IOP Conference Series: Earth and Environmental Science

New Energy Station Monitoring Equipment Rule Verification Model Based on Data Envelopment Analysis

To cite this article: Dongge Zhu *et al* 2021 *IOP Conf. Ser.: Earth Environ. Sci.* **645** 012004

IOP Conf. Series: Earth and Environmental Science **645** (2021) 012004 doi:10.1088/1755-1315/645/1/012004

New Energy Station Monitoring Equipment Rule Verification Model Based on Data Envelopment Analysis

Dongge Zhu[1*], Yiming Ma[1], Yongliang Li[1], Zhenhua Yan[1], Jia Liu[1] and Rui Ma[1]

[1]Electric Power Research Institute of State Grid Ningxia Electric Power Co.,Ltd., Ningxia, 750002, China

*E-mail: gen@bjcehh.com

Abstract. According to the functional and performance requirements of the new energy station monitoring equipment, an operating rule verification model based on data envelopment analysis is proposed, and the online automatic discovery of equipment and safe operation rules for key links are designed. The equipment adopts data envelopment analysis method to read real-time data of new energy station monitoring equipment, optimizes the automatic discovery function of key sections, and tests the operating parameters of wind turbines and wind farms online through remote testers, formulates monitoring rules, and forms data requirements , Use algorithms to acquire and process data, and build models. Realize the discrete processing and rule mining of key links and safety operating procedures to ensure the safe operation of new energy station monitoring equipment. Finally, the experimental results show that the equipment can quickly and automatically generate new energy station monitoring equipment safe operation knowledge and improve new energy Safety of station monitoring equipment.

1. Introduction

Because there are great differences in the power generation modes of wind energy, solar energy and traditional energy, the large-capacity and long-distance transmission of wind and photovoltaic power generation will have a significant impact on the voltage stability, frequency stability and power quality of the entire power grid[1]. At present, the operation status data of wind farm and photovoltaic power station are numerous, which often have deeper value and excavation demand. It is necessary to monitor the account information and equipment operation status of the station[2]. Under complex operating conditions, it is not only difficult for traditional energy station monitoring equipment to process the knowledge flow of its safe operation, but also the acquired knowledge cannot adapt to the changes of online operation mode and can not guarantee the safe operation of monitoring equipment. The operation of monitoring equipment in new energy station can make a preliminary judgment on the weak links and stability problems of the equipment through the operation experience accumulated by operators. However, it is impossible to be objective and comprehensive by human experience alone, which brings risks to the operation of new energy monitoring equipment[3]. This kind of judgment usually manifests as experience and intuition, and it is difficult to be comprehensive, stable and clear. Therefore, a new energy station monitoring equipment operation rule verification model based on data envelope analysis is proposed. The stable state of the equipment is obtained by machine learning method instead of manual experience judgment, so that the standard rules can more clearly reflect the operating state of the equipment, so as to establish scientific and reasonable verification rules to ensure the safe and reliable operation of the monitoring equipment.

2. Verification model of operation rules of monitoring equipment in new energy station

2.1. Discrete operation data of monitoring equipment in new energy station

Initial value characteristics of monitoring equipment in new energy station need to select data reflecting equipment steady state. Station account information, equipment status data, operation data and meteorological data are collected, and geographic information such as station telecontrol rate, wind abandonment rate, availability rate and energy utilization rate are acquired and analyzed, so as to evaluate station status and equipment operation status[4]. The sampling method was used to obtain the m-sample set. For each identical set, several features are randomly selected, and each decision package model is trained with randomly selected samples and features. When playback sampling is used, N samples are taken to build a grouping model for each decision, thus obtaining some results that have never been sampled. The limitations that can be considered are as follows:

$$\lim_{N\to\infty}\left(1+\frac{1}{N}\right)^{N}=e^{-1}\approx 0.368 \quad (1)$$

