基于SARIMA-BP融合模型的广东省月度用电量预测研究

吴　尧

（联通（广东）产业互联网有限公司，广州　510700）

摘　要： 针对广东省月度用电量序列存在长期增长趋势与复杂季节波动特征的问题，利用2006—2024年月度用电量数据，构建SARIMA-BP融合预测模型。通过季节性差分自回归滑动平均模型（SARIMA）提取用电量序列中的线性变化特征，并利用BP神经网络对预测残差进行拟合修正，实现线性趋势与非线性波动的联合预测。实验结果表明，该融合模型在测试集上的MAPE为7.62%，预测性能优于单一SARIMA模型和BP神经网络模型。在此基础上，对2025—2026年广东省月度用电量进行滚动预测，预测序列呈现出明显的季节变化特征。研究结果验证了SARIMA-BP融合模型在区域月度用电量预测中的有效性，可为电力负荷分析与资源配置提供参考。

关键词： 月度用电量预测；SARIMA模型；BP神经网络；模型融合

中图分类号： TM715　　**文献标识码：** A

Research on monthly electricity consumption forecasting in Guangdong Province based on SARIMA-BP hybrid model

WU Yao

（China Unicom（Guangdong）Industrial Internet Co.，Ltd.，Guangzhou 510700，China）

Abstract：To address the problems of long-term growth trends and complex seasonal fluctuations in the monthly electricity consumption series of Guangdong Province，a SARIMA-BP hybrid forecasting model was developed using monthly electricity consumption data from 2006 to 2024. The The Seasonal Differencing Autoregressive Integrated Moving Average（SARIMA）model was applied to extract linear variation features from the electricity consumption series，while the BP neural network was used to fit and correct the forecasting residuals，achieving joint prediction of linear trends and nonlinear fluctuations. Experimental results show that the hybrid model achieves a MAPE of 7.62% on the test set，outperforming the individual SARIMA model and BP neural network model in forecasting performance. Based on this model，rolling forecasting was conducted for the monthly electricity consumption of Guangdong Province from 2025 to 2026，and the predicted series exhibited obvious seasonal variation characteristics. The results demonstrate the effectiveness of the SARIMA-BP hybrid model in regional monthly electricity consumption forecasting and can provide a reference for power load analysis and resource allocation.

Key words：monthly electricity consumption forecasting，SARIMA model，BP neural network，model fusion

1　引言

随着经济发展和产业结构调整，区域用电需求持续变化，月度用电量预测逐渐成为电网运行分析和电力规划中的重要内容。受产业用电结构、气象条件、季节变化及节假日等因素影响，月度用电量序列通常具有明显的增长趋势、周期性波动和随机扰动特征。由于不同因素对用电需求的影响具有差异性，如何准确描述用电量变化规律并提高预测精度，成为电力负荷预测研究中的重要问题。

在时间序列预测研究中，统计模型因具有较好的可解释性和稳定性得到广泛应用。Box 等[1]提出的SARIMA能够通过差分、自回归和移动平均过程分析具有季节特征的时间序列，在负荷预测、能源需求预测等领域得到应用。然而，SARIMA模型主要依赖历史序列的线性关系，对复杂非线性变化和突发波动的描述能力有限。Rumelhart 等[2]提出的BP神经网络具有较强的非线性拟合能力，能够学习输入变量与预测目标之间的复杂映射关系，但其预测性能易受样本规模、网络结构和参数设置影响。

针对单一模型难以兼顾线性规律与非线性特征的问题，相关研究对统计模型与神经网络进行了融合。

Zhang[3]提出组合预测思想，证明了不同模型之间的信息互补能够提高预测效果；Chen 等[4]将 SARIMA 与 BP 神经网络结合用于电力需求预测，验证了融合模型在预测精度方面的优势。随后，相关研究进一步采用残差修正、权重组合等方式构建混合预测模型，取得了较好的应用效果[5-6]。

基于上述研究，本文针对广东省月度用电量序列中趋势变化与非线性波动并存的问题，构建 SARIMA-BP 融合预测模型；采用广东省历史月度用电量数据开展实验，对比分析不同模型的预测性能，并进一步预测未来月度用电量变化趋势，以验证融合模型在区域用电量预测中的适用性。

2　模型与方法

2.1　SARIMA 模型

SARIMA 可同时对非季节性效应和重复性季节性效应建模。该模型的一般形式为 SARIMA$(p,d,q)(P,D,Q)^i$，其中 p,d,q 分别表示非季节性自回归阶数、差分阶数和移动平均阶数，P,D,Q 分别表示季节性自回归阶数、差分阶数和移动平均阶数，s 为季节周期长度（本文取 s=12）。模型表示如下：

$$\Phi_P B^s \phi_p(B)(1-B)^d\left(1-B^s\right)^D y_t=\Theta_Q(B^s)\theta_q(B)\varepsilon_t \quad (1)$$

