applied sciences

MDPI

Article

Fault Prediction for Rotating Mechanism of Satellite Based on SSA and Improved Informer

Qing Lan [1,2,3], Ye Zhu [1,2,3,*], Baojun Lin [1,2], Yizheng Zuo [1,2,3] and Yi Lai [1,2,3]

[1] Innovation Academy for Microsatellites of Chinese Academy of Sciences, Shanghai 201304, China
[2] University of Chinese Academy of Sciences, Beijing 101408, China
[3] Key Lab for Satellite Digitalization Technology of Chinese Academy of Sciences, Shanghai 200031, China
* Correspondence: zhuye@microsate.com

Abstract: The rotational mechanism, which plays a critical role in energy supply, payload antenna pointing, and attitude stabilization in satellites is essential for the overall functionality and performance stability of the satellite. This paper takes the space turntable of a specific satellite model as an example, utilizing high-frequency and high-dimensional telemetry data. An improved informer model is used to predict and diagnose features related to the turntable's operational health, including temperature, rotational speed, and current. In this paper, we present a forecasting method for turntable temperature data using a hybrid model that combines singular spectrum analysis with an enhanced informer model (SSA-Informer), comparing the results with threshold limits to determine if faults occur in the satellite's rotational mechanism. First, during telemetry data processing, singular spectrum analysis (SSA) is proposed to retain the long-term and oscillatory trends in the original data while filtering out noise from interference. Next, the improved informer model predicts the turntable temperature based on the mapping relationship between the turntable subsystem's motor current and temperature, with multiple experiments conducted to obtain optimal parameters. Finally, temperature thresholds generated from the prediction results are used to forecast faults in the rotational mechanism over different time periods. The proposed method is compared with current popular time-series prediction models. The experimental results show that the model achieves high prediction accuracy, with reductions of at least 10% in both the MAE and MSE than CNN-LSTM, DA-RNN, TCN-SE and informer, demonstrating the outstanding advantages of the SSA and improved informer-based method in predicting temperature faults in satellite rotational mechanisms.

Keywords: rotation mechanism; SSA; informer; time-series prediction

Citation: Lan, Q.; Zhu, Y.; Lin, B.; Zuo, Y.; Lai, Y. Fault Prediction for Rotating Mechanism of Satellite Based on SSA and Improved Informer. *Appl. Sci.* **2024**, *14*, 9412. https://doi.org/10.3390/app14209412

Academic Editor: Angelo Luongo

Received: 20 August 2024
Revised: 4 October 2024
Accepted: 13 October 2024
Published: 15 October 2024

1. Introduction

Satellites are intricate systems integrating various instruments, such as optical, mechanical, and electronic devices, and they are characterized by their large scale and high complexity. Once a satellite is in orbit, ground-based intervention capabilities are extremely limited. Furthermore, the deteriorating space environment and numerous uncertainties present additional challenges. With the advancement of commercial space endeavors and the reduction in production costs, the number of satellite launches is expected to increase. However, if a satellite's operational status deviates from expectations while in orbit, effective maintenance is generally challenging. This can lead to system anomalies or even failures. Therefore, satellite fault prediction is crucial for enhancing the reliability and safety of satellites and has become a primary research focus in the aerospace field [1].

The satellite rotation mechanism is a critical component of the satellite attitude control system and is extensively utilized in satellite power management, payload operations, and attitude control [2]. The inertial execution mechanisms, which mainly include rotor-based flywheels and torque control gyroscopes, are essential for stabilizing the satellite's attitude. Acting as a bridge between the satellite's payload and platform modules, the rotation mechanism enables smooth, stable high-speed rotation while transmitting power

and signals. This ensures the satellite's stability during flight. However, the complexity of the rotation subsystem and the harsh operating environment pose significant fault risks. Such faults can compromise the satellite's stability, affecting both mission success and operational lifespan.

To enhance the reliability and performance of satellite rotation mechanisms, significant progress has been made in the research of satellite fault diagnosis and prediction. Fault diagnosis techniques can be broadly categorized into knowledge-based methods, model-based methods, and data-driven methods. Knowledge-based methods are relatively difficult to implement and have not been widely adopted. Z. Shi et al. [3] proposed a model-based fault diagnosis method. However, this approach is computationally complex and challenging to model, especially in the context of big data, where its shortcomings are further amplified. Consequently, data-driven fault diagnosis techniques have become the most widely applied methods today. Yin et al. [4] introduced a Just-In-Time Learning Data-Driven (JITL-DD) fault detection method. Zhang et al. [5] proposed an improved Probabilistic Principal Component Analysis (PPCA) nonlinear data-driven process monitoring method. With the further development of data-driven fault diagnosis techniques, they can be broadly divided into statistical methods, shallow learning, and deep learning.

