

# Research on underground cable abnormal vibration identification technology based on multi-model fusion

Congmin Wang<sup>a</sup>, Wen Xia<sup>b, \*</sup>, Yidong Yu<sup>c</sup>, Hao Zhang<sup>d</sup>, Liang Wang<sup>e</sup>

State Grid Zhejiang Electric Power Co., Ltd. Ningbo Power Supply Company, Ningbo 315000, China

<sup>a</sup>48084663@qq.com, <sup>\*</sup>, <sup>b</sup>Corresponding Author Email: dianliningbo@163.com,

<sup>c</sup>yanhuaheb@126.com, <sup>d</sup>xintonghahaha@yeah.net, <sup>e</sup>wuxiaozjnb@163.com

**Abstract.** With the full implementation of cable cabling in the urban core, thousands of kilometers of underground cables have placed new demands on cable operation management. In finding faults, fault identification by trial delivery of faulty lines section by section is inefficient and difficult, and there is no effective technical means to locate faults quickly. To address these problems, this paper constructs a multi-model fusion strategy of statistical learning + deep learning to achieve effective improvement of algorithm fitting effect and generalization ability, which can learn various types of features as comprehensively as possible, achieve differential extraction of distributed fiber optic timing features through differential feature construction techniques, and obtain spatio-temporal information of vibration events along the fiber optic cable for training to effectively solve the above problems and achieve more reliable and accurate prediction.

**Keywords:** Cable Monitoring, Pattern Recognition, Data Analysis.

## 1. Introduction

With the full implementation of the cable line cableization in the urban core area, the current core area power line into the ground rate has reached more than 95%, thousands of kilometers of underground cables bring new requirements for cable operation and management. In the actual operation process, there are many problems such as low level of informationization of power pipeline management, limited level of anti-external damage technology, and difficulty in finding fault points of faulty lines, which bring greater pressure on cable operation management.

In the prevention of external damage, the construction unit declaration, site inspection, key area control, etc., underground cable by external damage can not achieve early warning and real-time positioning; in fault finding, through the fault line section by section test delivery way to investigate, low efficiency, more power outages, difficult to investigate, no effective technical means to quickly locate the fault; in the operation and maintenance management information technology, it is difficult to open up with the existing business system; the existing comprehensive management process of underground cables is complex and lacks effective thematic graphic support, resulting in the collection results cannot effectively support the existing business system; in short, the existing collection and information management system can no longer meet the increasingly complex underground cable operation and maintenance and management needs.

In this paper, through the study of branch-type optical communication network vibration monitoring technology, the ground vibration in the region is monitored through the spare core of distribution automation communication cable with branch-type optical network access, and judged according to the characteristic value of the damage operation equipment, providing the operation personnel with accurate location and early warning, while the online decay and breakpoint detection of optical cable can be used to detect abnormalities in time and prevent the urban underground distribution lines caused by external damage. It can prevent accidents such as large-scale power outages caused by external damage to underground distribution lines in cities. At the same time, it can use the optical cable's light-transmitting characteristics to monitor the displacement of pipelines, improve the geographic information management of underground pipelines, build a smart grid system, improve the efficiency of operation and maintenance, realize the digital intelligent operation and

maintenance of power pipelines, and build a digital power supply station management system. It is mainly reflected in the following aspects:

(1) Effectively solve the current low performance and high cost problems of cable channel control. Urban power cables are widely distributed and large in number, while the current operation and management department mainly adopts the manual inspection approach. For areas where theft cases are frequent and construction is frequent, the operation and maintenance department cannot fully guarantee the safety of cable equipment even if they arrange maintenance personnel to strengthen inspections, and cannot achieve real-time monitoring and fast and efficient management in the face of abnormal conditions. This project can greatly improve the level of cable channel safety operation, save the pressure of cable channel safety inspection, and realize the transformation of urban cable channel safety control from passive defense to active warning by integrating branch-type fiber optic vibration for cable channel anti-intrusion monitoring and early warning, as well as cable channel anti-intrusion early warning prediction mechanism based on multi-source sensing perception, which can realize the digital twin panoramic control of cables in urban core areas.