式中，y_t 为观测值；ε_t 为白噪声序列；B 为滞后算子；$\phi_p(B)=1-\phi_1 B-\cdots-\phi_p B^p$ 为非季节性自回归多项式；$\theta_q(B)=1-\theta_1 B-\cdots-\theta_q B^q$ 为非季节性滑动平均多项式；$\Phi_P(B^s)$ 和 $\Theta_Q(B^s)$ 为对应的季节性多项式。SARIMA 模型通过历史序列信息建立线性预测关系，能够有效描述用电量序列中的趋势变化和季节规律。本文基于 AIC 准则选择模型参数，具体参数设置及检验结果将在实验部分给出。

2.2　BP 神经网络

反向传播（Back Propagation，BP）神经网络是一种基于误差反向传播算法训练的多层前馈网络，能够建立输入变量与输出结果之间的非线性映射关系。BP 神经网络由输入层、隐藏层和输出层组成，相邻层神经元之间通过权值连接。

对于三层 BP 神经网络，隐藏层第 j 个神经元的输出表示为：

$$h_j=f\left(\sum_{i=1}^{n} w_{ij}x_i|b_j\right) \quad (2)$$

式中，x_i 为输入变量；w_{ij} 为输入层至隐藏层的连接权值；b_j 为偏置项；$f(\cdot)$ 为激活函数（本文采用 ReLU 函数）。此外，输出层采用线性激活函数，以满足连续型用电量预测需求。

2.3　SARIMA-BP 融合模型

为综合利用 SARIMA 模型的线性建模能力和 BP 神经网络的非线性拟合能力，构建 SARIMA-BP 融合预测模型。该模型采用分阶段建模方式，首先利用 SARIMA 描述用电量序列中的趋势和季节变化，其次利用 BP 神经网络对预测残差进行修正。

（1）利用 SARIMA 模型对原始用电量序列进行拟合和预测，得到拟合值 $\hat{y}_t^{\mathrm{SAR}}$，并计算残差序列：

$$e_t=y_t-\hat{y}_t^{\mathrm{SAR}} \quad (3)$$

式中，残差序列 e_t 包含 SARIMA 模型未能捕捉的非线性成分；y_t 为实际用电量；$\hat{y}_t^{\mathrm{SAR}}$ 为 SARIMA 拟合值。

（2）将残差序列 e_t 作为 BP 神经网络的训练目标，以月份、当前月用电量、SARIMA 拟合值为输入特征，建立非线性映射关系：

$$\hat{e}_t=f_{\mathrm{BP}}(\mathrm{Month}_t, y_t, \hat{y}_t^{\mathrm{SAR}}) \quad (4)$$

式中，$\hat{e}_t$ 为 BP 网络对残差的预测值；$f_{\mathrm{BP}}(\cdot)$ 为 BP 网络学习的非线性映射函数。

（3）将 SARIMA 预测值与 BP 残差预测值融合，得到最终预测结果：

$$\hat{y}_t=\hat{y}_t^{\mathrm{SAR}}+\hat{e}_t \quad (5)$$

3　实验与结果分析

3.1　数据来源与预处理

采用广东省 2006 年 3 月—2024 年 12 月的月度全社会用电量数据作为研究对象，数据来源于广东省统计年鉴与电力部门公开数据。按照时间顺序划分数据集，其中训练集包含 2006 年 3 月—2022 年 12 月的用电量数据，共 202 个月；测试集包含 2023 年 1 月—2024 年 12 月的用电量数据，共 24 个月。在此基础上，对用电量数据进行预处理，包括缺失值线性插补、季节性调整和 Min-Max 标准化。模型预测结果经反标准化处理后恢复至原始量纲（亿千瓦时），用于计算各项误差指标。

3.2　SARIMA 实验结果分析

SARIMA 运行结果如表 1 所列。平稳性检验结果表明，原始序列的 ADF 统计量为 1.051 0（p=0.994 8），无法拒绝非平稳原假设；一阶差分后，ADF 统计量降至 -5.287 2（p<0.000 1），序列达到平稳状态，表明原始用电量序列存在非平稳特征，需要进行差分处理。

采用 AIC 准则对 $(p,d,q)(P,D,Q)_{12}$ 进行网格搜索，最优模型为 $(1,1,2)(2,1,2)_{12}$，AIC 值为 1 326.44。Ljung-Box 检验结果表明，Q6，Q12 和 Q18 对应的 p 值分别为 0.507 1，0.833 8，0.723 8，均大于 0.050 0，说明残差序列不存在明显自相关关系。

进一步采用 Jarque-Bera 检验分析残差分布特征，结果显示 JB 统计量为 1 822.999 6（p=0.000 0），表明残差序列偏离正态分布，仍存在部分未被 SARIMA 模型解释的复杂变化特征。因此，引入 BP 神经网络对残差进行进一步修正。

表 1 SARIMA 模型参数确定与检验结果

诊断项目	统计量/参数	检验结果	结论
平稳性检验	ADF 统计量（原始序列）	1.051 0（p=0.994 8）	不平稳
	ADF 统计量（一阶差分）	−5.287 2（p<0.000 1）	平稳
模型定阶	最优阶数（AIC 准则）	SARIMA $(1,1,2)(2,1,2)_{12}$	AIC=1 326.44
残差自相关检验	Ljung-Box Q6	p=0.507 1	不存在自相关
	Ljung-Box Q12	p=0.833 8	不存在自相关
	Ljung-Box Q18	p=0.723 8	不存在自相关
残差正态性检验	Jarque-Bera	1 822.999 6（p=0.000 0）	偏离正态分布

