

基于大数据分析的新能源功率数据修复技术

田坤

国网宁夏电力有限公司调度控制中心

摘要: 大规模新能源发电并网对电网规划、生产和运行等产生重大影响, 不仅会造成电网电压、电流和频率的波动, 影响电网的电能质量, 还会影响发电计划的制定, 对新能源的消纳产生不利影响。本论文利用大数据思维深入分析新能源发电运行数据挖掘相关规律, 提出一种新能源发电数据异常修复技术, 可以为电网规划发展与运行决策提供一种新的理论支持。

关键词: 新能源发电; 数据修复技术; 大数据分析

一、引言

由于新能源场站多处于荒漠之中, 受复杂环境的影响, 导致数据采集系统的通信中断, 造成数据丢失。同时受制于温度的影响, 数据的监测设备、高精度传感器等仪器的灵敏度及精度下降, 从而造成数据大范围异常, 除上述环境影响外, 数据在整个输入、转换、传输过程中都会出现误码或数据丢失的情况。在数据处理过程中, 由于未采用统一的规范化数据处理方式, 人为误差也将导致数据异常或丢失。数据的丢失或者异常, 严重影响新能源发电功率预测的准确度, 进而影响电网规划或调度决策, 降低新能源高效利用的同时, 严重威胁电网安全稳定, 本文针对以上问题, 提出一种数据处理技术, 将新能源出力的异常数据修复准确率提高90%, 从而准确为调度决策提供依据。

二、光伏电站数据异常识别及修复技术

(一) 光伏电站非正常数据鉴别技术

(1) 数据丢失。表现为数据缺失, 其主要原因由通信中断造成;

(2) 数据超出正常范围。主要由错码率、数据存储不正常等原因引起;

(3) 数据与实际值偏差较大。除了第一类和第二类异常数据外, 由于监测异常等原因导致记录数据在正常范围内, 但与实际值偏差较大。

(二) 光伏电站数据异常判别

(1) 第一类异常数据识别

通过数据的几组样本辨别是否存在一些空缺值, 如下所示:

设 $X = \{x_1, x_2, \dots, x_n\}$ 为新能源场站时序输出功率, 当存在 x_i 时:

$$x_i = \emptyset \quad (i \in [1, n]) \quad (1)$$

判断 x_i 存在数据丢失。

(2) 第二类异常数据识别

通过数据辨别可否超出正常数据上限和下限, 如下所示:

$$x_i > x_h \mid x_i < x_l \quad (2)$$

式中, x_h 和 x_l 为正常数据的上限和下限。

(3) 第三类异常数据识别

结合同一地域相邻电站的功率数据, 与观测电站的功率数据相互校验, 从而区分出观测电站的不正常数据, 如下所示:

设 $X_A = \{x_{A1}, x_{A2}, \dots, x_{An}\}$ 为相邻电站的时序输出功率, 则有:

$$\bar{X} = \{\bar{x}_1, \bar{x}_2, K, \bar{x}_n\} \quad (3)$$

$$\bar{X}_A = \{\bar{x}_{A1}, \bar{x}_{A2}, K, \bar{x}_{An}\} \quad (4)$$

$$\bar{x}_i = \frac{x_i - x_{\min}}{x_{\max} - x_{\min}} \quad (5)$$

设两个电站归一化后的功率差值为 e_i

$$e_i = |\bar{x}_i - \bar{x}_{Ai}| \quad (6)$$

不同新能源场站在不同时刻的功率数据是不一样的, 为了

得到较高可信度的识别结果, 根据第一类异常数据识别和第二类异常数据识别, 计算出相邻新能源场站输出功率的平均值和标准差, 最后根据莱以特准则判断数据是否异常, 即 e_i 与功率差值的平均值相差大于标准差三倍时, 就可以认为 x_i 为第三类不正常数据。

(4) 光伏电站异常数据识别流程

根据三类异常数据的异同, 本论文采用大数据出力的方式对异常数据进行识别和筛选, 主要原理图如下所示:

