征缺乏依据和准确性。因此，直接采用注意力机制方法获得的数据间影响程度权重的准确性较差。那么，如果能够通过计算数据间向量权重的方式，自动提取数据特征的向量，增强数据特征提取结果的可靠性，则可以避开人为提取特征向量导致的主观问题，为已有样本数据与相似数据的影响程度权重量化提供客观依据，从而提高数据预测的准确性。

而对于异常数据的问题，在进行基于数据分布特点的数据预处理后，考虑到 GIS 设备 LCC 数据按照产生时间的先后顺序进行收集排列，属于时序数据，因此采用具有时序关系捕捉能力的循环神经网络(recurrent neural network, RNN)进行时序数据的预测[23]。RNN 在处理 LCC 等时序数据的过程中，根据反向传播算法和链式法则[13]，计算调整神经网络参数的梯度，使得损失函数取到极小值，从而降低预测结果的总体误差。但 RNN 中各参数的梯度表达式均为链式法则线性连乘公式，这导致计算结果经常会随着时间跨度的增加成指数般收敛于 0 或趋向于无穷，从而致使 LCC 数据的梯度出现梯度消失或爆炸现象，导致预测结果准确性存疑。RNN 中产生梯度消失或爆炸的根本原因在于其信息传递结构的线性关系。为此，若能优化网络的内部结构，使其有选择性地保留长时间跨度的信息，则可以增强数据信息间的非线性关联程度，抑制网络训练更新过程中出现的梯度消失或爆炸问题。

综上，为获得准确的变电站 GIS 设备 LCC 数据预测结果，首先可通过引入自注意力机制，强化已有样本数据特征，量化已有样本数据与相似数据的影响程度权重，以此扩充样本数据规模；然后采用 GRU 模型，利用其门控结构抑制梯度消失或爆炸问题，使得数据预测结果的准确性得到进一步提升，实现 LCC 数据的准确预测。

## 2 用于 LCC 数据预测的 SA-GRU 模型

### 2.1 基于自注意力机制的样本规模扩充

根据已有样本数据数量较少的特点，可引入相同电压等级、相近环境的多个同规模变电站 GIS 设备 LCC 相似数据，以扩充样本数据规模。由于相似数据并不能完全代表已有样本数据特征，若将相似数据直接加入已有样本数据中进行成本预测，显然无法实现 LCC 数据的精确测算。因此，需要根据已有样本数据特征对相似数据进行权重处理，解决相似数据与已有样本数据特征不完全一致情况而导致的数据预测准确性降低问题。

由于变电站 GIS 设备 LCC 数据分布规律复杂，为寻找数据间的相关性以进行权重处理，需要将已有样本数据与相似数据转换成矩阵的形式。将不同年份作为数据的行标，不同种类作为矩阵的列标，依此构建已有样本数据的矩阵表达式 $\boldsymbol{H}_{\mathrm{LCC}}$ 为：

$$\boldsymbol{H}_{\mathrm{LCC}}=\begin{bmatrix} h_{11} & \cdots & h_{1n} \\ \vdots & & \vdots \\ h_{t1} & \cdots & h_{tn} \end{bmatrix} \tag{1}$$

式中：$h_{tn}$ 为 GIS 设备的第 $t$ 年第 $n$ 种 LCC 数据；$n$=1, 2, 3, …, $N$；$t$=1, 2, 3, …, $T$。

考虑到相似数据包括多个同规模变电站 GIS 设备 LCC 数据，以不同变电站为矩阵列分块依据，每块子矩阵表达式的构建方式同已有样本数据矩阵表达式保持一致。为便于进行不同数据间的相关性比较，将已有样本数据作为参照值并加入相似数据，从而相似数据的矩阵式可写成如下形式：

$$\boldsymbol{L}_{\mathrm{LCC}}=[\boldsymbol{H}_{\mathrm{LCC}},\boldsymbol{H}^{1}_{\mathrm{LCC}},\boldsymbol{H}^{2}_{\mathrm{LCC}},\cdots,\boldsymbol{H}^{m-1}_{\mathrm{LCC}},\boldsymbol{H}^{m}_{\mathrm{LCC}}] \tag{2}$$

式中：$\boldsymbol{L}_{\mathrm{LCC}}$ 为构建的 LCC 数据相似矩阵；$\boldsymbol{H}^{m}_{\mathrm{LCC}}$ 为第 $m$ 个同规模变电站 GIS 设备 LCC 数据构成的相似数据子矩阵，$\boldsymbol{H}^{m}_{\mathrm{LCC}}$ 的行列数与已有样本数据矩阵一致。

由式(1)和式(2)的矩阵构成可知，已有样本数据与相似数据的相关性寻找问题，已转变为矩阵列向量间的相关性求解问题。根据注意力机制的求解算法[24-25]，矩阵列向量间的相关性可通过注意力打分函数进行求解，即：

$$\boldsymbol{\alpha}_m = N\left(\mathrm{s}\left(\boldsymbol{H}_{\mathrm{LCC}}^{m}, \boldsymbol{H}_{\mathrm{LCC}}\right)\right) \tag{3}$$

式中：$\boldsymbol{\alpha}_m$ 为第 $m$ 个相似数据子矩阵与已有样本数据矩阵间的权重系数矩阵，即矩阵间的相关性；$N(\cdot)$ 为归一化函数；$s(\cdot)$为注意力打分函数，考虑 LCC 数据的时序较长，且点积模型适用矩阵间的向量权重系数计算，故式(3)中注意力打分函数使用缩放点积模型。

在利用打分函数进行矩阵相关性求解的过程中，由于注意力打分函数直接利用矩阵元素进行运算，且 GIS 设备 LCC 数据具有数据特征不明显的特点，因此注意力打分函数的打分结果不准确。为此，引入自注意力机制的算法[24]求解 $\boldsymbol{H}_{\mathrm{LCC}}$ 矩阵元素间的权重关系，以加强数据特征的关联。自注意力机制在求解权重关系时，为更好捕捉矩阵内任意两元素的相关性，需通过线性映射的方式将 $\boldsymbol{H}_{\mathrm{LCC}}$ 矩阵转化为 3 个矩阵：

$$\begin{cases} \boldsymbol{Q}_{\mathrm{LCC}} = \boldsymbol{W}_{\mathrm{q}}^{\mathrm{LCC}} \boldsymbol{H}_{\mathrm{LCC}} \\ \boldsymbol{K}_{\mathrm{LCC}} = \boldsymbol{W}_{\mathrm{k}}^{\mathrm{LCC}} \boldsymbol{H}_{\mathrm{LCC}} \\ \boldsymbol{V}_{\mathrm{LCC}} = \boldsymbol{W}_{\mathrm{v}}^{\mathrm{LCC}} \boldsymbol{H}_{\mathrm{LCC}} \end{cases} \tag{4}$$