OOB data can be used as test data for each decision package model to test the classification accuracy of each base class. The test error rate of the out-of-bag data of all decision package models, i. e. the classification error rate of the whole random validation model, was averaged. After calculating the error rate of each decision package model, the value of feature X is randomly changed, and then the error rate of the outer bag data of the model is calculated. Before and after noise addition, the difference of predicted value of outsourcing data is an important factor for feature determination. Based on the characteristic importance of this data packet Xj, the specific algorithm is as follows:

$$FIM_{ij}^{OOB}=\frac{\mathrm{X_j}\sum_{p=1}^{n_{OOB}^{i}} I\left(Y_p=Y_p^i\right)}{n_{OOB}^{i}}-\frac{\mathrm{X}\sum_{p=1}^{\mathrm{n^i_{oob}}} I\left(Y_p=Y_{p,s}^i\right)}{n_{OOB}^{i}} \quad (2)$$

In the above algorithm, $\mathrm{n^i_{OOB}}$ indicates the number of observation samples of the ith data packet of monitoring equipment data of new energy station, and I is a small function of monitoring equipment operation of new energy station. Yp∈{0,1} is the real results of the p-th observation of the monitoring equipment operation data of the new energy station, $\mathrm{Y_p}^{\mathrm{i}} \in \{0,1)$ is the predicted result of the i-th data packet of the Rule Verification Model to the p-th of the out-of-bag data.

The association rule algorithm is applied to the safety and stability evaluation of monitoring equipment in new energy station. Where possible, all possible operating states of the equipment are covered based on a large amount of simulated and historical data. According to the statistical principle, the relationship between equipment status and safety and stability is analyzed, and the real-time status of equipment is judged accordingly. In order to meet the requirements of association rule algorithm, the key and necessary step is to discretize continuous static features before the rule runs stably[5]. Finding out the discrete points between continuous attribute subdomains, reasonably determining the breakpoints and transforming them into finite intervals are the key to construct stable operation rules. On the basis of K-means weighted clustering and information ordering, a feature discretization algorithm is proposed. The importance of feature selection model was evaluated, and the importance of each attribute in classification was determined, which was regarded as the weight of data envelope spatial clustering. When N samples and C target classes constitute a set of data, the discretization algorithm discretizes the continuous attribute A in the data set into m discrete intervals:

$$\mathrm{D}=CNA\,/\,dFIM_{j}^{OOB} \quad (3)$$

$d=\{[d_0,d_1],(d_1,d_2],\ldots,(d_{m-1},d_m]\}$, The lowest value in feature A is d0 and the highest value is dm. This discrete result D is called attribute P, because data envelope analysis method is the most classical clustering algorithm and the most widely used clustering method, so the difference of sample space data is usually measured by Euclidean distance, so data envelope is generally expressed by error square sum as objective function and SEE:

$$\mathrm{SSE} = D\sum_{i=1}^{k}\sum_{x\in C_i} q\left(C_i - Px\right)^2 \quad (4)$$

In the above algorithm, Ci is the number of clusters, q is the value of cluster center, and x is the location of sample data in the cluster. The smaller the number, the better the choice of the center point.

Different discretization algorithms have different classification results. This method has better performance, reduces or simplifies data, improves learning speed, produces more accurate and concise results, and shorter calculation time; discrete attributes are easier to understand, use and interpret for researchers and practitioners. However, any discretization process usually leads to the loss of information, so minimization is the main purpose of discretization[6]. The design of unsupervised discretization is simple, but the lack of classification information of samples in the process of discretization makes it difficult to improve the degree of discretization. Compared with the unsupervised discretization method, the supervised discretization method refers to the discretization process of class attributes and combines several classification algorithms, so the discretization effect is generally better than the unsupervised discretization method.

2.2. Rule mining of monitoring equipment operation data of new energy station

Further, Python language is used to implement data envelope analysis, and a discretized database of stability evaluation data samples is mined; this method only needs to obtain the association rules related to stability, without the need to associate feature attributes. In order to improve the efficiency of the algorithm, it is first necessary to determine whether class attributes are included in the conditional pattern library[7]. If included, the FP-Tree condition continues to be generated. Otherwise, the next frequent pattern library will be entered without generating the conditional FP-Tree. In this way, useless rules can be filtered out and the efficiency of the algorithm can be further improved. Before Mining Association rules, stable scheduling operation rules need to be found by setting appropriate minimum support and reliability. Pragmatism reflects the importance of generating rules and reliability reflects the reliability of rules, while the size of both affects the size of generating rule base. To ensure the reliability of extracting transient stability rules, we set the threshold below 80% confidence.

