A Load Modeling Method Based on Machine Learning

Lei Song¹, Chun-guang Lu¹, Xiao-qiong Huang¹, Qing-juan Wang¹, Chen-hui Zhou¹, Chi-Jiang¹, Chao-liang Wang¹, Yi DING², Rui-peng DIAO², Wei-Liu¹, Yi-long LI¹

¹Marketing Service Center, State Grid Zhejiang Electric Power Co., Ltd. Hang Zhou 310014, China; ²Qingdao Topscomm Communication Co., Ltd. Qingdao 266109, China

Abstract—More and more high-power impact loads in the power grid are put into use. When the power is impacted, these loads will cause the power grid bus voltage fluctuations and reduce the power quality. Therefore, in order to accurately analyze the changes in the power system, it is particularly important to establish a reasonable impact load model. This paper takes the impact load data generated by the CT machine during exposure as an example to analyze the characteristics of the impact load. Based on the active power characteristics of the impact load, machine learning methods (such as support vector machines and long and short-term memory networks) are introduced into the impact load modeling, and the power waveform change law of the impact load is accurately established by this method. Finally, the effectiveness of the method is verified by simulation.

1. Introduction
The problem of load modeling in power system has been restricting the development of power system analysis. There are a large number of loads with impulsive power in the power supply system. These loads will bring huge impacts to the system during production, causing voltage fluctuations, that is, "flicker". However, due to the different internal mechanisms and characteristics of this kind of impact load, the operating state changes with the seasonal climate change, making the modeling work more difficult. Accurate modeling of impact load is the need for reasonable and stable operation of the power grid.

In recent years, the power industry has conducted a lot of researches on the modeling of impact load and achieved certain results. Literature⁴ puts forward a model combining power demand and self-healing dynamic model through the study of impact load characteristics; Literature⁵ discusses and studies the time-varying nature of load, and improves the model structure of comprehensive load. Literature⁶ combines the dynamic model of the asynchronous motor with the initiative of the actual running power demand of the rolling mill, which reflects the periodic regular changes of the power demand of the rolling mill. Literature⁷ adds the frequency variable to the load model and proposes a model equation with frequency characteristics. Literature⁸ adopts a method of identifying multiple curve fitting parameters of the model and verifies its effectiveness.

In this paper, two non-mechanical modeling methods based on machine learning are proposed on the basis of in-depth analysis of the impact load model described by predecessors and actual measurement data. First, the regression theory of Support Vector Machine (SVM) is adopted to establish the model relationship between voltage and active power P and reactive power Q. Then, considering that the power system data is a kind of time series data, the current moment will be affected by the state of the historical...
moment, using the memory characteristics of the Long Short Term Memory (LSTM) to establish a deep learning model. Finally, by analyzing the effects of the two modeling methods on the measured data, the feasibility of the method is verified.

2. Impact load characteristics and its model

2.1. Impact load characteristics

Large-scale impact loads that can produce impact loads often have the following characteristics: the connected voltage level is high, the amplitude of power changes is large, and the frequency of power changes is high. Impact load is a load that the peak power generated by a large impact load during startup or work is several times or even dozens of times during normal operation. This type of load has a greater impact on the power system and is likely to cause continuous oscillations in the frequency of the power system. In the actual power grid, the voltage fluctuation range is relatively stable, and the impact load characteristics are mainly reflected in the current change characteristics, which have nonlinear, initiative of power demand, dynamic time-varying, and fluctuating frequently.

This paper focuses on the common CT machine in impact load [6], and analyzes its production characteristics and modeling.

The CT machine scans a certain part of the human body with an X-ray beam. When the X-ray is directed to the human tissue, part of the radiation is absorbed by the tissue, and part of the radiation passes through the human body and is received by the detector to generate a signal, which is then converted into digital information and then performed by the computer processing, output to the display screen to display the image. When the CT machine is working, due to the exposure of the X-ray machine, the current amplitude during exposure is much larger than in the normal state, which makes the load of the CT machine change from the normal state to the transient state. After the exposure, it returns to the normal state and the transient duration is shorter. It is normal most of the time. The recorded current and power waveforms of the CT machine before and after exposure are shown in Figure 1.