式中：$\boldsymbol{Q}_{\mathrm{LCC}}$、$\boldsymbol{K}_{\mathrm{LCC}}$、$\boldsymbol{V}_{\mathrm{LCC}}$ 分别为线性映射后得到的 LCC 数据查询向量矩阵、LCC 数据键值向量矩阵和 LCC 数据表征向量矩阵；$\boldsymbol{W}_{\mathrm{q}}$、$\boldsymbol{W}_{\mathrm{k}}$ 和 $\boldsymbol{W}_{\mathrm{v}}$ 分别为矩阵 $\boldsymbol{Q}_{\mathrm{LCC}}$、矩阵 $\boldsymbol{K}_{\mathrm{LCC}}$ 和矩阵 $\boldsymbol{V}_{\mathrm{LCC}}$ 的参数矩阵，在自注意力机制线性映射过程中由参数学习得到。

其中，矩阵 $\boldsymbol{Q}_{\mathrm{LCC}}$ 和 $\boldsymbol{K}_{\mathrm{LCC}}$ 的维度相同，用于计算矩阵 $\boldsymbol{H}_{\mathrm{LCC}}$ 列向量间的权重系数，而矩阵 $\boldsymbol{V}_{\mathrm{LCC}}$ 用于与权重系数的结果进行计算以获得矩阵输出。此时，矩阵 $\boldsymbol{Q}_{\mathrm{LCC}}$ 与矩阵 $\boldsymbol{K}_{\mathrm{LCC}}$ 间的权重系数即代表 $\boldsymbol{H}_{\mathrm{LCC}}$ 矩阵元素间的权重系数，该权重系数同样可通过打分函数进行求解：

$$\boldsymbol{A}_{\mathrm{LCC}} = N\left(s\left(\boldsymbol{K}_{\mathrm{LCC}}, \boldsymbol{Q}_{\mathrm{LCC}}\right)\right) \tag{5}$$

式中：$\boldsymbol{A}_{\mathrm{LCC}}$ 为已有样本数据的权重系数矩阵；由于矩阵 $\boldsymbol{Q}_{\mathrm{LCC}}$ 与矩阵 $\boldsymbol{K}_{\mathrm{LCC}}$ 也属于时序数据，且同为矩阵间的向量权重系数计算，故式(5)中同样使用缩放点积公式。

由于式(5)得到的权重系数矩阵 $\boldsymbol{A}_{\mathrm{LCC}}$ 仅能代表矩阵元素间的相关性，若要得到数据特征加强后的矩阵，还需对矩阵 $\boldsymbol{A}_{\mathrm{LCC}}$ 与矩阵 $\boldsymbol{V}_{\mathrm{LCC}}$ 进行矩阵相乘，从而得到数据特征矩阵：

$$\boldsymbol{H}'_{\mathrm{LCC}} = \boldsymbol{V}_{\mathrm{LCC}} \boldsymbol{A}_{\mathrm{LCC}} \tag{6}$$

式中：$\boldsymbol{H}'_{\mathrm{LCC}}$ 为数据特征矩阵，矩阵的行列同矩阵 $\boldsymbol{H}_{\mathrm{LCC}}$ 一致。

进一步，将特征强化后的数据特征矩阵 $\boldsymbol{H}'_{\mathrm{LCC}}$ 代入式(3)中进行求解，获得第 $m$ 个相似数据子矩阵与数据特征矩阵间的权重系数矩阵 $\boldsymbol{\alpha}'_m$。在此基础上，为降低矩阵 $\boldsymbol{H}'_{\mathrm{LCC}}$ 与矩阵 $\boldsymbol{H}_{\mathrm{LCC}}^{m}$ 数据特征不完全一致对预测准确度的影响，将矩阵 $\boldsymbol{\alpha}'_m$ 与矩阵 $\boldsymbol{H}_{\mathrm{LCC}}^{m}$ 相乘，由此可得到扩充样本数据后的变电站 GIS 设备 LCC 数据矩阵：

$$\boldsymbol{Y}_{\mathrm{LCC}} = [\boldsymbol{\alpha}'_1 \boldsymbol{H}_{\mathrm{LCC}}^{1}, \cdots, \boldsymbol{\alpha}'_m \boldsymbol{H}_{\mathrm{LCC}}^{m}] \tag{7}$$

式中：$\boldsymbol{Y}_{\mathrm{LCC}}$ 为扩充样本数据矩阵。

至此，最终构建出如图 2 所示的基于自注意力机制的 LCC 样本数据扩充模型。

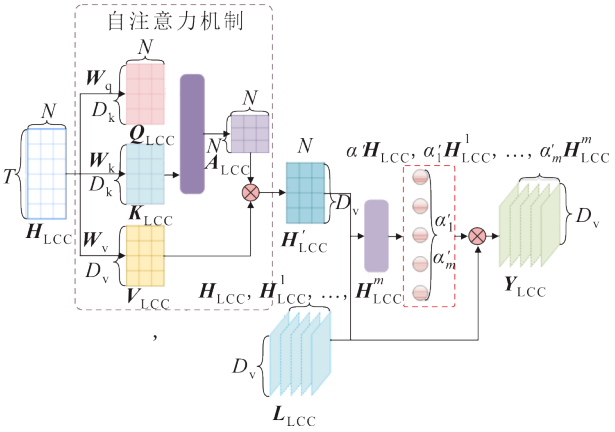



图 2　基于自注意力机制的 LCC 样本数据扩充模型

Fig.2　LCC sample data augmentation model based on self-attention mechanism algorithm

## 2.2　基于 GRU 的数据预测

由于 LCC 数据中存在异常数据，若直接使用扩充后的样本数据进行预测，仍会对数据预测的准确性造成影响。因此，为从根源上提高数据预测的准确性，还需对 LCC 数据中的异常数据进行处理。由于 LCC 数据中的异常值仅出现在时序中的某一时刻，属于点异常数据[4]，为削弱原时序序列中点异常数据的影响，采用加权移动平均法对扩充样本数据进行处理：

$$\begin{cases} \boldsymbol{Y}_{\mathrm{LCC}}^{t+1} = \dfrac{\sum_{i=1}^{t} w_i \boldsymbol{Y}_{\mathrm{LCC}}^{i}}{\sum_{i=1}^{t} w_i} \\ \boldsymbol{Y}'_{\mathrm{LCC}} = [\boldsymbol{Y}_{\mathrm{LCC}}^{1}, \cdots, \boldsymbol{Y}_{\mathrm{LCC}}^{T}]^{\mathrm{T}} \end{cases} \tag{8}$$

式中：$\boldsymbol{Y}_{\mathrm{LCC}}^{i}$ 为第 $i$ 年的扩充样本数据值；$w_i$ 为第 $i$ 年的数据权重；$i\in[1,T]$；$T$ 为采用加权移动平均法进行数据处理的年份；$\boldsymbol{Y}'_{\mathrm{LCC}}$ 为处理后的扩充样本数据矩阵。