Statistical-based fault diagnosis techniques include Qualitative Trend Analysis (QTA) [6], Principal Component Analysis (PCA) [7], and Partial Least Squares (PLS) [8], among others. These methods eliminate the complex modeling processes of traditional methods and require minimal prior knowledge. Shallow learning fault diagnosis methods include Support Vector Machines (SVMs) [9], k-Nearest Neighbors (kNNs) [10], Gaussian Mixture Models (GMMs) [11], Artificial Neural Networks (ANNs) [12], and similar techniques. Corresponding hybrid improvement methods also demonstrate good performance. Some researchers in the satellite field have attempted to combine neural networks with statistical methods or model-based methods [13]. While statistical and shallow learning-based fault diagnosis methods have made significant advancements over traditional fault diagnosis algorithms, these methods often require domain-specific knowledge to extract fault features, and their recognition accuracy still needs improvement.

In recent years, the rapid development of machine learning and artificial intelligence has made deep learning a popular topic. Deep learning, characterized by network structures with multiple hidden layers, has excellent feature extraction and learning capabilities. It can effectively analyze and extract hidden feature information from large amounts of data, establishing mapping relationships between data and features. Deep learning algorithms overcome many of the inconveniences of traditional algorithms and have been successfully applied in fields such as image, text, and video analysis [12]. Gradually, many researchers have attempted to use predictive methods for fault detection. Liu L et al. [14] employed Gaussian Process Regression to detect sensor anomalies in aircraft engines, demonstrating good predictive performance. S. Ghasemi et al. used the Extended Kalman Filter method for fault diagnosis and isolation in spacecraft attitude control systems [15]. The advantage of this method lies in its ability to estimate both the system state and the uncertainty of the state. Other researchers have used system identification methods to predict target variables. Lu et al. [16] proposed a model combining wavelet approximation with system identification theory for anomaly detection in network signals. Chen et al. [17] utilized an Autoregressive with Exogenous Inputs (ARX) model for anomaly detection in wireless communication network data systems. Ouadine et al. [18] combined neural networks with the Hammerstein–Wiener system identification model to estimate the state of a quadrotor system and establish a set of residuals to detect sensor anomalies. AutoEncoder, a type of artificial neural network, can learn effective representations of input data through unsupervised learning, thus performing data dimensionality reduction. Zhou et al. [19] used deep autoencoders for anomaly detection. However, relevant studies seldom address telemetry data similar to those of in-orbit satellites, nor do they consider seasonal trends, periodic trends, noise, and the limitations of on-board real data sampling frequency [20]. Additionally, most of these methods simply stack different networks sequentially, which

can lead to some degree of feature loss when further extracting features across different networks. Therefore, although these algorithms may show performance improvements compared to previous algorithms, they still require further refinement due to the potential loss of information.

A novel approach for fault prediction in space rotational mechanisms based on telemetry data is proposed. By analyzing the causes, manifestations, and influencing factors of rotational mechanism faults, a hybrid signal processing method combined with a self-attention network for fault prediction is introduced. This approach is better suited to addressing practical engineering issues, such as the operational states of satellite rotational mechanisms under special working conditions and the potential noise interference during telemetry data transmission.

This paper proposes a fault prediction method for satellite rotation mechanisms, integrating singular spectrum analysis (SSA) and an improved informer model. Telemetry data from the satellite turntable subsystem were analyzed. First, min-max normalization mapped the data to the [0,1] range. SSA was then used for data decomposition to enhance prediction accuracy. The dataset was trained using the optimized informer model, and key parameter impacts on prediction were evaluated. Through experiments, the best parameter combination was identified to maximize CNN-Informer performance. Predicted data were used to generate thresholds for fault detection, and results were validated against RNN, LSTM, and transformer models, demonstrating improved prediction accuracy.

The remainder of this paper is structured as follows. In Section 2, we introduce the working principles and background of the rotating mechanism involved in fault prediction. Section 3 presents the signal processing and fault prediction methods proposed in this paper. Section 4 describes the simulation experiments and analysis. Finally, Section 5 provides the conclusion of this paper.