3.3 融合模型实验结果分析

在训练集上，以 SARIMA 残差为目标训练 BP 神经网络。训练过程中采用 L2 正则化和早停机制控制模型过拟合，正则化系数设置为 1.0。残差拟合结果的 R^2 为 −0.070 9，*RMSE* 为 32.78 亿千瓦时，方差解释率仅为 0.09%。该结果表明，残差序列具有较强随机性和较弱自相关性，BP 神经网络对残差变化规律的拟合能力有限。

在测试集上评估各模型性能，评价指标包括 *MAPE*，*RMSE*，*MAE*，R^2。同时，引入季节朴素模型（Naive-12，即预测值等于上一年度同月实际值）作为对比基准，结果如表 2 所列。

表 2 各模型在测试集上的预测精度对比

模型	*MAPE*/%	*RMSE*/亿千瓦时	*MAE*/亿千瓦时	R^2
Naive-12	27.69	239.33	207.73	−3.904
单一 SARIMA	8.20	80.48	59.06	0.445
单一 BP	10.55	96.78	76.15	0.198
SARIMA-BP 融合模型	7.62	77.24	54.87	0.489

由表 2 可知，季节朴素模型预测误差最大且 R^2 为负值，表明仅利用历史同期数据难以充分描述月度用电量变化规律；单一 BP 模型预测效果相对较弱（*MAPE*=10.55%，R^2=0.198），可能与样本规模有限以及单一非线性模型对序列周期变化特征的描述能力不足有关；单一 SARIMA 模型表现较好（*MAPE*=8.20%，R^2=0.445），说明统计模型能够有效捕获用电量序列中的趋势和季节变化；SARIMA-BP 融合模型在各项评价指标中均取得最优结果，*MAPE* 降低至 7.62%，*RMSE* 和 *MAE* 分别降低至 77.24 亿千瓦时与 54.87 亿千瓦时，R^2 提升至 0.489。以上结果表明，融合模型能够综合利用统计模型和神经网络模型的预测优势，提高月度用电量预测精度。

3.4 未来月度用电量预测

基于训练完成的融合模型，对广东省 2025 年 1 月—2026 年 12 月的月度用电量进行滚动预测。如图 1 所示，2025 年和 2026 年的用电量延续增长趋势，夏季仍为全年高峰，2025 年 7 月预计达到 1 020 亿千瓦时，2026 年 7 月进一步增至 1 055 亿千瓦时；冬季出现次高峰，春季和秋季用电量相对平缓。该预测结果与广东省历史用电量的季节变化特征保持一致，可为区域用电负荷分析提供参考。

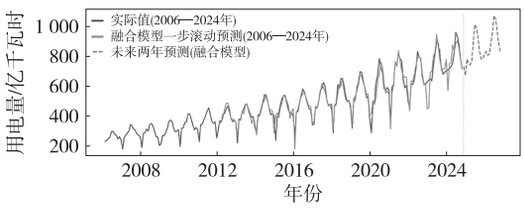

图 1 月度用电量预测结果

4 结束语

针对广东省月度用电量序列中趋势变化与季节波动并存的问题，本文构建了 SARIMA-BP 融合预测模型。通过 SARIMA 模型分析序列中的线性变化规律，并结合 BP 神经网络对预测残差进行修正，提高了模型对复杂用电量变化特征的描述能力。实验结果表明，融合模型在测试集上的 MAPE 为 7.62%，预测精度优于单一 SARIMA 模型和 BP 神经网络模型。未来研究可进一步引入气象、经济等外部影响因素，完善多源信息融合预测机制，以提升不同场景下的用电量预测能力。

参考文献：

[1] BOX G E P，JENKINS G M，REINSEL G C，et al.Time series analysis：forecasting and control［M］. State of New Jersey：John Wiley & Sons，2015.

[2] RUMELHART D E，HINTON G E，WILLIAMS R J.Learning representations by back-propagating errors［J］.nature，1986，323(6088)：533-536.

[3] ZHANG G P.Time series forecasting using a hybrid ARIMA and neural network model［J］. Neurocomputing，2003，50：159-175.

[4] CHEN Y，LIU C，GE J，et al. Deep learning models for forecasting electricity demand in green low-carbon supply chains［J］.International Journal of Low-Carbon Technologies，2024，19：2375-2382.

[5] 王宁，谢敏，邓佳梁，等.基于支持向量机回归组合模型的中长期降温负荷预测［J］.电力系统保护与控制，2016，44(3)：92-97.

[6] 宋鑫磊，黎莫林，詹勤辉，等.基于 SARIMA 与 ANFIS 组合方法的电力物资需求预测［J］.机械设计，2022(6)：66-72.

作者简介：

吴尧(1992—)，本科，工程师，研究方向：具身智能。