The stable operation rules generated by the test set samples can be used for stable identification and classification. For most unstable samples, rules can be correctly identified, while misclassified samples are mostly stable samples. By establishing a hierarchical correlation analysis model of empirical data and historical data, acquiring key factors of equipment stability problems at all levels in advance, and carrying out fine management, providing operators with auxiliary decision-making information, the stability analysis ability and control level of monitoring equipment in new energy stations can be improved[8]. According to the line stability limit calculated by the dispatching center of monitoring equipment operation mode of new energy station, the risk degree of each part of the whole network is evaluated. In order to achieve "dimension reduction" control of monitoring equipment of complex new energy station, it is usually necessary to analyze monitoring equipment of new energy station offline and obtain its critical section limit by simulation calculation.

2.3. Verification of Rules for Operational Data of Monitoring Equipment in New Energy Stations

New energy station monitoring equipment status parameters online monitoring remote control device detects leakage of ground wire at sleeve end by signal acquisition terminal and measuring device, transmits measurement signal to field monitoring device in real time by analog-to-digital conversion device and wireless communication device, transmits collected current signal to field display terminal by wireless communication, and displays leakage current of main insulation of sleeve in real time; signal acquisition The transmission equipment of communication device is powered by solar cell and backup battery, and the power supply control is controlled by photoelectric remote end, so that the operator can start up at a long distance to reach a safe distance. To ensure the reliability of the extracted[9]. Association rules, the minimum confidence threshold was set to 80%. Because monitoring equipment of new energy station seldom overloads seriously in actual operation, if the minimum support value set based on effective overload rule is too high, then the minimum support value will be set to 10%. The central station main thread realizes centralized management and balanced scheduling of specific task threads, and manages and distributes them according to the running load of the threads. In view of the problems existing in the existing monitoring equipment and the practicability, advancement and reliability of the equipment requirements, the overall design includes monitoring the main insulation leakage current with the photoelectric control switch at the high voltage side of the power equipment, including the composition of the optical control circuit and the sampling and monitoring of the control circuit[10].The equipment mainly collects the relative capacitance and leakage current of the bushing, calculates the relative capacitance and insulation resistance of the bushing according to the capacitance and sensitivity components of the current, realizes the on-line monitoring and on-line operation of the insulation level of the bushing, thus realizing more rapid and accurate verification of the operating rules of the monitoring equipment of new energy station, and ensuring the safe operation of the equipment in the station. Study requirements.

3. Analysis of experimental results

In order to validate the running effect of this model, experimental tests are carried out. The test software runs on common operating devices and hardware platforms, especially in Windows, Linux and UNIX environments widely used in practice. Considering the company's inheritance of previous software, C++ high-level program language is adopted, and Oracle, SQL Server, through standardized data access interface are used. Mainstream database server integration such as Sybase. Programs closely related to data management are written directly in standard SQL language and compiled to a specific database platform when deployed, with process space running on the database server. The device is a distributed application device, which adopts the mode of centralized management and decentralized acquisition.

Based on the above environment, carry out experimental testing, and record the test results, as shown in the following figure:

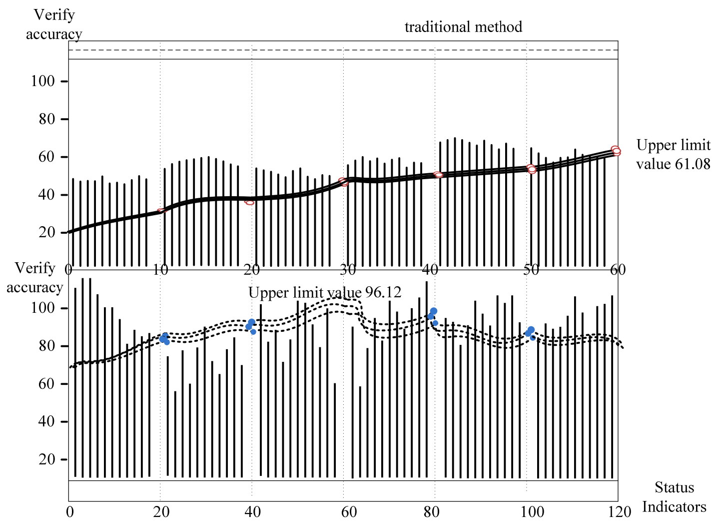

Figure.1 Comparison of experimental results

Based on the results of the above-mentioned analysis, compared with the traditional method, the validation accuracy of rule data validation of the proposed new energy station monitoring equipment operation rule validation model based on data envelope analysis in the actual application process is significantly improved compared with the traditional method, and can better carry out the rapid detection of large-scale index data, thus confirming thatBased on data envelope analysis, the new energy station monitoring equipment operation rules validation model has relatively better practical application effect and fully meets the research requirements.

4. Concluding remarks

Through the comprehensive utilization of equipment operation data and equipment operation data, the health assessment of equipment operation status is realized, and the data support for the healthy operation and scheduling decision of new energy equipment is provided.The design scheme and safety operation rules of automatic detection equipment for key components of new energy station are proposed, the business contents of monitoring are sorted out, the monitoring rules are formulated, the data requirements are formed, the collected data are processed algorithmically, and the monitoring model is established. Finally, the rules are verified and adjusted, the model results are output, the status of the site is evaluated, and the real-time data of the site is read by data envelope analysis method, which ensures the safety monitoring of the new energy station.

Reference

[1] Rouzegari N , Hassanzadeh Y , Sattari M T. (2019) Using the Hybrid Simulated Annealing-M5 Tree Algorithms to Extract the If-Then Operation Rules in a Single Reservoir[J]. Water Resources Management, 33(10):3655-3672.

[2] Abbas M , Dalia S , Tomas B , et al. (2018) Data Envelopment Analysis in Energy and Environmental Economics: An Overview of the State-of-the-Art and Recent Development Trends[J]. Energies,11(8):2002-2008.

[3] Hosseinzadeh-Bandbafha H , Nabavi-Pelesaraei A , Khanali M , et al. (2018) Application of data envelopment analysis approach for optimization of energy use and reduction of greenhouse gas emission in peanut production of Iran[J]. Journal of Cleaner Production, 172(2):1327-1335.

[4] Yang W , Li L . (2018) Efficiency evaluation of industrial waste gas control in China: A study based on data envelopment analysis (DEA) model[J]. Journal of Cleaner Production, 179(APR. 1):1-11.

[5] Mehlawat M K , Gupta P , Kumar A , et al. (2020) Multi-Objective Fuzzy Portfolio Performance Evaluation Using Data Envelopment Analysis Under Credibilistic Framework[J]. IEEE Transactions on Fuzzy Systems, (99):1.

[6] Feng M , Li X . (2020) Evaluating the efficiency of industrial environmental regulation in China:A three-stage data envelopment analysis approach[J]. Journal of Cleaner Production, 242(Jan. 1):118535. 1-118535. 13.

[7] Ann M H , Kim J , Kim M , et al. (2019) Device Design Rules and Operation Principles of High-Power Perovskite Solar Cells for Indoor Applications[J]. Nano Energy, 68(4):104321.

[8] Zhang Y , Wu J , Yu H , et al. (2019) Formulation and Implementation of Short-Term Optimal Reservoir Operation Schemes Integrated with Operation Rules[J]. Water, 11(5):944.

[9] Chen B , Wan J , Xia M , et al. (2020) Exploring Equipment Electrocardiogram Mechanism for Performance Degradation Monitoring in Smart Manufacturing[J]. IEEE/ASME Transactions on Mechatronics, (99):1.

[10] Zelikman M I , Kruchinin S A , Mikheev D V . (2018) Assessment of the Technical Condition of X-Ray Equipment at Healthcare Facilities in the Russian Federation Based on Periodic Monitoring of Operational Parameters[J]. Biomedical Engineering, 51(6):1-3.