在此基础上，考虑到 RNN 算法的循环结构可以很好地捕捉时序数据整体变化趋势，故引入 RNN 算法对处理后的扩充样本数据进行预测[21]。将矩阵 $\boldsymbol{Y}'_{\mathrm{LCC}}$ 作为数据输入，通过 RNN 算法的梯度计算过程对输入数据进行网络训练和预测：

$$\frac{\partial L_t}{\partial V}=\sum_{k=0}^{t}\frac{\partial L_k}{\partial y_k}\frac{\partial y_k}{\partial h_k}\left(\prod_{j=k-1}^{t}\frac{\partial h_j}{\partial h_{j-1}}\right)\frac{\partial h_k}{\partial V}=\sum_{k=0}^{t}\frac{\partial L_k}{\partial y_k}\frac{\partial y_k}{\partial h_k}\left(\prod_{j=k-1}^{t}\frac{\partial\left(\tanh(Wx_t+Uh_{t-1}+b)\right)}{\partial h_{t\text{-}1}}\cdot U\right)\frac{\partial h_k}{\partial V}\tag{9}$$

式中：$L_t$ 为 RNN 网络 $t$ 时刻损失函数；$y_k$ 为 RNN 网络 $k$ 时刻的输出；$h_j$ 为 RNN 网络隐藏层 $j$ 时刻的隐状态；tanh 为 RNN 隐藏层的激活函数；$U$ 为 RNN 隐藏层之间的权重参数；$b$ 为偏置项；$V$ 为 RNN 隐藏层到其他层的权重参数；$W$ 为隐藏层对输入量 $x$ 的权重参数。

由式(9)中存在的线性连乘关系可知，若 $\partial\left(\tanh(Wx_t+Uh_{t-1}+b)\right)U/\partial h_{t-1}<1$，随着输入的 LCC 数据时间跨度增加，会导致梯度成指数般收敛于 0；若 $\partial\left(\tanh(Wx_t+Uh_{t-1}+b)\right)U/\partial h_{t-1}>1$，则随着时间跨度增加则梯度会趋向于无穷。因此，在进行长时间跨度的梯度计算时，容易出现梯度消失或爆炸现象，导致预测结果的精度下降。针对该问题，考虑到变电站 GIS 设备 LCC 数据不属于大样本数据集，为便于计算，在 RNN 算法的结构基础上增加 GRU 门控结构，以解决因线性连乘关系导致的梯度消失爆炸问题。构成的新网络结构如图 3 所示。

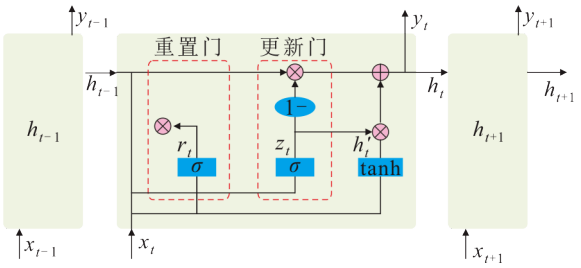



图 3 增加 GRU 门控结构的网络示意图

Fig.3 Schematic diagram of the network with the GRU gating structure

GRU 门控结构的网络结构表达式为：

$$\begin{cases}r_t=\sigma(W_{\mathrm{r}}x_t+U_{\mathrm{r}}h_{t-1}+b_{\mathrm{r}})\\ z_t=\sigma(W_{\mathrm{z}}x_t+U_{\mathrm{z}}h_{t-1}+b_{\mathrm{z}})\\ h'_t=\tanh(W_{\mathrm{h}}x_t+U_{\mathrm{h}}(h_{t-1}\odot r_t)+b_{\mathrm{h}})\\ h_t=(1-z_t)\odot h_{t-1}+z_t\odot h'_t\end{cases}\tag{10}$$

式中：$\boldsymbol{r}_t$ 为 $t$ 时刻的重置门；$\boldsymbol{z}_t$ 为 $t$ 时刻的更新门；$\boldsymbol{h}'_t$ 为 $t$ 时刻的候选状态信息；$\boldsymbol{x}_t$ 为 $t$ 时刻的输入信息；$\boldsymbol{h}_t$ 为传递给 $t$+1 时刻的隐状态信息；$\boldsymbol{h}_{t-1}$ 为 $t$–1 时刻的刻隐状态信息；$\sigma$ 为 sigmoid 激活函数；tanh 为激活函数；$\odot$是 Hadamard Product 算法，即矩阵对应元素相乘；$\boldsymbol{W}_{\mathrm{r}}$、$\boldsymbol{W}_{\mathrm{z}}$、$\boldsymbol{W}_{\mathrm{h}}$、$\boldsymbol{U}_{\mathrm{r}}$、$\boldsymbol{U}_{\mathrm{z}}$、$\boldsymbol{U}_{\mathrm{h}}$ 分别为重置门、更新门和候选状态对应的权值参数矩阵；$\boldsymbol{b}_{\mathrm{r}}$，$\boldsymbol{b}_{\mathrm{z}}$，$\boldsymbol{b}_{\mathrm{h}}$ 分别为重置门、更新门和候选状态的偏置参数矩阵。

由此，GRU 结构中梯度计算时的连乘关系表达式转变为：

$$\begin{cases}\displaystyle\prod_{j=k-1}^{t}\frac{\partial \boldsymbol{h}_j}{\partial \boldsymbol{h}_{j-1}}=\prod_{j=k-1}^{t}\left(\left(1\text{-}\boldsymbol{z}_j\right)+\left(\boldsymbol{U}_{\mathrm{z}}\boldsymbol{\delta}_{\boldsymbol{z}_j}+\boldsymbol{U}_{\mathrm{h}}\left(\boldsymbol{\delta}_{\boldsymbol{h}_j}\boldsymbol{r}_j\right)\right)+\left(\boldsymbol{U}_{\mathrm{r}}\boldsymbol{\delta}_{\boldsymbol{r}_j}\right)\right)\\ \boldsymbol{\delta}_{\boldsymbol{z}_j}=\dfrac{\partial \boldsymbol{L}_j}{\partial \boldsymbol{z}_j}=\dfrac{\partial \boldsymbol{L}_j}{\partial \boldsymbol{h}_j}\left[\boldsymbol{h}_j-\boldsymbol{h}_{j-1}\right]\boldsymbol{z}_j(1-\boldsymbol{z}_j)\\ \boldsymbol{\delta}_{\boldsymbol{r}_j}=\dfrac{\partial \boldsymbol{L}_j}{\partial \boldsymbol{r}_j}=\boldsymbol{\delta}_{\boldsymbol{h}_j}\left(\boldsymbol{U}_{\mathrm{h}}\boldsymbol{h}_{j-1}\right)\boldsymbol{r}_j(1-\boldsymbol{r}_j)\\ \boldsymbol{\delta}_{\boldsymbol{h}_j}=\dfrac{\partial \boldsymbol{L}_j}{\partial \boldsymbol{h}_j}=\dfrac{\partial \boldsymbol{L}_j}{\partial \boldsymbol{h}_j}\boldsymbol{z}_j\left(1-\boldsymbol{h}_j^{\,2}\right)\end{cases}\tag{11}$$