![Fig. 1. Impact characteristics of CT machine](image)

Taking active power as an example, after normalization, the impulsive waveform is shown in Figure 2.
It can be seen from Figure 2 that the peak power of the CT machine during impact is about 10 times the peak power of the steady state, the duration is extremely short, and it has typical impact characteristics. Discrete Fourier transform analysis is performed on the inrush current to obtain the current harmonic content. As shown in Figure 3, when the CT machine is working, once the machine rays are exposed, it will maintain a relatively stable state in a short period of time. When the current signal is impacted, the low-order odd-numbered harmonics account for a higher component.

2.2. Impact load model
The operation process of impact load is divided into stable operation state and power shock state. The stable running state corresponds to the load characteristic model of the impact load. At this time, the power absorbed by the load changes gently with the bus voltage and frequency. At this time, the static model can be used for analysis. When the impact load itself requires a larger power, the system will have a power shock. This is the difference between the impact load and the conventional load, that is, "proactive". The power absorbed from the system is determined by its own production characteristics[7-9]. When analyzing impact characteristics, dynamic models are generally used for analysis. At present, most of the modeling of impact load is based on the actual voltage, current, power and other data measured on site for modeling and analysis[10-11].
There is no need to know the internal characteristics, the model has low interpretability and can characterize static and dynamic characteristics.

3. Machine learning impact load modeling method

3.1. Impact load model based on support vector machine regression

In this section, the support vector machine regression (SVR) algorithm is introduced into the impact load modeling. The measured data is used as the training set, and the model is regarded as a black box. There is no need to understand the internal mechanism of the relevant impact load and establish the node voltage of the impact load[13].

Support Vector Machine (SVM) is a statistical learning method with rigorous theoretical and mathematical derivation. It seeks to minimize the structured risk to improve the generalization ability of the learning machine, and to minimize the experience risk and confidence range. This method is an algorithm that can be used for data regression. The corresponding support vector regression (SVR) model is obtained by training the input continuous label data.

In view of the difficulty of collecting most electric impact load data, the data has the characteristics of privacy, complexity, and high confidentiality. The available reference data are often relatively small. The SVR method has low dependence on the number of samples. In the case of small statistical samples, it can better fit high-dimensional nonlinear data, so as to achieve the purpose of good statistical law. Therefore, in view of modeling requirements and realistic conditions, the SVR regression model is introduced into the electrical impact load modeling.

For the input power load variables \(((t_1, P_1), (t_2, P_2), ..., (t_m, P_m))\), \(t \in \mathbb{R}^d\), \(P \in \mathbb{R}\), there is a high-dimensional linear regression function \(\xi(t) = w \cdot t + b\) that can fit the relationship between \(P\) and \(t\). In the function fitting, an appropriate kernel function \(\kappa(t, t')\) is used to represent the high-dimensional inner product space operation after nonlinear mapping to avoid complex nonlinear transformations. The final SVR form is:

\[
\xi(t) = \sum_{i=1}^{m} (\beta_i - \beta_i^*) \kappa(t, t_i) + b(1),
\]

where \(\beta_i, \beta_i^*\) are Lagrange multipliers.

For impact load modeling, the initialization of the SVR model is affected by the kernel function parameters and relaxation factor coefficients. The SVR model uses traversal to find the best parameters by default. This method is relatively time-consuming, and if you want to find better SVR model parameters in a larger range, the time complexity will be very high.

In this section, two heuristic methods of genetic algorithm (GA) and particle swarm optimization (PSO) are used to optimize the parameter search method of SVR. The genetic algorithm encodes the parameters \(C\) and \(\gamma\) to be solved into chromosomes, and obtains the optimal solution through \(k\) generations of heredity, mutation, crossover, and replication operations. The particle swarm algorithm encodes the parameters to be solved into the optimization space, and uses the information sharing of individuals in the group to make the movement of the whole group produce an evolution process from disorder to order in the problem solving space, thereby obtaining the optimal solution. The two heuristic methods can avoid large-scale traversal of all data, and use less time to search for the optimal SVR model initialization parameters in a larger range.
The overall modeling process of SVR is shown in Figure 4. First, the measured data is imported into the model, and then the appropriate model initialization parameters are obtained through two optimization methods, and then the best fitting curve is obtained through training and testing, and the support vector is determined and then the model is saved.