2. Background

The space turntable is a crucial component for realizing payload functions. Its reliability not only affects the normal execution of payload functions but also plays a significant role in maintaining the satellite's attitude stability due to the large inertia components it drives and the despinning functions it performs [21]. The satellite turntable subsystem is an essential part of the satellite spatial rotation mechanism that is used to control the direction and angle of the satellite turntable to ensure that the satellite antenna is aligned with the desired ground station. The turntable subsystem primarily consists of the turntable body, the turntable electronic control box, balance wheels, the locking and release mechanism, and supporting software. The structural form of the turntable body is shown in Figure 1. The base is thermally insulated and mounted on the top surface of the satellite platform module, with the bearing housing encased by the payload module. The support shaft system uses a pair of back-to-back angular contact ball bearings, preloaded during installation. The shaft system employs a parallel configuration with bearings located outside the entire turntable. Inside the bearing space, from top to bottom, conductive slip rings and torque motors are sequentially installed.

The satellite rotation mechanism links the payload and platform modules, enabling smooth, high-speed rotation while transmitting power and signals. The platform module supplies power and commands to the payload via the slip ring assembly, while the turntable subsystem sends rotational angle data back to the payload manager. Figure 2 illustrates these connections and control relationships.

The stability of the turntable temperature is crucial for ensuring the effective operation of the satellite rotation mechanism. It is influenced by several factors, including the turntable motor current, the satellite's attitude and orbit, and the position of the sun. Firstly, the different operational modes of the entire satellite necessitate varying operational modes for the turntable, resulting in different operating conditions. During these conditions, the turntable motor generates a certain amount of current, which in turn produces heat within the motor. Excessive current or poor motor cooling can lead to an increase in the

turntable temperature [22]. Secondly, the satellite's attitude and orbit also impact the turntable temperature. Adjustments and movements of the satellite's attitude in space impose different thermal loads on the turntable under various operating states. Therefore, when studying turntable system faults, it is crucial to consider the satellite's attitude characteristics and adjust the turntable's temperature thresholds to maintain an appropriate temperature under various attitude and orbital conditions. Additionally, the position of the sun significantly affects the turntable temperature. Solar radiation is one of the main heat sources, and direct exposure to the turntable can cause its temperature to rise. Consequently, the angle and intensity of solar exposure need to be considered in different satellite missions and working environments.

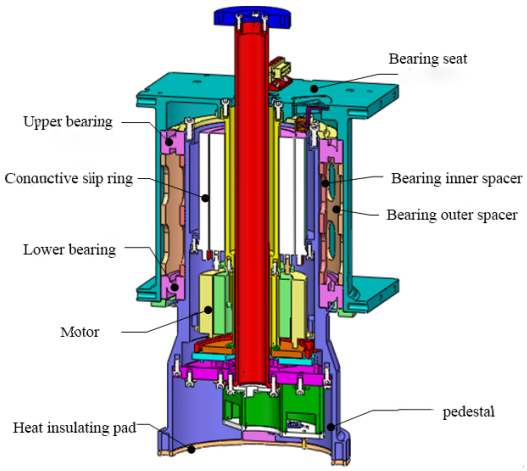


Figure 1. Structure ofturntable body.

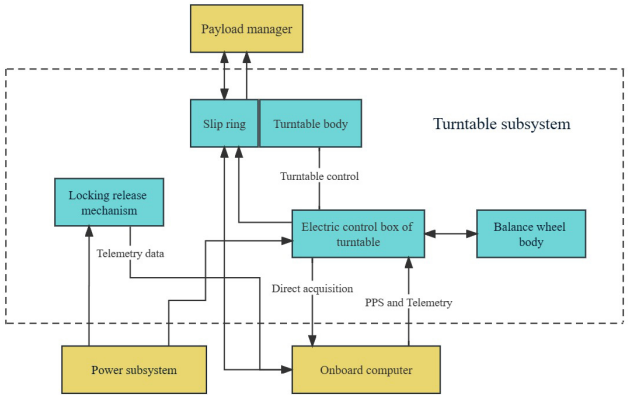


Figure 2. Satellite payload connection and control relationship.

Among these factors, the current of the turntable motor is relatively easy to detect and analyze numerically. Hence, this study aims to quantify the variations in the turntable motor current and analyze its impact on the turntable temperature. Based on the current variation patterns, this paper predicts turntable temperature faults to enable timely adjustments and repairs. By thoroughly investigating the relationship between turntable current and temperature, this study allows us to better understand the operational state and performance characteristics of the turntable and anticipate potential temperature faults in advance. This enables effective control and management of the turntable temperature, which is crucial for enhancing the reliability and performance of the turntable and ensuring the normal operation of the satellite.