式中：$\boldsymbol{\delta}_{\boldsymbol{z}j}$ 为第 $j$ 时刻损失函数 $\boldsymbol{L}$ 对更新门 $\boldsymbol{z}$ 的偏导；$\boldsymbol{\delta}_{\boldsymbol{r}j}$ 为第 $j$ 时刻损失函数 $\boldsymbol{L}$ 对重置门 $\boldsymbol{r}$ 的偏导；$\boldsymbol{\delta}_{\boldsymbol{h}_j}$ 为第 $j$ 时刻损失函数 $\boldsymbol{L}$ 对隐状态信息 $\boldsymbol{h}$ 的偏导。

将式(11)中的连乘代入式(9)，可以看出，原来式(9)中连乘的线性关系式已被非线性关系式代替，因此数据预测中产生的梯度消失爆炸问题得以解决。此时，再将处理后的扩充样本数据矩阵作为网络的输入，通过 GRU 网络进行数据的预测。

得到 GRU 网络的预测值后，选取时间序列预测中最为常用的均方根误差 $E_{\mathrm{r}}$、平均绝对误差 $E_{\mathrm{ma}}$ 以及决定系数 $R^2$ 指标进行误差校验：

$$E_{\mathrm{r}}=\sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(y_i-\hat{y}_i\right)^2}\tag{12}$$

$$E_{\mathrm{ma}}=\frac{1}{n}\sum_{i=1}^{n}\left|\frac{y_i-\hat{y}_i}{y_i}\right|\times100\%\tag{13}$$

$$R^2=1-\frac{\displaystyle\sum_{i=1}^{n}\left(y_i-\hat{y}_i\right)^2}{\displaystyle\sum_{i=1}^{n}\left(y_i-\bar{y}_i\right)^2}\tag{14}$$

式中：$y_i$ 为样本真实值；$\hat{y}_i$ 为预测值；$n$ 为样本真

实值的数量；$\bar{y}_i$为预测值的平均值。

### 2.3　数据预测流程

结合上述对自注意力机制和 GRU 的分析，首先基于自注意力机制对相似数据进行权重处理，得到扩充样本数据；然后利用扩充样本数据 GRU 模型训练模型参数，最后输出 LCC 数据预测结果并进行数据预测结果的准确性校验。完整的 LCC 数据预测流程如图 4 所示。

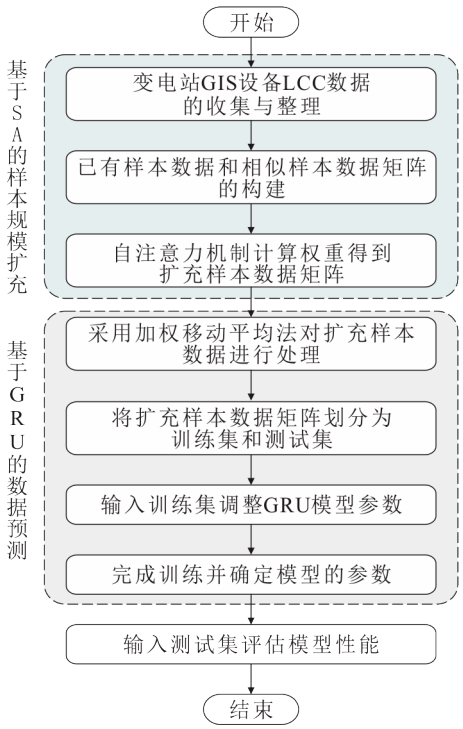



图 4　LCC 数据预测流程图

Fig.4　LCC data prediction flow chart

## 3　算例验证与分析

### 3.1　变电站实际情况

以如图 5 所示的中部省某 220 kV 变电站中 110 kV 侧的 GIS 设备为例。选取该站为样本变电站，将本文所提出的 SA-GRU 模型应用于样本变电站的 GIS 设备 LCC 数据预测。

该变电站现有主变压器 2 台，110 kV 侧配电装置采用户外 GIS 设备，地处 E 级污秽区，样本变电站的电压等级、环境和规模等基本信息如表 1 所示。

为证明本文所提出的 SA-GRU 模型适用于变电站 GIS 设备 LCC 数据的预测，根据实例变电站的电压等级、电气规模和周围环境等基本条件，引入 3 座相似变电站进行 LCC 数据样本的扩充。引入的相似变电站基本信息如表 2 所示。

对比表 1 和表 2 的变电站基本信息可知，引入



图 5　中部省某 220 kV 变电站内 110 kV 侧 GIS 设备

Fig.5　GIS equipment on the 110kV side of a 220kV substation in Henan Province of China

表 1　样本变电站基本信息

Table 1　Basic information of the sample substation

| 项目 | 数值 |
|---|---|
| 电压等级/kV | 220 |
| 主变台数/台 | 2 |
| 主变容量/MVA | 180 |
| 110 kV 进出线/回 | 12 |
| 110 kV 侧 GIS 设备布置方式 | 户外 |
| 海拔高度/m | <1 000 |
| 污秽等级 | E 级 |

表 2　相似变电站基本信息

Table 2　Basic information about similar substations

| 项目 | 变电站 1 | 变电站 2 | 变电站 3 |
|---|---|---|---|
| 电压等级/kV | 220 | 220 | 220 |
| 主变台数 | 3 | 3 | 2 |
| 主变容量/MVA | 240 | 180 | 180 |
| 110 kV 进出线/回 | 12 | 16 | 10 |
| 110 kV 侧 GIS 设备布置方式 | 户外 | 户外 | 户外 |
| 海拔高度/m | <1 000 | <1 000 | <1 000 |
| 污秽等级 | E 级 | E 级 | E 级 |

的 3 个相似变电站电压等级与样本变电站相同，海拔均小于 1 000 m，地处 E 级污秽区，且 GIS 设备均为户外布置。变电站 1 和变电站 2 的主变台数比样本变电站多 1 台，变电站 3 的主变台数同样本变电站的一致。在进出线回数方面，变电站 1 的回数与样本变电站一样，变电站 2 的回数多出 4 回，变电站 3 的回数少了 2 回。其中，变电站 3 和样本变电站的基本信息相似度最高，变电站 2 和变电站 1 的基本信息相似度较低。

确定样本变电站和相似变电站后，通过 PMS、ERP 系统平台和多维精益报表获取 LCC 原始数据，

主要收集电压等级、运行年限、资产原值、日常运维检修材料费、大修材料费、检修费、人工费等成本相关数据，共收集到 GIS 设备的 LCC 样本数据 886 条，引入 3 座相似变电站 GIS 设备的 LCC 样本数据共 4 101 条，进行 LCC 数据的样本扩充。收集整理得到部分的 110 kV GIS 设备成本相关数据样本如表 3 所示。

表 3 110 kV GIS 设备成本相关数据样本(部分)

Table 3 110 kV GIS equipment cost related data sample (partial)

| 序号 | 运行年限/a | 资产原值/元 | 运维检修材料费/元 | 大修材料费/元 | 检修费/元 | 人工费/元 |
|---|---|---|---|---|---|---|
| 1 | 38 | 143 435 | 414 | 1 | 2 959 | 2 632 |
| 2 | 32 | 359 302 | 2 101 | 10 676 | 589 | 35 819 |
| 3 | 18 | 1 557 504 | 1 515 | 1 515 | 3 680 | 15 062 |
| 4 | 23 | 363 904 | 222 | 0 | 4 864 | 5 059 |
| 5 | 4 | 171 930 | 708 | 0 | 0 | 4 484 |
| 6 | 12 | 873 264 | 135 | 0 | 1 140 | 2 585 |
| 7 | 2 | 124 368 | 7 468 | 5 193 | 7 509 | 1 857 |
| 8 | 9 | 645 125 | 8 415 | 9 456 | 1 180 | 53 565 |