(2) Effectively solve the current problem of low refinement of manual alerts. The traditional monitoring method relies on the monitoring of cable wells to achieve crude monitoring of underground cables, which cannot monitor the overall cable channel safety status in real time. With the continuous upgrading of urban construction, rail transit and municipal engineering road excavation is frequent, resulting in a number of cable channel external damage events, 10 kV cables were dug, and a large number of users lost power due to external breakage of vicious events. Traditional manual inspection and simple information monitoring can no longer meet the requirements of lean control of distribution network cables. Based on the existing distribution automation fiber optic cable construction, this project develops network-based remote control optical switch equipment by fusing suitable spare cores to form a branching vibration sensing network, upgrading fiber optic vibration sensing technology from line defense to mesh defense; providing timely early warning and construction location, thus realizing environmental monitoring of underground cables in urban core areas without dead ends throughout.

(3) Create a digital twin mode cable channel control system. Develop an operation and maintenance management system based on digital twin technology, connect sensing, mapping and underground electric (optical) cable operation and maintenance business management channels, establish a relational database with spatial data, real-time data and object data, and realize intelligent sensing and management of underground cables and optical cables. Safety equipment managers at all levels can log on to the platform to remotely monitor the status of any equipment and gain a comprehensive understanding of the operating status of all types of equipment.

## **2. Data-driven spatio-temporal feature reconstruction technology**

### **2.1 Characterization of underground fiber disturbance data**

Facilities along urban underground cables are very complex, for example, they may produce through densely populated central areas, viaducts with high traffic flow, schools, factories, etc. The amount of vibration along different geographical characteristics varies greatly, such as along the highway with more disruptive behavior, the general waveform intensity of 2000 may not even be triggered by external damage, while in the relatively secluded suburban section, the waveform intensity of 500 has been a disruptive event construction behavior. How to accurately identify which actions need to be alarmed and which actions do not need to be alarmed in the process of long-distance cable branch network safety monitoring along the fiber optic cable has become a major task of the city cable fiber optic cable early warning leak detection system.

Pattern recognition[1,2] is one of the core technologies in the early warning leak detection system for optical fiber cables of urban underground distribution lines, and the extraction of detection signal feature vectors is one of the most critical links in the pattern recognition module. This technique extracts the process that reflects the pattern characteristics from the localized disturbance signal to

provide feature vectors for classification recognition[3], and the feature vectors of the fiber optic disturbance signal need to satisfy the following three conditions:

- (1) The features need to be unique, different categories of perturbation signals should have their own unique characteristics and overlap with the characteristics of other types of perturbations.
- (2) The features need to have stability, the features will not change with the passage of time, the number of samples, the external environment and other factors, that is, the inherent characteristics of the signal are always present and stable.
- (3) The features need to have universality, i.e., different categories of events can characterize the differences between events under the same feature. Especially in distributed fiber optic warning systems, the feature data can be quantified to facilitate quantitative processing.

At this stage, most of the fiber waveform features are extracted from the waveform itself, from the time-frequency domain perspective, based on statistical features and spectral characteristics to distinguish between external damage and various types of interference events. In the time-domain dimension, due to the characteristics of the vibration signal itself, the effect on the length and refractive index of the fiber also has the characteristics of vibration. The vibrational characteristics are indirectly reflected in the phase change function, which in turn affects the coherent intensity of the backward Rayleigh scattered light.

Some of the time domain features can be intuitively felt by the naked eye, such as the maximum and minimum values of the waveform, while some features need to be obtained by certain operations, such as the average value, variance, and short-time energy of the waveform over a period of time. As shown in Table 1, the time domain features of waveforms include the following 15 features:

**Table 1.** Time domain features that can be extracted for distributed fiber waveforms

Features	Calculation formula
Maximum value (PK)	Max(X)
Minimum value (Min)	Min(X)
Peak to Peak (PK-PK)	$[\max(X)-\min(X)]/2$
Energy (E)	$10\log[\sum_{i=1}^n (x_i - \bar{x})^2]$
Average value (Mean)	$\frac{1}{n} \sum_{i=1}^n x_i$
Absolute mean (Arv)	$\frac{1}{n} \sum_{i=1}^n  x_i $
root of the mean square (RMS)	$\sqrt{\frac{1}{n} \sum_{i=1}^n x_i^2}$
Variance (var)	$\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2$
Standard deviation (SD)	$\sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2}$
Peak factor (Kur)	$\sum_{i=1}^n x_i^4 / nX_{RMS}^4$
Peak factor (Skew)	$\frac{1}{n} \sum_{i=1}^n \left( \frac{x_i - \bar{x}}{\sqrt{x_{var}}} \right)$
Gap factor (CL)	$X_{PK} / \left( \frac{1}{n} \sum_{i=1}^n \sqrt{ x_i } \right)^2$
Waveform factor (Sf)	$X_{sf} = X_{RMS} / X_{arv}$
Pulse factor (If)	$X_{If} = X_{PK} / X_{arv}$
Yield factor (Cf)	$X_{Cf} = X_{PK} / X_{RMS}$

Frequency domain characteristics is the use of short-time Fourier and other related calculation methods, the interference waveform of the time series to do frequency domain analysis, analysis of the characteristics of the low-frequency and high-frequency parts of different vibration waveforms, capture in the frequency domain third-party damage, relative to other sources of vibration waveform characteristics of the difference. Short-time Fourier transform (STFT) is a commonly used signal time-frequency analysis method, whose basic idea is to localize the integration interval of signal Fourier transform, which is an effective tool to retain signal time-domain information and frequency-domain information at the same time. The short-time Fourier transform of the sensing signal can effectively analyze the frequency information of the sensing signal in different perturbation time regions, so as to distinguish the high-frequency perturbation from the low-frequency perturbation. The fiber optic sensing signal is a finite and discrete digital signal, which needs to be transformed into the frequency domain using the discrete Fourier transform to analyze the characteristics of the frequency domain. For a segment of discrete fiber optic sensing signal  $x(i)$ , its frequency domain expression is:

$$X(k) = \sum_{i=0}^{L-1} x(i)e^{-j\frac{2\pi}{L}ik} \quad k = 0, 1, 2, \dots, L-1 \quad (1)$$

where  $L$  denotes the length of the signal that does the DFT, also known as the DFT at point  $L$ . The larger its value the higher the frequency resolution of the signal, and  $x(k)$  denotes the value of the  $k$ th frequency point on the spectrum.

The maximum, minimum, median, mean and peak difference of amplitude corresponding to the frequency spectrum are extracted, while the frequency, center of gravity frequency, mean square frequency and root mean square frequency, frequency variance and frequency standard deviation corresponding to the peak of amplitude are obtained. The calculation formula is shown in Table 2.

**Table 2.** Frequency domain feature extraction extremely calculation formula

Frequency domain characteristics	Calculation formula
Amplitude maximum	$\max(\text{fft})$
Amplitude minimum	$\min(\text{fft})$
Minimum amplitude	$\text{median}(\text{fft})$
Average amplitude	$\frac{1}{n} \sum_{i=1}^n \text{fft}_i$
Amplitude peak difference	$[\max(\text{fft})-\min(\text{fft})]/2$

However, under these existing methods, because of the various facilities along the fiber optic cable, geological and disturbing events, these traditional feature extraction schemes are difficult to portray the differences between third party damage and strong disturbing events such as heavy vehicles, traffic honking and factory roaring in complex disturbing environments. It is a very important task to rationalize the use of time-domain features to achieve a high degree of distinction between third-party sabotage and various types of strong interference events.

It is worth mentioning that the geographic information along the underground cables has not been reasonably utilized in any existing technical applications, while the geographic information of the defense zones actually contains a lot of useful data, such as suburban areas and school sections with less activities in general, where the invasion waveforms of third-party sabotage events are clearer, while those along highways and intersections are accompanied by great disturbances. How to reasonably introduce the geological and geological information of each defense zone is one of the effective ways to provide highly accurate third-party damage identification for the whole line section.

Here, this technical solution proposes a data-driven spatio-temporal feature reconstruction technology to maximize the extraction of temporal features that can characterize actual third-party sabotage and various types of interference events, learn the geographic correlation of fiber optic cable

defense zones, and help achieve high accuracy in preventing third-party sabotage of fiber optic cable along the whole line.

