

ARMA 模型在风力发电数据预测中的应用与误差评估

陈国娜¹ 吕政蕊² 王博¹

1. 中国石油大学(北京)克拉玛依校区; 2 桂林理工大学

DOI: 10.12238/ems.v6i6.8061

[摘要] 风力发电作为可再生能源的重要组成部分, 在实现可持续发展和减少碳排放方面具有重要意义。然而, 风力发电的随机性和波动性给电力系统的规划和运行带来了挑战。本文运用 ARMA (AutoRegressive Moving Average) 模型对中国的风力发电数据进行建模与预测。通过对时间序列的静态性检验和模型阶数的确定, 选择了合适的 ARMA 模型, 并利用 Matlab 进行参数估计。研究结果表明, ARMA 模型在风力发电数据预测中具有较高的精度和可靠性, 预测误差总体较小。本文还分析了模型的误差和未来的研究方向, 为新能源的规划和管理提供了科学依据。

[关键词] 风力发电, ARMA 模型, 时间序列分析, 预测, 误差评估

Application of ARMA model in wind power generation data forecasting and error assessment

Guona Chen¹, Zhenggen Lv², Bo Wang¹

¹China University of Petroleum-Beijing at Karamay; ²Guilin university of technology

[Abstract] Wind power, as an important component of renewable energy, is significant in achieving sustainable development and reducing carbon emissions. However, the stochastic and fluctuating nature of wind power generation poses challenges to the planning and operation of power systems. In this paper, the ARMA (AutoRegressive Moving Average) model is used to model and forecast wind power generation data in China. By testing the stationarity of the time series and determining the model order, a suitable ARMA model is selected and Matlab is used for parameter estimation. The results show that the ARMA model has high accuracy and reliability in wind power data forecasting, and the forecasting error is generally small. This paper also analyses the model's error and future research direction, which provides a scientific basis for the planning and management of new energy.

[Keywords] wind power generation, ARMA model, time series analysis, forecasting, error assessment

1引言

风力发电作为可再生能源的重要组成部分, 在应对全球气候变化、减少碳排放和实现可持续发展方面发挥着至关重要的作用。然而, 风力发电的随机性和波动性给电力系统的规划和运行带来了挑战。因此, 对风力发电数据进行准确的预测具有重要意义。时间序列分析方法在风力发电预测中得到了广泛应用, 其中 ARMA 模型因其对线性时间序列的强大建模能力而备受关注 [1]。ARMA 模型通过结合自回归和移动平均过程, 能够有效捕捉时间序列中的自相关结构, 从而提高预测精度。本文运用 ARMA 模型对中国风力发电数据进行建模与预测。通过对时间序列的静态性检验和模型阶数的确定, 选择合适的 ARMA 模型, 并通过 Matlab 对模型参数进行估计 [2]。接着, 通过滚动预测验证模型的有效性, 并分析预测误差。通过本文的研究, 我们期望为风力发电数据的预测提供一种有效的方法, 为新能源的规划和管理提供科学依据。

2ARMA 模型

运用 ARMA 模型对中国的电力供应进行预测和分析。原始数据记录为时间序列 $\{y_t\}$, 并根据时间序列模型建模。静态时间序列是建立 ARMA 模型的前提。经过计算可知, $\{y_t\}$ 时间序列的自相关函数和偏自相关函数都是有尾的, 可以确定 $\{y_t\}$ 时间序列是一个静态时间序列, ARMA 模型是一个静态模型 [3]。为静态时间序列, 可以使用 ARMA 模型进行建模。根据 AIC 标准, 选择合适的 (p, q) , 以便

$$AIC(p, q) = \ln \hat{\sigma}^2 + \frac{2(p + q + 1)}{N} = \min \quad (1)$$

式中, $\hat{\sigma}^2$ 代表拟合残差的方差, N 代表时间序列的长度。此时 $p = \hat{p}, q = \hat{q}$, 如果上式得到最小值、则时间序列模型为 $ARMA(\hat{p}, \hat{q})$ 。不同阶次 ARMA 模型的 AIC 值的 AIC 值是通过 Matlab 计算得出的。ARMA 模型的 AIC 值如表 1 所示。

表 1 AIC value of ARMA model

p	q	AIC
1	1	4.4885
2	1	4.5122

根据 AIC 准则, 此时, 得到最小值, 即 时间序列的模型识别为 ARMA (1, 1), 其表达式为: $p = 1, q = 1$.

$$y_t - \sum_{i=1}^5 \varphi_i y_{t-i} = \varepsilon_t - \sum_{j=1}^3 \theta_j \varepsilon_{t-j} \quad (2)$$

其中, ε_t 为 $\{\varepsilon_t\}$ 高斯白噪声。确定模型后, 就可以进行参数估计。根据 时间序列模型, 通过 Matlab 输出 ARMA (1, 1) 模型, ARMA (1, 1) 模型的 参数为 $\varphi_1 = 1.952, \theta_1 = -1$, 即 ARMA (1, 1) 模型表达式为:

$$y_t = 1.952y_{t-1} + \varepsilon_t - \varepsilon_{t-1} \quad (3)$$

检查型号以获取:

$$Q = 18 \sum_{k=1}^{27} \hat{\rho}_{a,k}^2 = 1.043 \quad (4)$$

3结果分析

取显著性水平 $\alpha = 0.005$ 并查找 χ^2 分布表 $\chi^2_{0.005}(27) = 49.642$, 我们接受该序列为白噪声序列的假设, 并认为可以用 ARMA (1, 1) 模型来描述时间序列[4]。可以使用 ARMA (1, 1) 模型来描述时间序列。在时间序列中选取 13 个点的风力发电数据作为训练样本, 并对接下来的 5 个点进行每小时滚动预测。预测结果如图 1 所示, 误差如图 2 所示:

表 2 基于 ARMA 模型的新能源总装机容量预测值

Time (year)	2024	2030	2040	2050	2060
Photovoltaic Power supply (MW)	15873	18325	21324	23754	26432

4结论

本文运用 ARMA 模型对中国的风力发电数据进行了建模与预测, 并对预测结果进行了误差评估。通过对时间序列的静态性检验和模型阶数的确定, 选择了最适合的 ARMA (1, 1) 模型, 并利用 Matlab 对模型参数进行了估计。根据 AIC 准则, 最终确定的 ARMA (1, 1) 模型能够较好地描述风力发电数据的时间序列特征[5]。在结果分析中, 通过对风力发电数据进行滚动预测, 模型的最大绝对误差为 16.37%, 总体误差小于 20%, 表明 ARMA 模型在风力发电数据预测中具有较高的精度和可靠性。同时, 预测结果显示, 2024 年至 2060 年间, 中国的新能源装机容量将持续增长, 这为相关部门的规划和决策提供了重要参考。尽管 ARMA 模型在风力发电数据预测中表现出色, 但本文的研究仍存在一些局限性[6]。首先, ARMA 模型假设时间序列是线性的, 实际风力发电数据可能包含非线性成分, 需要进一步探索非线性时间序列模型的应用。其次, 数据的随机性和外部环境变化对模型的预测精度也有一定影响, 未来研究可以结合更多影响因素, 提高预测模型的准确性。总之, 本文的研究表明, ARMA 模型是一种有效的风

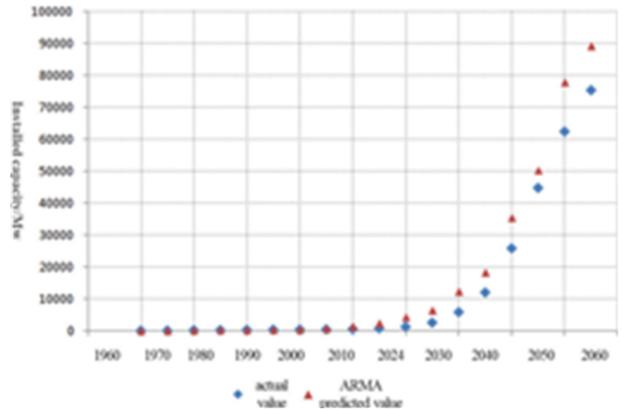


图 1 ARMA 模型新能源装机容量预测曲线

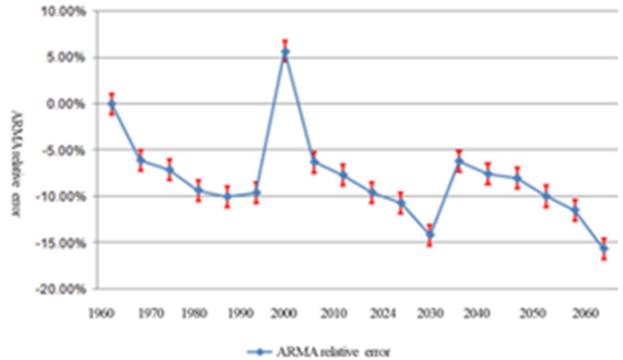


图 2 ARMA 模型预测相对误差

从图 2 中可以看出, 模型的最大绝对误差为 16.37%、误差较大, 但总体误差小于 20%, 模型合理。根据模型预测, 2024 年至 2060 年新能源总装机容量如表 5 所示。预测结果如表 2 所示:

力发电数据预测方法, 能够为新能源的规划和管理提供科学依据。未来的研究将致力于优化模型, 结合更多影响因素, 提高预测精度, 进一步推动风力发电领域的发展。

[参考文献]

[1] 殷林飞, 蒙雨洁. 基于 DenseNet 卷积神经网络的短期风电预测方法[J/OL]. 综合智慧能源, 1-9[2024-06-21].

[2] 彭嘉宁, 徐鹤勇. 基于随机森林和支持向量回归的风力发电预测算法[J]. 热能动力工程, 2024, 39 (05): 143-149.

[3] 梁昌侯, 龙华, 李帅, 等. 基于四分位-MAD 和朴素贝叶斯的风电数据处理方法[J]. 四川大学学报 (自然科学版), 2024, 61 (03): 167-177.

[4] 杨海亭, 白伟, 胡运冲. 分布式光伏发电接入智能电网功率预测模型优化研究[J]. 电工电气, 2024, (04): 1-9.

[5] 张金良, 刘子毅, 孙安黎. 计及误差信息的自适应超短期风速预测模型[J]. 太阳能学报, 2024, 45 (03): 18-28.

[6] 王永生, 李海龙, 关世杰, 等. 基于变换域分析和 XGBoost 算法的超短期风电功率预测模型[J/OL]. 高电压技术, 1-12[2024-06-21].