3.2. Impact load modeling based on long and short-term memory network

The power load data itself is data about time and can be regarded as a time series. For time series, the current moment data is not only related to the current moment input, but may also be affected by the previous moments. The output of Long Short Term Memory (LSTM) in the current period and other data in the next period will be used as input to the current neuron, so that historical data will also affect the characteristics of the current sequence[14].

LSTM is mainly composed of cell nucleus (cell) one by one, and the state of the current cell nucleus can be transmitted to the next cell nucleus. In order to selectively discard and add to the transmitted information, the concept of gates is introduced. Finally, after continuous loop iteration[15], the output gate determines the information to be output.

Through these operations, LSTM can achieve both short-term and long-term dependence. In other words, when performing regression analysis, the current target value is not only related to the current input value but also related to the historical input value.

Impact load modeling is to establish the relationship between voltage U, active power P, and reactive power Q, so as to reflect the internal change law of the equipment when the impact occurs. At the same time, the output at the current moment is not only related to the state at the current moment, but also affected by the previous moment. Therefore, when modeling, use the long and short-term memory capabilities of LSTM to establish the following relationship model:

\[
P(t) = f(U(t), U(t-1), ..., U(t-n), P(t-1), ..., P(t-n))
\]

\[
Q(t) = f(U(t), U(t-1), ..., U(t-n), Q(t-1), ..., Q(t-n))
\]
When returning the active power $P$ at the current moment, the voltage at the current moment, the voltage at the historical moment, and the active power are used as input, and the return of the reactive power $Q$ is the same. The block diagram of the algorithm steps is shown in Figure 5. First, analyze the characteristics of the data when the impact load occurs, and collect the data for preprocessing; Then, transform the data set into a supervised learning problem, fully consider the time sequence information in the data, and use the current time and historical time information to establish a power prediction model at the current time; Then, input the training data and update the network parameters to obtain the trained model by minimizing the loss function; Finally, analyze the results of training data and test data to prove the effectiveness of the algorithm and the generalization of the model.

4. Experimental comparison

4.1. Analysis of SVR experiment results

In the actual calculation of SVR model parameters optimized by genetic algorithm, with the help of the LIBSVM toolbox of Matlab[16]. In the experiment, The maximum number of iterations of the optimization algorithm $t$ is selected as $t=120$, the range of parameter $C$ of the two optimization methods is set to $(0.1, 100)$, the range of parameter $g$ is set to $(0.1, 1000)$. In the calculation of GA algorithm and PSO algorithm, combined with the error requirements of the impact load data model, the fitness function is the minimum mean square error (MSE) function:

$$\text{MSE} = \sqrt{\frac{\sum_{i=1}^{k} (\hat{\rho}_i - \rho_i)^2}{k}}$$

In the formula, $k$ is the number of statistical mean square error points, $\hat{\rho}_i$ is the predicted value, and $\rho_i$ is the real sample.

The optimization and fitting results of the Grid Search traversal method are shown in Figure 6, Figure 7, and Figure 8.

---

**Fig.5  Functional block diagram for energymeteringunit**
Fig. 6  Grid Search traversal method

Grid Search finds results in a fixed interval through a certain step length traversal. The x and y axes are parameter intervals, representing the values of $\gamma$ and C, the z axis represents the value of MSE, and the step size is set to 0.5. It can be seen that in a certain range, the $\gamma$ value is too large, and the C value is too small, which will lead to very poor fitting accuracy. Moreover, the fitting effect found by the model under the fixed step size of this area is not ideal, and the fitting effect of the area beyond the parameter is also unknown.

Fig. 7  Grid Search active power fitting result

Fig. 8  Grid Search reactive power fitting results
Using heuristic search method, where the range of parameter C is set to (0.1, 100), and the range of parameter g is set to (0.1, 1000). This range is much larger than the traversal search interval, and there is no step size limit.

In addition, PSO iterated quickly to a better parameter interval. After 120 iterations, the optimal effect was found, and the parameters were also in a more reasonable interval.