## 2.2 Time domain feature reconstruction based on differentiation analysis

It is difficult to use signal data directly for classification because the data itself carries information about the characteristics of the perturbed signal, but the data also contains a lot of information that has commonality and is not valuable for signal classification. Therefore, extracting effective features as target samples for classification can help improve the reliability of the discriminative model.

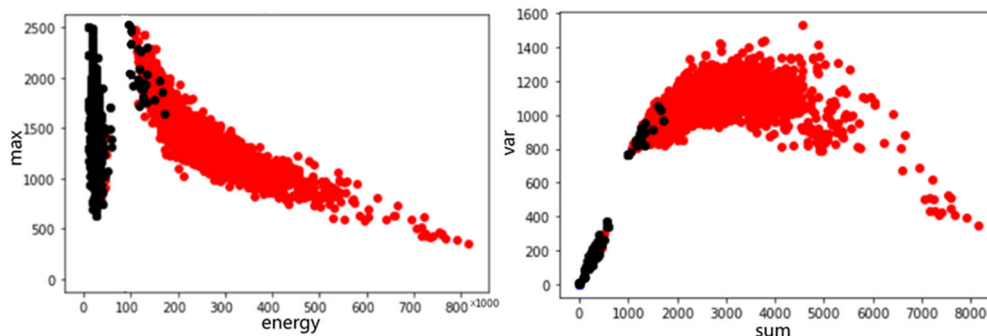
The original signal after the difference algorithm can be based on the amplitude threshold range of the vibration signal for preliminary classification prediction[4]: the vibration signal of the car driving and excavator digging vibration signal amplitude than the personnel walking and manual shovel digging amplitude, personnel walking signal amplitude convex change is more continuous, conforming to the pace when walking on the fiber landing, and gradually enhance the amplitude with time transformation; car driving in gradually approaching the fiber buried position in the process of optical fiber can be sensed through the ground propagation of the vibration signal of the car driving, generating part of the car vibration signal, but the amplitude is relatively low, and through the buried fiber when the time is short, fast, from the waveform can be observed to leave the signal transient fluctuations, with impact signal, high amplitude; shovel excavation signal envelope transformation is fast, amplitude changes show regularity, in line with the frequency of manual excavation. Excavator excavation because of the wide range of vibration, excavator start will have a vibration signal, to the excavator began to dig generated by the amplitude are relatively large, the time domain waveform changes clearly; and the noise signal amplitude is small, the waveform without obvious envelope changes, adjacent signal points of the great minimal value without convex change phenomenon, always in a smooth state. In summary, different intrusion vibration signals can be distinguished by waveform change characteristics and amplitude energy of different vibration signals.

## 2.3 Variability in Episodic Spike Event Characteristics

However, in the distributed fiber optic often receive some occasional spikes of interference waveform[5]. At present, because many long-distance oil and gas cables are laid along the road, even along the elevated, highway, the flow of traffic on the road will give rise to a lot of vibration interference, especially when the car idle vibration, multi-vehicle sirens, etc. will bring a lot of spike interference to the fiber optic sensing; in the township along the line, some firecrackers, fireworks will also bring a burst of episodic spike waveform; in the muddy In the muddy road or near the factory, some heavy trucks loading goods, traffic start and stop or road skidding, its episodic spike phenomenon is also very obvious.

## 2.4 Differentiated Features Reconfiguration

Through the analysis of large batch data, we propose a truncated difference feature reconstruction method based on the important characteristic that most of the episodic spikes within 120 units of data are within 5-10. The innovative normalization scheme selects the 10th largest value of the waveform as the upper limit, and the exceeding part is uniformly weakened as the upper limit to effectively avoid shock interference events. At the same time, for the characteristics of frequent low noise activities in some release sections, the base noise is set, and the part below the base noise is deflated and uniformly set to zero value.