在采用移动平均加权法时，点异常处按照窗口距离远近分配权重，可得到从远到近的权重分别为 0.17、0.33、0.50。此外，本文 GRU 选用 2 层堆叠结构，每层 GRU 的隐藏单元数为 128 个，以保证模型在训练过程中的收敛稳定性。活函数方面，隐藏层采用 tanh 函数以增强非线性建模能力，输出层采用 Linear 激活函数，以满足连续型 LCC 数值预测的需求。

### 3.2 基于自注意力机制的权重计算结果

由自注意力机制计算得到的已有样本 LCC 数据特征权重分布如图 6 所示。

图 6 中每个 LCC 数据方块的颜色表征该特征与整体 LCC 数据之间的相关性权重。结合自注意力机制的计算原理可知，某一方块在对应行中颜色越深、越偏蓝，表示该类 LCC 数据在特征交互过程中被赋予的注意力权重越高，反映其与其他 LCC 成本要素之间的关联程度越强。其中，数据 1—数据 9 输入分别对应：大修项目材料费和检修费、小修项目材料费和检修费、日常运维检修项目材料费、日常运维检修项目检修费、人工费、车辆使用费、财产保险费、巡视次数和检修次数。从整体分布特征来看，数据 1、数据 2 及数据 9 对应方块整体呈现偏红色，表明其与其他 LCC 成本要素的相关性相对较弱。这主要是由于大修和小修项目具有明显的阶段性与计划性特征，其发生频率和费用规模受专项检修安排影响较大，与日常运行维护成本之间的联动关系相对有限；同时，巡视次数和检修次数更多反映作业强度指标，其与具体费用类数据之间并不呈现严格的线性对应关系。

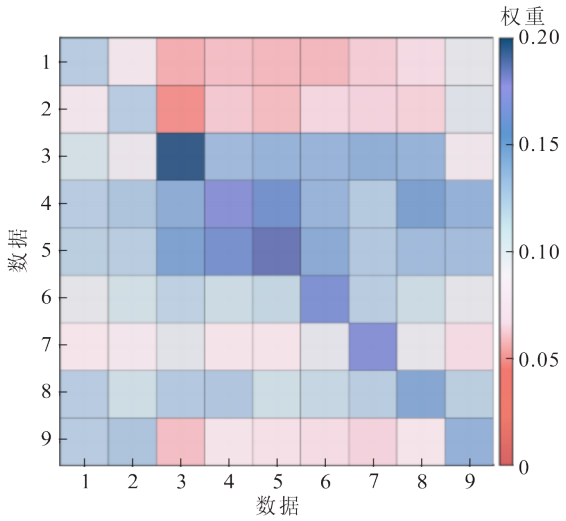



图 6 自注意机制得到的特征权重分布

Fig.6 Feature weight distribution obtained by the self-attention mechanism

相比之下，数据 4 与数据 5 在相互对应的特征权重分布中呈现明显偏蓝颜色，表明二者之间具有显著的相关性。这是由于在实际运维作业中，日常检修项目往往伴随人工投入的同步变化，两类成本具有较强的同源性和时序同步性。此外，数据 3、数据 4、数据 5、数据 6 及数据 8 之间整体呈现较高的自注意力权重分布，这些特征共同构成了 LCC 中日常运行维护成本的主要组成部分，其变化趋势在年度尺度上高度一致。由数据 3、数据 4、数据 5、数据 6 及数据 8 之间实际联动关系可知，自注意力机制能够有效捕捉这种多特征间的内在关系。

对于数据 7，其对应行的权重分布整体偏低，表明该类成本与其他 LCC 成本要素之间的相关性较弱。这主要是由于财产保险费通常以计划方式确定，具有较强的稳定性和固定性特征，其变化更多受保险政策和资产规模影响，而与具体运维 LCC 数据强度关联较弱。上述结果表明，自注意力机制能够学习 LCC 各成本要素的作业与业务属性，自适应地提取不同特征之间的关联关系，从而为后续预测模型提供具有物理与管理意义的特征加权结果。

在此基础上，再根据自注意力机制的 LCC 样本数据扩充模型，可得到不同变电站 GIS 设备 LCC 数据的影响程度权重，并与直接采用注意力机制得到的影响程度权重进行对比，结果如图 7 所示。

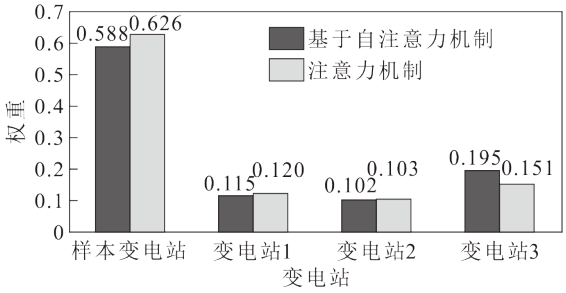



图 7　不同模型得到的影响程度权重

Fig.7　The weights of the degree of influence obtained by different models

从图 7 中可以看出，样本变电站、变电站 1、变电站 2 和变电站 3 基于自注意力机制的影响程度权重分别为 0.588、0.115、0.102、0.195。从权重结果可以看出，除样本变电站外，变电站 3 的权重最大，可能是由于变电站 3 的主变台数、主变容量和进出线回数与样本变电站最为相近，仅在进出线回数上少了 2 回。变电站 1 和变电站 2 比变电站 3 的权重小了近 50%，应是由于变电站 1 的主变台数和主变容量、变电站 2 的主变台数和进出线回数等数量大于样本变电站，这些数量上的差异导致变电站 1 和变电站 2 的 LCC 数据数值较大。变电站 2 的权重最小，可能是 110 kV 侧进出线回数对 LCC 数据的数值影响更大。

对比图 7 中注意力机制的影响程度权重结果，可以看出，样本变电站基于自注意力机制的权重比注意力机制的权重减少了 0.038，这可能是由于注意力机制直接将原始样本数据作为样本特征，而自注意力机制需要从原始样本数据中提取样本特征，两种样本特征存在一定的差异。除了变电站 3，其他相似变电站基于自注意力机制的影响程度权重均有所降低，产生这种情况的原因可能是变电站 3 的 LCC 数据特征更贴近通过自注意力机制提取的数据特征。此外，变电站 2 的权重结果仍为最小，变电站 3 的权重结果仍为最大。