The final fitting results of active and reactive power are shown in Figure 11 and Figure 12.
The results of the SVR model parameters selected by the GA-SVR and PSO-SVR algorithms and the comparison results of the mean square error value with the Grid Search method are shown in Table I:

| Optimization algorithm | Average number of iterations (/time) | Kernel function parameters γ | Punishment factor C | Mean square error MSE |
|-------------------------|-------------------------------------|-----------------------------|--------------------|----------------------|
| Grid Search             | /                                   | 3                           | 7.5                | 0.2152               |
| GA-SVR                  | 60                                  | 0.017                        | 41.897             | 0.0107               |
| PSO-SVR                 | 120                                 | 0.01                         | 0.1                | 0.0131               |

GA-SVR iterates to a better parameter interval uses more algebra than PSO-SVR, but has less termination algebra than PSO-SVR method. According to the analysis of the fitting results, the GA-SVR optimization results pay more attention to outliers in a certain range. Therefore, the value of the parameter C found is larger than that of the PSO method, and the fitting accuracy is slightly higher, but the generalization performance is not as good as the former.
4.2. Analysis of LSTM experiment results

The model of the LSTM modeling method is based on the Tensorflow 1.13.1 framework. Keras is the top-level package, and the programming language is Python 3.6. Using the network structure of \([1,100,100,1]\) to build a four-layer network model, the input layer contains 50 sequences, and the LSTM layer contains 100 neurons. The Dropout layer is a Dropout layer to prevent overfitting, and finally a neuron in the fully connected layer. The mean square error (MSE) is sensitive to the mutation point, and the LSTM modeling method also takes the MSE method as the model loss function. See Table 3 for specific parameter Settings.

Figure 13 shows the loss of the model on the training set and validation set. It can be seen from the figure that, due to the selection of the appropriate model structure, modeling method and model parameters, the first epoch on the training set and validation set reached 0.0098 and 0.0063, and the loss decreased rapidly under the action of appropriate optimization algorithm. After 7 epochs, the normalized losses of the training set and the validation set reached 0.00084 and 0.00037. With the increase of training times, the normalized losses reached the minimum of 0.00014 and 0.00013.

| Parameter names     | Parameter value       |
|---------------------|-----------------------|
| Network structure   | LSTM+Dropout+Dense    |
| batchsize           | 100                   |
| timesteps           | 50                    |
| epoch               | 100                   |
| learning_rate       | 0.001                 |
| beta1               | 0.9                   |
| beta2               | 0.999                 |
| epsilon             | 1e-08                 |

Fig.13 LSTM Model loss graph

Fig.14 LSTM algorithm fitting active power results
Figure 14 shows the fitting results of the model on the active power $P$ of the original data, which reflect the regression results of the model on the training set and the validation set. The predicted value of the model can well reflect the change of the actual value of the model, and the accuracy rate is 99.98%. Figure 15 shows the fitting results of the model on the reactive power $Q$ of the original data. Although the variation law of reactive power is different from that of active power, the prediction result of the model can also follow this change accurately, and the accuracy rate has reached 99.99%.

4.3. Comparison between SVM method and LSTM modeling method

In order to verify the robustness of the algorithm, a new set of measured impact load data had been selected for testing in this paper. Figure 16 and Figure 17 respectively show the regression results of the model for active power and reactive power on the new test data. It can be seen from the figure that the model still has a good regression result for unknown shocks, which proves the effectiveness of the algorithm.

It can be seen from the fitting result graph that for the SVR method in the case of small sample training, the optimal initialization parameters were obtained by introducing the kernel function and different parameter optimization methods, and a good test effect was obtained under the complex and nonlinear real measured data.

For the LSTM modeling method, the current moment information and historical moment information are fully used in the modeling. The results show that this method can well reflect the change law of active power $P$ and reactive power $Q$ with voltage when an impact occurs, and has good regression results on the training data and test data, and the algorithm has strong robustness. Relatively speaking, compared with SVR, LSTM model is more complex and has a longer running time, and over-fitting may occur. Its training requires higher computing capabilities of the equipment.