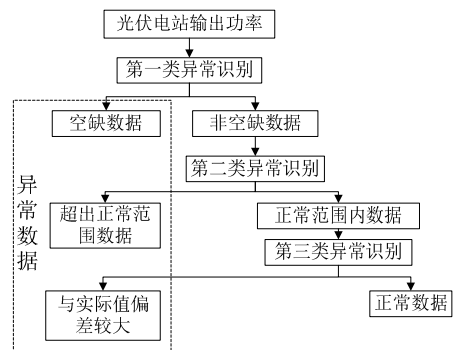


图1 新能源场站异常数据识别流程

第一步: 数据输入;

第二步: 取新能源场站的功率数据特性, 可以确定不同类型异常数据辨别的阈值;

第三步: 筛选电站数据中的空缺数据;

第四步: 在没有缺少数据中, 选出超过正常范围的数据;

第五步: 在其他数据中选出与实际值相差比较大的数据。

三、光伏电站数据修复方法

(一) 基本思路

新能源场站输出功率在不考虑设备运行状态, 当天气条件相似时, 新能源场站的输出功率也存在一定的相似性, 所以可以选择历史数据中的“相似日”或“相似天气”的输出功率对新能源场站的输出功率来进行修复。此时, 新能源电站的输出功率数据修正的可信度主要依靠历史数据的准确度。

获取待修复光伏出力序列所在的目标光伏电站及与其相邻光伏电站的历史数据; 基于所述历史数据计算目标光伏电站及与其相邻光伏电站的相关系数; 基于所述相关系数, 修复光伏出力序列在设定时段的光伏出力数据进行修复; 所述待修复光伏出力序列为带有异常数据的光伏出力序列。

(1) 仅存在“相似日”的状况下, 可以利用现有方法进行数据修复;

(2) 不存在“相似日”的状况下, 借助“相关电站”的方法解决数据修复;

(3) 当“相似日”和“相关电站”同时存在, 可以借助“相关电站”的数据对原方法的修复数据进行进一步验证。

(二) 练习样本

由历史数据特点, 用于光伏电站发电功率数据修复的练习样本有出现下列三种情况的可能:

(1) 仅存在“相似日”时, 练习样本是要修复电站在“相似日”的历史发力数据;

(2) 仅存在“相似电站”时, 练习样本为“相似电站”的历史发力数据;