### 3.3　不同方法和模型的训练结果对比分析

根据常规变电站 GIS 设备的寿命年限，本次选取 40 年作为 LCC 的整体时间，以年作为预测步长进行数据预测。考虑扩充样本数量后的变电站 GIS 设备 LCC 数据仍然属于小规模数据集，因此不划分验证集，而是将式(7)处理后的样本数据按照传统比例划分为训练集和测试集，通过训练集训练模型参数，通过测试集验证模型的训练结果。而为了在训练中保证时序数据的连续性，采用分段方式划分数据集，选择前 32 年的连续数据进行模型参数训练，采用后 8 年的连续数据进行验证。将划分的训练集作为 GRU 模型的输入，只输出 LCC 数值。

考虑到本文中模型的组成部分，选取 SA-GRU 模型、ATT-GRU 模型(Attention-GRU 模型)和 GRU 模型进行不同模型训练结果的对比分析。由此得到不同模型对变电站 GIS 设备 LCC 数据的训练结果，如图 8 所示，评估指标 $E_r$、$E_{ma}$ 和 $R^2$ 的结果见表 4。

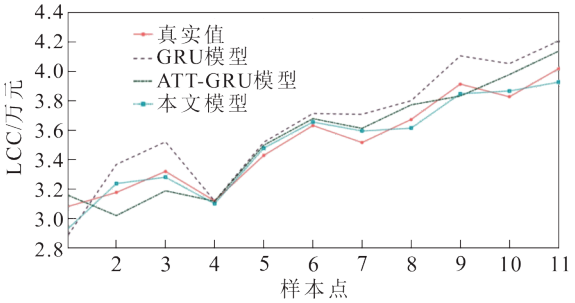



图 8　不同模型训练结果对比

Fig.8　Comparison of training results of different models

表 4　不同模型的 $E_r$、$E_{ma}$ 和 $R^2$ 指标结果

Table 4　Results of $E_r$, $E_{ma}$ and $R^2$ indicators of different models

| 预测模型 | $E_r$/元 | $E_{ma}$/% | $R^2$/% |
|---|---|---|---|
| GRU 模型 | 1 677.33 | 3.83 | 83.32 |
| ATT-GRU 模型 | 1 039.18 | 2.35 | 91.78 |
| 本文方法 | 715.93 | 1.54 | 94.86 |

由图 8 可知，从整体训练结果来看，GRU 模型的整体变化趋势与真实值虽然基本一致，但是预测结果与真实值的差距明显大于 SA-GRU 模型和 ATT-GRU 模型。而由于 ATT-GRU 模型是对 GRU 模型进行的改进，在建模时引入了相似数据，在小样本数据预测方面有一定优势，其单点预测结果虽然不如 SA-GRU 模型，但整体预测结果仍是优于 GRU 模型，比较接近真实值，仅在部分样本点与真实值有所偏差。本文所用的 SA-GRU 模型相较于 ATT-GRU 模型，结果更接近真实值，原因在于 SA-GRU 模型利用自注意力机制增强了样本数据特征，使得接近样本数据特征的相似变电站被赋予的权重有所提高，减小了相似数据与样本数据特征不

完全一致情况对预测准确度带来的负面影响。

从表 4 中可以看出，SA-GRU 模型的 $E_r$ 为 715.93 元，$E_{ma}$ 为 1.54%，$R^2$ 值为 94.86%；ATT-GRU 模型的 $E_r$ 为 1 039.18 元，$E_{ma}$ 结果为 2.35%，$R^2$ 值为 91.78%；GRU 模型的 $E_r$ 为 1 677.33 元，$E_{ma}$ 结果为 3.83%，$R^2$ 值为 83.32%。SA-GRU 模型在 $E_r$ 准确度上相较于 ATT-GRU 模型提高了 31.11%，在 $E_{ma}$ 准确度上相较于 ATT-GRU 模型提高了 34.47%，而 $R^2$ 值相较于 ATT-GRU 模型从 91.78%提升至 94.86%。

为了验证本文所提方法在 LCC 数据预测中的准确性，将本文所提方法同引言中所提到的优化偏最小二乘法以及模糊神经网络法进行对比，训练结果如图 9 所示。

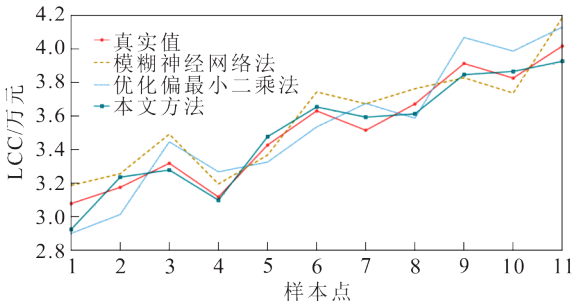



图 9 不同方法训练结果对比

Fig.9 Comparison of training results of different methods

依据评估指标计算式，得到不同预测方法的 $E_r$、$E_{ma}$ 和 $R^2$ 指标结果如表 5 所示。结合图 9 可知，同为时序数据预测方法，基于 SA-GRU 模型预测方法的训练效果和评估指标结果均优于其他方法。优化偏最小二乘法在训练过程中受数值变化影响，所得结果难以准确反映数据变化趋势，而模糊神经网络法虽然能减少数值变化对训练带来的影响，但对数据特征的捕捉能力较弱，这些情况导致这两种方法预测结果的准确度均低于本文所提方法。由表 4 可知，本文所提方法相较于优化偏最小二乘法，$E_r$ 指标下降 672.18 元，$E_{ma}$ 指标下降 1.83%，$R^2$ 指标提高 7.55%；相较于模糊神经网络法，$E_r$ 指标下降 444.64 元，$E_{ma}$ 指标下降 1.19%，$R^2$ 指标提高 9.18%。

此外，为进一步增强模型的准确性，本文还与极端梯度(eXtreme gradient boosting, XGBOOST)、时域卷积网络(temporal convolutional network, TCN)以及长短期记忆神经网络(long short-term memory, LSTM)共 3 种主流时序预测模型进行了对比分析，结果如图 10 所示。

表 5 不同方法的 $E_r$、$E_{ma}$ 和 $R^2$ 指标结果

Table 5 Results of $E_r$, $E_{ma}$ and $R^2$ indicators of different methods

| 预测模型 | $E_r$/元 | $E_{ma}$/% | $R^2$/% |
|---|---|---|---|
| 优化偏最小二乘法 | 1 388.11 | 3.37 | 87.31 |
| 模糊神经网络法 | 1 160.57 | 2.73 | 85.68 |
| 本文方法 | 715.93 | 1.54 | 94.86 |