(a) Energy value and maximum value distribution chart (b) Sequence sum and variance distribution graph

**Figure 1.** Distribution of differential characteristics

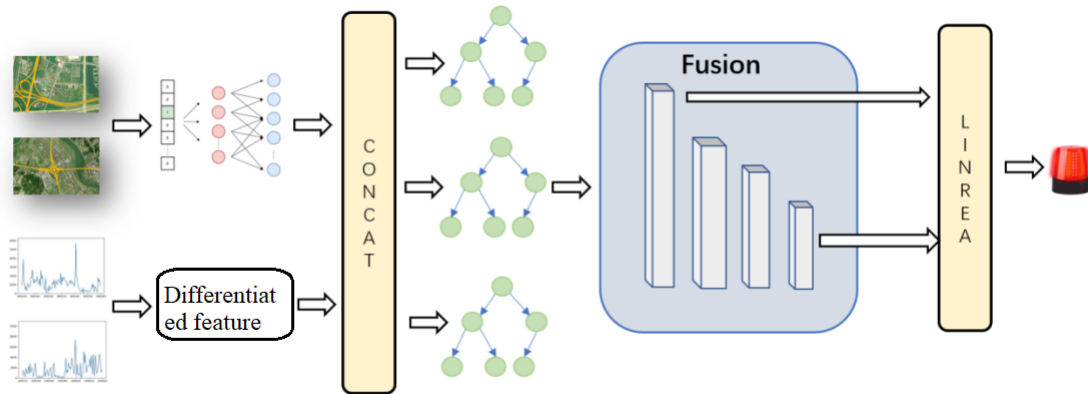
The features such as energy value, maximum value, variance value, sequence sum, etc. of the truncated improved waveform are selected as time domain features. The data features extracted after truncation are displayed and analyzed by data visualization techniques, as shown in Figure 1. Such as the statistical distribution of variance features and sequence sum after truncation, the initial distinction between third-party damage events and strong interference events is achieved, which provides a better quality of time series data feed for subsequent model identification.

### 3. Intelligent recognition algorithm based on multi-model fusion

#### 3.1 Multi-model fusion training

Due to the existence of complex human activity noise along the cable, which increases the recognition difficulty of the model, it is difficult to achieve accurate recognition of each third-party damage by using a simple single model under multiple noise interference. The use of a single recognition algorithm will often be difficult to achieve accurate recognition of the state of certain situations, the lack of learning efforts, there will be a certain recognition of blind areas, and the use of a variety of different principles of machine learning methods, can effectively reduce the recognition of blind areas, to improve the reliability of the overall prediction results. A single model often appears to be more sensitive to some categories of data features and can achieve relatively good classification, while it is difficult to classify a certain type of data, at this time this type of data is not distinguishable in the model prediction, and it may present a random classification result, that is, the judgment classification result is any of the three, and the prediction result accuracy is low; multiple models with low distinguishability tend to have different types of data. Therefore, machine learning algorithms with multiple different mechanisms can learn all kinds of features as comprehensively as possible to effectively solve the above problems and achieve more reliable and accurate prediction.

This technique adopts the training method of single model mentioned above in the structure with the recognition technique of neural network, and builds the multi-model fusion strategy of statistical learning + deep learning to realize the effective improvement of the fitting effect and generalization ability of the algorithm. The multi-model fusion process is shown in Figure 2. Firstly, the geographic information of defense zones along the fiber optic cable obtained in the previous subsection is learned by graph embedding to obtain the geographic association information of each defense zone, and then the differential feature construction technique realizes the differential extraction of distributed fiber optic timing features to obtain the spatio-temporal information of vibration events along the fiber optic cable. The spatio-temporal information is given to each model as the input of the model for training and learning, so as to achieve timely and accurate identification of third-party damage under multi-noise interference.



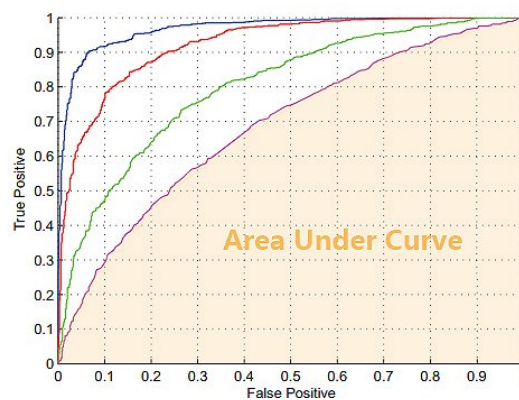
**Figure 2.** Schematic diagram of multi-model fusion technology process