(3) “相似日”和“相关电站”都有时, 练习样本为“相关电站”在“相似日”的历史发力数据。

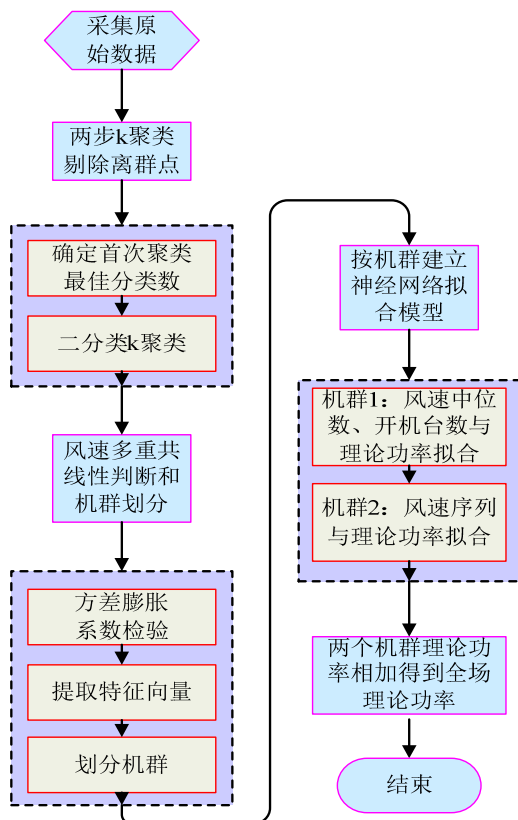


图2 算法流程图

其中，情况1的修复数据可信度在于电站自身数据的可信度，情况2的修复数据可信度取决于“相关电站”输出功率数据的可信度，情况3的修复数据可信度最高。

(三) 算法流程

算法流程如2所示。首先通过多步k聚类算法对原始数据进行异常点的剔除，然后对各风机的机头风速序列进行多重共线性检验，根据检验结果将风电场分为线性强相关机群和线性弱相关机群，对于线性强相关机群，抽取各风机风速的中位数作为机群的代表风速，对于线性弱相关机群则直接采用原始风速序列，进而通过人工神经网络建立各机群风速和理论功率的映射模型。

四、总结

本论文利用大数据思维深入分析新能源发电运行数据挖掘相关规律，提出一种新能源发电数据异常修复技术，为电网规划发展与运行决策提供一种新的理论支持，因此本论文具有以下特点：

- 优化电网规划、生产和运行；
- 降低电网电压、电流和频率的波动；
- 提高电网的电能质量；
- 优化调度发电计划的，提升新能源消纳。

参考文献

[1]刁瑞盛. 风力发电对电网的影响研究[D]. 浙江大学, 2006.
 [2]刘姝. “风力发电原理与应用”课程教学改革与探索[J]. 现代企业教育, 2014.
 [3]张建勋. 发展潜力巨大的风力发电[J]. 内蒙古科技与经济, 2006.
 [4]朱家胜,陈旭,张炜. 中国风力发电前景探讨[J]. 科技资讯, 2012.

(上接第340页)

持。所以，施工单位需根据自身实际情况积极组建规范且职业素养高的安管团队，由此避免施工现场出现安全事故而影响施工方的经济效益。一方面，施工单位应适当引入高新技术为安全管理人员提供便捷条件；另一方面，可设置考核机制，促使安全管理人员能够认识到安全管理工作的重要性。同时，还可拓宽安全管理人员的招聘渠道，以此扩大人才规模。现如今比较常见的安全管理技术包括BIM技术、虚拟仿真技术等。由于施工现场存在许多不可见的危险因素，一旦安全管理人员经验不足很容易忽视潜在隐患。

(六) 加强安全技术交底，强化检查监督力度

由于分部分项工程施工技术要点和安全管理要点不一，为确保建筑工程施工安全，督促分包单位加强技术交底，尤其是脚手架、土方、钢筋工程等安全风险较高的分项工程，要求必须按要求编写专项安全技术方案，细化专项技术方案和安全技术措施，技术交底要求落实到个人，并签字确认，实现安全技术交底全覆盖，切实提高全员安全意识，防范和控制施工安全风险。

(七) 落实安全管理制度执行情况检查监督

为防止安全管理制度形式化、表面化，改善施工现场安全管理执行情况不到位的问题，应加强安全管理制度执行情况检查监督，明确分部分项工程安全管理人员责任，建立安全管理

人员奖惩考核制度，强化现场安全管理人员责任意识。同时，每天检查现场安全管理记录，对照施工现场安全风险识别记录，检查安全记录完整性、准确性和真实性，确保现场安全管理制度落实到位。

三、结束语

总而言之，建筑工程建设期间影响质量安全的因素较多，在应对时需深入分析目前建筑工程质量安全管理方面存在的常见问题，针对性采取可行对策，通过提升质量安全管理水平、健全质量安全管理机制、全面开展监督工作、综合分析自然因素影响等方法，保证建筑工程施工有序落实，顺利实现工程安全质量管理目标。

参考文献

[1]陈绍军. 建筑施工现场安全管理存在的问题及对策分析[J]. 科技展望, 2016(019):188.
 [2]侯增光. 谈新时期如何加强建筑工程安全监督管理[J]. 山西建筑, 2017,(1):240-241.
 [3]武华东. 新时期加强建筑工程安全监督管理的方略[J]. 科技经济导刊, 2018,(03).
 [4]周劲进. 建筑工程施工现场管理及其优化措施[J]. 建材与装饰, 2018(46):187-188.