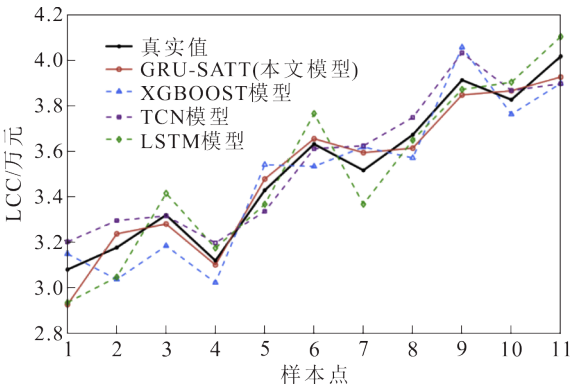



图 10 时序预测模型训练结果对比

Fig.10 Comparison of training results for time series prediction models

从图 10 可知，XGBOOST 模型在部分样本点(如样本点 5 和 9 附近)对局部变化的响应较为敏感，但在样本点 1—4 及 10 附近存在一定程度的偏差，预测曲线在个别区段出现振荡或过度修正。由于 LCC 数据运维、检修和故障成本的发生具有不确定性，故其在一定区间内持续波动，该模型对 LCC 数据这类非平稳时序数据的长期依赖关系刻画有限。TCN 模型整体预测曲线相对平滑，能够较好地跟随真实值的变化趋势，但在样本点 6—9 区间对峰值的拟合存在一定滞后，局部极值点附近预测精度有所下降。LCC 数据作为生产活动的经济数据，其产生和变化具有随机性，而 TCN 在捕捉这种短期突变特征时仍存在不足。相比之下，LSTM 模型在样本点 3 和 6 附近对局部峰值具有较好的拟合能力，但在样本点 7 附近出现明显低估。此外，LSTM 在前期拟合时与真实值差异较大，随着时间的增加，后期拟合曲线更贴近真实值。但 LCC 数据不属于长时序数据，在短期预测时候 LSTM 的稳定性仍有提升。

综合对比，本文模型针对 LCC 数据波动性、随机性和短期性的特点，预测结果在整体趋势跟随能力和局部波动刻画方面均表现出更高的一致性。本文模型的预测曲线在多数样本点上与真实值保持较小偏差，既避免了过度平滑导致的峰值损失，也减

少了对短期噪声的过拟合，体现出较好的鲁棒性和泛化能力。与此同时，XGBOOST 模型的 $E_r$ 为 1 108.4 元，$E_{ma}$ 为 3.08%，$R^2$ 值为 87.22%，与本文模型相比预测误差明显较大，拟合优度相对较低，如表 6 所示，表明传统树模型在刻画 LCC 时序特征和长期依赖关系方面存在一定局限性。TCN 模型与 LSTM 模型的预测性能优于 XGBOOST 模型，两者均能够较好地捕捉 LCC 数据的时间序列特征，但整体预测精度仍低于本文提出的方法。相比之下，本文模型在 $E_r$、$E_{ma}$ 及 $R^2$ 指标上均表现出最优性能，在预测精度和泛化能力方面均具有明显优势，能够更加准确地刻画 LCC 年度变化趋势。

表 6　不同时序模型的 $E_r$、$E_{ma}$ 和 $R^2$ 指标结果

Table 6　Results of $E_r$, $E_{ma}$ and $R^2$ indicators of different methods

| 时序预测模型 | $E_r$/元 | $E_{ma}$ | $R^2$ |
|---|---|---|---|
| XGBOOST 模型 | 1 108.4 | 3.08% | 87.22% |
| TCN 模型 | 915.68 | 2.35% | 91.28% |
| LSTM 模型 | 1 001.00 | 2.65% | 89.58% |
| 本文方法 | 715.93 | 1.54% | 94.86% |

综上所述，本文所提出的预测方法可以有效提升变电站 GIS 设备 LCC 数据预测的准确度。此外，本文还对相似变电站 2、3 的 LCC 数据进行预测，验证了结果的稳定性。

## 4　结论

1）SA-GRU 模型在 $E_r$ 指标、$E_{ma}$ 指标和 $R^2$ 指标结果上均优于 Attention-GRU 模型和 GRU 模型，验证并分析了采用自注意力机制提取数据特征的有效性，同时证明了采用自注意力机制处理相似变电站数据的必要性。

2）相较于传统的优化偏最小二乘法和模糊神经网络法，本文方法的 $E_r$ 指标分别下降了 672.18 元和 444.64 元，$E_{ma}$ 分别指标下降了 1.83%和 1.19%，$R^2$ 指标分别提高了 7.55%和 9.18%，验证了本文所提方法能有效提高变电站 GIS 设备 LCC 数据的准确性，解决了变电站 GIS 设备 LCC 数据难以准确预测的问题。

## 参考文献　References


[1] 蔡亦竹，柳璐，程浩忠，等. 全寿命周期成本(LCC)技术在电力系统中的应用综述[J]. 电力系统保护与控制，2011，39(17)：149-154.
CAI Yizhu, LIU Lu, CHENG Haozhong, et al. Application review of life cycle cost (LCC) technology in power system[J]. Power System Protection and Control, 2011, 39(17): 149-154.

[2] 罗逸涵，杨兰均，成林，等. 变电站隔离开关操作产生的电磁暂态特征量研究[J]. 高电压技术，2024，50(4)：1801-1811.
LUO Yihan, YANG Lanjun, CHENG Lin, et al. Research of electromagnetic transient characteristics in substations due to disconnector switching[J]. High Voltage Engineering, 2024, 50(4): 1801-1811.

[3] 石蓉，张磊，李永毅，等. 基于实物 ID 的电网工程设备费智能化统计算法研究[J]. 电网与清洁能源，2020，36(2)：68-74.
SHI Rong, ZHANG Lei, LI Yongyi, et al. Research on intelligent statistical algorithm of power grid engineering equipment cost based on physical ID[J]. Power System and Clean Energy, 2020, 36(2): 68-74.

[4] 齐冬莲，李启，陈毅，等. 面向电力设备智慧运维的数字孪生关键技术研究现状与展望[J/OL]. 高电压技术，2025：1-17 [2026-03-22]. https://doi.org/10.13336/j.1003-6520.hve. 20251291.
QI Donglian, LI Qi, CHEN Yi, et al. Research status and prospects of digital twin key technologies for intelligent operation and maintenance of power equipment[J/OL]. High Voltage Engineering, 2025: 1-17 [2026-03-22]. https://doi.org/10.13336/j.1003-6520.hve. 20251291.

[5] 杜浙. 基于模糊理论的地下变电站全寿命周期成本分析与管理[D]. 上海：上海交通大学，2013：8-10.
DU Jian. Analysis and management of life-cycle cost of underground substation based on fuzzy theory[D]. Shanghai, China: Shanghai Jiao Tong University, 2013: 8-10.

[6] 张媛. 基于灰色系统理论的电力设备全寿命周期成本评估及模型[D]. 重庆：重庆大学，2008：23-27.
ZHANG Yuan. Electric power equipment's life cycle cost assessment and model based on grey system theory[D]. Chongqing, China: Chongqing University, 2008: 23-27.

[7] 林健宇. G 电网公司变压器的状态评估预测与全寿命周期成本管理研究[D]. 广州：华南理工大学，2016：41-42.
LIN Jianyu. Research on condiction assessment prediction and life cycle cost for transfomer of G power grid[D]. Guangzhou, China: South China University of Technology, 2016: 41-42.

[8] LEE S H, LEE A K, KIM J O. Determining economic life cycle for power transformer based on life cycle cost analysis[C]//2012 IEEE International Power Modulator and High Voltage Conference (IPMHVC). San Diego, USA: IEEE, 2012: 604-607.