The base learners in the first layer eventually yield new test data sets as well as new training data sets, which will be applied in the training of the second layer model, so the input in the second layer is the evaluation result of each base model output in the first layer, and the output is the adjusted final evaluation result. The second layer model uses a fully connected cascade neural network with the network structure shown in Figure 2. Five base learners are connected to five neurons through the input layer, and the first four neurons use the  $\tanh()$  activation function, and the last neuron is a linear summation function. With the same number of neurons, the FCC neural network can provide more connections and get more weight relationships than the traditional neural network, and through such structural advantages, the FCC neural network can obtain better results.

The shallow structured FCC[6] neural network used in this technique is able to provide sufficient learning capability while avoiding the over-fitting problem. the FCC neural network has a direct mapping relationship between each input and the latent variables of each neuron.

### 3.2 Model evaluation index and identification effect

In order to balance the recognition accuracy of the algorithm recognition and to minimize the systematic underreporting problem on the basis of the recognition accuracy, we use AUC [7] (Area Under Curve) as our model evaluation metric here, which is defined as the area under the ROC curve. We tend to use the AUC value as a model evaluation criterion because often the ROC curve does not clearly indicate which classifier is more effective, and as a value, the classifier corresponding to a larger AUC is more effective.



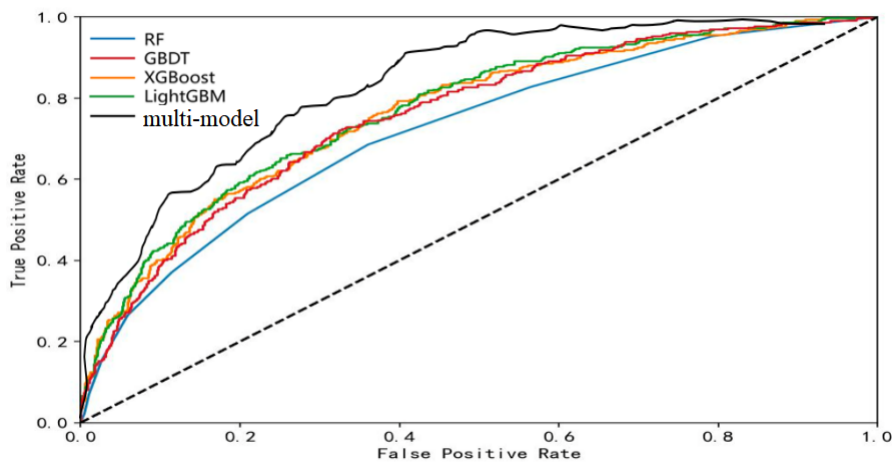
**Figure 3.** ROC curve and AUC area

The ROC curve and AUC area are shown in Figure 3. The ROC curve is a curve based on a series of different dichotomies (cut-off values or decision thresholds) with the true positive rate (sensitivity) as the vertical coordinate and the false positive rate (1-specificity) as the horizontal coordinate. the horizontal coordinate of the ROC curve is the False Positive Rate (also called False Positive Rate),

the vertical coordinate is the True Positive Rate (True Positive Rate), and correspondingly the True Negative Rate (True Negative Rate) and the False Negative Rate (False Negative Rate). These four types of indicators are calculated as follows.

False positive rate (FPR): the probability of a positive case but not a true case, i.e., the probability of a positive case among true negative cases; true positive rate (TPR): the probability of a positive case but also a true case, i.e., the probability of a positive case among true cases (i.e., positive case recall rate); false negative rate (FNR): the probability of a negative case but not a true negative case, i.e., the probability of a negative case among true cases; true negative rate (TNR): the probability of a negative case but also a true negative case, i.e., the probability of a negative case among true negative cases.

For the practical difficulties of interference of multiple noises along urban cables, this technical study collects data on third-party damage events along cables, strong interference such as highway and factory noise, and static non-disturbance to build a sample library with large data size. In the local training, 15000 dangerous intrusion samples, 15000 strong disturbance samples, 15000 non-disturbance samples were trained, and the final comprehensive recognition algorithm was obtained after parameter tuning. Compared with the conventional single model training for effect, the comparison of AUC indexes of each scheme is shown in Figure 4, and the recognition effect of this scheme is substantially improved compared with the general recognition scheme.