In short, whether it is SVR or LSTM, it can effectively establish its model for impact loadsto characterize the operating state of the loads when the shocks occur.
Fig. 16 Comparison of partial active power of P and its amplification results

Fig. 17 Comparison of partial reactive power of Q and its amplification results

5. Conclusion

This paper analyzes the characteristics of impact loads and expounds the impact characteristics of CT machines; it summarizes current impact load modeling methods and compares the characteristics of different modeling methods; first proposing to introduce related algorithms for machine learning into impact load modeling, using the SVR method, optimizing the parameters through the kernel function technique and two heuristic methods to fit the measured impact load data; using the LSTM (long and short-term memory) characteristics to introduce historical information to build a regression model. Finally, analyzing and comparing the modeling characteristics and regression results of SVM and LSTM to prove the effectiveness of the two algorithms.

Through the modeling results of the two methods in this paper on the measured data of CT machine, the feasibility of the machine learning method in the impact load modeling is proved, but the results of other machine learning modeling methods need to be further studied. In addition, only the measured data of CT machine were studied in this paper, and the generalization performance of machine learning modeling method on different types of impact loads also needs further experimental verification.
Acknowledgments
Supported by: Science and Technology Project of State Grid Zhejiang Electric Power Co., Ltd. (5211DS19003K)

References
[1] HE Ren-mu, WANG Ji-li, SHI Ke-qin. Analysis and Modeling of Shock Load via Measurement[J]. Proceedings of The Chinese Society for Electrical Engineering, 2010, 030(025):59-65.
[2] SHI Jing-hai. Load Modeling of Large-scale Power Grid Considering Time-variant Characteristic.[D].North China Electric Power University, 2004.
[3] LI Zhi-qiang, HAN Zhi-yong, An Ning. A Periodic Shock Load Model Incarnating Power Demand Initiative of Rolling Mill[J]. Power System Technology, 2011(12):72-76.
[4] ZHANG Xiao, JU Ping, CHEN Qian, et al. Study and Application of Load Models Considering Frequency Characteristics[J]. Journal of Hohai University (Natural Science), 2010, 038(003):353-358.
[5] TANG Yong-hong, ZHENG Xiao-yu. Research on Load Model Synthesis Based on Multi-curve Fitting Method[J]. Modern Electric Power, 2009,26(5):7-10.
[6] YANG Fu-li, HOU Xing-zhe, MA Shan-shan, WANG Xue-wei, LIU Xing-zhi. The Influence of Dynamic Load Characteristics on the Operating State of Smart Electricity Meter[J]. Electrical Measurement & Instrumentation, 2019,56(12):140-145.
[7] ZHANG Bin. Research of Electric Power Metering of Impact Loads[D]. South China University of Technology, 2014.
[8] QIN Zhen-liang. Shock load Modeling and Analysis of Power Quality[D]. Shandong University, 2014.
[9] NING Yu-bao. Research on the Impact of Large-Scale Shock Loads on the Power Grid and Key Technologies for Governance[D]. Southeast University, 2016.
[10] SUN Zhi-da, CHEN Hong-kun, HU Pan, DENG Yun-fei, ZENG Shuang. Electrified Railway Impact Load Modeling Based on Measured Data[J]. Engineering Journal of Wuhan University, 2016, 49(02): 236-241.
[11] XIE Yan. Study on the Characteristics and Load Modeling of Typical Impulsive Loads in Power Systems[D]. North China Electric Power University, 2014.
[12] Karlsson D , Hill D J . Modelling and Identification of Nonlinear Dynamic Loads in Power Systems[J]. IEEE Transactions on Power Systems, 1994, 9(1):157-166.
[13] ZHANG Hao-ran, HAN Zheng-zhi, LI Changgang. Support Vector Machine Based Nonlinear Systems Identification[J]. Journal of System Simulation, 2003(01):119-121.
[14] HOCHREITER S, SEHMIDHUBERJ. Long short-term memory[J]. Neural Computation, 1997, 9(8): 1735-1780.
[15] GRAVES A. Supervised sequence labelling with recurrent neural networks[D]. Ph. D Dissertation. Manno, Switzerland: Technical University of Munich, 2008.
[16] GU Feng, WANG Yao-hui, ZHAN Feng. Research and Simulation of Subway Traction Model Based on MATLAB[J]. Industrial Control Computer, 2017, 030(012): 39-41.