[9] 王绵斌，张妍，耿鹏云，等. 基于优化正交偏最小二乘法的变电站全寿命周期成本预测分析[J]. 智慧电力，2020，48(5)：119-124.
WANG Mianbin, ZHANG Yan, GENG Pengyun, et al. Substation life cycle cost prediction model based on optimized orthogonal partial least squares[J]. Smart Power, 2020, 48(5): 119-124.

[10] 孔祥玉，罗家宇，张琪，等. 基于正交信号修正与高效偏最小二乘的质量相关故障检测方法[J]. 控制与决策，2020，35(5)：1167-1174.
KONG Xiangyu, LUO Jiayu, ZHANG Qi, et al. Quality-related fault detection method based on orthogonal signal correction and efficient PLS[J]. Control and Decision, 2020, 35(5): 1167-1174.

[11] 杨磊，应黎明，王玉磊. 变电站建设全生命周期成本估算研究[J]. 计算机仿真，2017，34(1)：123-128.
YANG Lei, YING Liming, WANG Yulei. Research on life cycle cost estimation of substation construction[J]. Computer Simulation, 2017, 34(1): 123-128.

[12] 朱广贺，朱智强，袁逸萍. 基于时序模型和深度学习的设备故障上限评估算法[J]. 吉林大学学报(工学版)，2024，54(8)：2348-2354.
ZHU Guanghe, ZHU Zhiqiang, YUAN Yiping. Device fault ceiling evaluation algorithm based on timing model and deep learning[J]. Journal of Jilin University (Engineering and Technology Edition), 2024, 54(8): 2348-2354.

[13] 熊一，詹智红，柯方超，等. 基于改进 BP 神经网络的变电站检修

运维成本预测[J]. 电力科学与技术学报，2021，36(4)：44-52.
XIONG Yi, ZHAN Zhihong, KE Fangchao, et al. Overhaul operation and maintenance cost prediction of substation based on improved BP neural network[J]. Journal of Electric Power Science and Technology, 2021, 36(4): 44-52.

[14] 彭小圣，熊磊，文劲宇，等. 风电集群短期及超短期功率预测精度改进方法综述[J]. 中国电机工程学报，2016，36(23)：6315-6326.
PENG Xiaosheng, XIONG Lei, WEN Jinyu, et al. A summary of the state of the art for short-term and ultra-short-term wind power prediction of regions[J]. Proceedings of the CSEE, 2016, 36(23): 6315-6326.

[15] 阎洁，许成志，刘永前，等. 基于风速云模型相似日的短期风电功率预测方法[J]. 电力系统自动化，2018，42(6)：53-59.
YAN Jie, XU Chengzhi, LIU Yongqian, et al. Short-term wind power prediction method based on wind speed cloud model in similar day[J]. Automation of Electric Power Systems, 2018, 42(6): 53-59.

[16] 杨胜杰，唐杰，李超良，等. 基于 TCN-FAM-Informer 的短期风电功率预测[J/OL]. 高电压技术，2026：1-11[2026-03-22]. https://doi.org/10.13336/j.1003-6520. hve. 20251604.
YANG Shengjie, TANG Jie, LI Chaoliang, et al. Short-term wind power forecasting based on TCN-FAM-informer[J/OL]. High Voltage Engineering, 2026: 1-11[2026-03-22]. https://doi. org/10.13336/j. 1003-6520. hve. 20251604.

[17] 张承圣，邵振国，陈飞雄，等. 基于条件深度卷积生成对抗网络的新能源发电场景数据迁移方法[J]. 电网技术，2022，46(6)：2182-2189.
ZHANG Chengsheng, SHAO Zhenguo, CHEN Feixiong, et al. Renewable power generation data transferring based on conditional deep convolutions generative adversarial network[J]. Power System Technology, 2022, 46(6): 2182-2189.

[18] 王陈恩，殷豪，陈顺，等. 考虑空间耦合的少数据风电功率预测方法[J]. 南方电网技术，2022，16(6)：75-81.
WANG Chen'en, YIN Hao, CHEN Shun, et al. Wind power forecasting method with few data considering spatial coupling[J]. Southern Power System Technology, 2022, 16(6): 75-81.

[19] 杨子民，彭小圣，熊予涵，等. 计及邻近风电场信息与 CNN-BiLSTM 的短期风电功率预测[J]. 南方电网技术，2023，17(2)：47-56.
YANG Zimin, PENG Xiaosheng, XIONG Yuhan, et al. Short-Term wind power prediction based on information in neighboring wind farms and CNN-BiLSTM[J]. Southern Power System Technology, 2023, 17(2): 47-56.

[20] 叶林，李奕霖，裴铭，等. 寒潮天气小样本条件下的短期风电功率组合预测[J]. 中国电机工程学报，2023，43(2)：543-554.
YE Lin, LI Yilin, PEI Ming, et al. Combined approach for short-term wind power forecasting under cold weather with small sample[J]. Proceedings of the CSEE, 2023, 43(2): 543-554.

[21] YIN H, OU Z H, FU J J, et al. A novel transfer learning approach for wind power prediction based on a serio-parallel deep learning architecture[J]. Energy, 2021, 234: 121271.

[22] 郑颖颖，李鑫，陈延旭，等. 基于 Stacking 多模型融合的极端天气短期风电功率预测方法[J]. 高电压技术，2024，50(9)：3871-3882.
ZHENG Yingying, LI Xin, CHEN Yanxu, et al. Short-term wind power forecasting method in extreme weather based on Stacking multi-model fusion[J]. High Voltage Engineering, 2024, 50(9): 3871-3882.

[23] 李洁，林永峰. 基于多时间尺度 RNN 的时序数据预测[J]. 计算机应用与软件，2018，35(7)：33-37，62.
LI Jie, LIN Yongfeng. Prediction of time series data based on multi-time scale RNN[J]. Computer Applications and Software, 2018, 35(7): 33-37, 62.

[24] 张直政. 神经网络的注意力机制研究[D]. 合肥：中国科学技术大学，2021：58-61.
ZHANG Zhizheng. Research on attention mechanism for neural networks[D]. Hefei, China: University of Science and Technology of China, 2021: 58-61.

[25] 宋立业，孙梦宇，闫云凤，等. 基于改进注意力机制与对比学习的电力设备缺陷图文融合分类方法[J]. 高电压技术，2025，51(9)：4610-4619.
SONG Liye, SUN Mengyu, YAN Yunfeng, et al. Image-text fusion classification method for power equipment defects based on improved attention mechanism and contrastive learning[J]. High Voltage Engineering, 2025, 51(9): 4610-4619.



**唐　波**
1978—，男，博士，教授，博导
主要从事超特高压输电技术与输变电工程电磁环境研究
E-mail: tangboemail@sina.com

TANG Bo
Ph.D., Professor



**尚智宇**(通信作者)
1997—，男，博士生
主要从事电力系统资产管理与超电大尺寸的电磁兼容研究
E-mail: 1064091568@qq.com

SHANG Zhiyu
Ph.D. candidate
Corresponding author

收稿日期　2025-10-27　修回日期　2026-05-14　　编辑　曹昭君