**Figure 4.** Comparison of AUC metrics between multi-model fusion technology and other solutions

A comparison of the identification specific AUC metrics for each model is shown in Table 3.

**Table 3.** Comparison of recognition effect of each scheme

Classifier	Recognition effect (AUC)
RF	0.8212
GBDT	0.8832
XGBoost	0.9286
LightGBM	0.9153
Multi-model fusion	0.9667

This multi-model fusion recognition technology solution in the cable line of fiber optic data for a long time short recognition verification, it can be seen that this technology solution can achieve excellent recognition effect.

We get the vibration waveform data for the presence of third-party damage along the cable. It can be seen that this multi-model fusion algorithm achieves timely and accurate identification and determination of third-party sabotage events. We also get the vibration waveform data of a strong disturbance event without third-party damage, but triggered by the plant. The green line in the figure represents the distributed fiber waveform data for the current time period, which is shown to be in a

large jitter for a long time; the orange line model recognition results, the orange line at the top indicates that the model recognizes this moment as the moment of intrusion, and at the bottom it is considered that there is no damage event at this time. It can be seen that this multi-model fusion algorithm reduces the false alarm situation as much as possible under the strong interference event without actual damage.

#### 4. Conclusion

In this paper, through the study of branch-type optical communication network vibration monitoring technology, the ground vibration in the region is monitored through the spare core of distribution automation communication cable with branch-type optical network access, and judged according to the characteristic value of the damage operation equipment, providing the operation personnel with accurate location and early warning, while the online decay and breakpoint detection of optical cable can be used to detect abnormalities in time and prevent the urban underground distribution lines caused by external damage. It can prevent accidents such as large-scale power outages caused by external damage to underground distribution lines in cities. At the same time, it can use the optical cable's light-transmitting characteristics to monitor the displacement of pipelines, improve the geographic information management of underground pipelines, build a smart grid system, improve the efficiency of operation and maintenance, realize the digital intelligent operation and maintenance of power pipelines, and build a digital power supply station management system.

#### Acknowledgments

This work is supported by State Grid Zhejiang Electric Power Co., Ltd. Ningbo Power Supply Company Innovation Consulting Project "Safety pre-warning and prediction technology of underground cable combining optical fiber vibration sensor and urban digital twin and its application", Ningbo Science and Technology Special Project of China under Grant No. 2022Z095.

#### References

- [1] Min WANG, Jiantao ZHENG, Teng ZHOU. Research on pattern recognition technology for a highperformance optical fiber vibration sensor. *Journal of Physics: Conference Series*,2021,2005(1):012110.
- [2] Li J, Wang Y, Wang P, et al. Pattern recognition for distributed optical fiber vibration sensing: A review. *IEEE Sensors Journal*, 2021, 21(10): 11983-11998.
- [3] Zhou Zichun, Liu Kun, Jing Junfeng, Xu Tianhua, Wang Shuang, Sun Zhenshi, Guo Hairuo, Liu Tiegeng. Optical Fiber Vibration-Sensing Event Recognition Based on CLDNN. *ACTA OPTICA SINICA*,2021,41(13):1306019.
- [4] Wang Y, Wang P, Ding K, et al. Pattern recognition using relevant vector machine in optical fiber vibration sensing system. *IEEE Access*, 2019, 7: 5886-5895.
- [5] Lyu C, Huo Z, Cheng X, et al. Distributed optical fiber sensing intrusion pattern recognition based on GAF and CNN. *Journal of Lightwave Technology*, 2020, 38(15): 4174-4182.
- [6] Liu X, Wang N. A novel gray wolf optimizer with RNA crossover operation for tackling the non-parametric modeling problem of FCC process. *Knowledge-Based Systems*, 2021, 216: 106751.
- [7] Luo D, Zhou C X. A Brief Discussion about Application of Artificial Intelligence in Computer Network Technology in the Era of Big Data. *Journal of Physics Conference Series*, 2020, 1684:012001.