3. Methods

3.1. SSA

This paper employs the singular spectrum analysis (SSA) algorithm to decompose and reconstruct satellite turntable telemetry data. SSA is a time-series analysis method based on Singular Value Decomposition (SVD) [23]. The core idea of SSA is to decompose a time series into components of different frequencies. This allows for noise removal and helps identify periodic patterns in the signal, achieving effective denoising [24]. Through the analysis of eigenvectors, it is possible to determine the frequency, amplitude, and phase of the periodic components present in the signal, thereby revealing the periodic variations of the signal. Generally, SSA is divided into two parts: matrix decomposition and series reconstruction [25].

The analysis object of SSA is one-dimensional time series with finite length. Firstly, we need to select the appropriate window length to arrange the original time series with delay to obtain the trajectory matrix:

Step 1: Decomposition

The original time series is mapped into a vector sequence of length L to form $i + L - 1$ vectors of the following length:

$$\mathrm{X_i} = (x_i, \ldots, x_{i+L-1})^{\mathrm{T}} \quad (1 \le \mathrm{i} \le \mathrm{K}) \tag{1}$$

These vectors form a trajectory matrix, as follows:

$$\mathbf{X} = \begin{bmatrix} X_1 & X_2 & \cdots & X_{i-L+1} \\ X_2 & X_3 & \cdots & X_{i-L+2} \\ \vdots & \vdots & \ddots & \vdots \\ X_L & X_{L+1} & \cdots & X_i \end{bmatrix} \tag{2}$$

The trajectory matrix $\mathbf{X}$ is decomposed of singular value decomposition. By calculating $\mathbf{X}\mathbf{X}^{\mathrm{T}}$ and performing Singular Value Decomposition (SVD), L eigenvalues are obtained, as follows:

$$\mathbf{X} = \mathbf{X}_1 + \ldots + \mathbf{X}_d \tag{3}$$

$$\mathbf{X_i} = \sqrt{\lambda_\mathrm{i}}\mathrm{U_i}\mathrm{V_i^T} \tag{4}$$

where λ_i is the eigenvalue of the matrix $\mathbf{X}\mathbf{X}^{\mathrm{T}}$; $\mathrm{U_i}$ and $\mathrm{V_i}$ are orthogonal eigenvectors of the trajectory matrix $\mathbf{X}$.

Step 2: Reconstruction

Reconstruction involves grouping the sub-matrices obtained from SVD to form several sub-matrices and calculating the contribution rate of each sub-matrix to determine its variability. The sub-matrices include high-frequency sub-matrices, low-frequency sub-matrices, and noise sub-matrices [26]. These sub-matrices are then restored to a reconstructed sequence with a shape similar to the original signal through diagonal averaging. The contribution rate can be expressed as follows:

$$\eta = \sum_{i=1}^{r} \lambda_i / \sum_{i=1}^{d} \lambda_i \tag{5}$$

The sub-matrices obtained by grouping are reconstructed into $L \times K$ matrix $\mathrm{X_{ij}}$ via the diagonal average formula, and the expression of reconstructed time series Y is as follows [24]:

$$\tilde{y} = \begin{cases} \frac{1}{k}\sum_{p=1}^{k} X^*_{p,k-p+1} (1 \le k \le L^*) \\ \frac{1}{L^*}\sum_{p=1}^{L^*} X^*_{p,k-p+1} (L^* \le k \le K^*) \\ \frac{1}{N-k+1}\sum_{p=k-K^*+1}^{N-K^*+1} X^*_{p,k-p+1} (K^* \le k \le N) \end{cases} \tag{6}$$

where $L^* = \min\{L, K\}$, $L^* = \min\{L, K\}$, $N = L + k - 1$, the sum of the sequences obtained after reconstruction is the total time series after singular spectrum analysis decomposition.

3.2. CNN-Informer

In the proposed model of this paper, the informer is an improved version based on the transformer framework [27]. This framework is specifically designed to address the prediction of long time series. In this paper, a Convolutional Neural Network (CNN) is introduced before the overall informer framework. The CNN primarily includes convolutional layers, pooling layers, and fully connected layers. Due to its characteristics of weight sharing and spatial pooling, the parameters and complexity of the network model are significantly reduced [28]. The architecture diagram of the CNN-Informer algorithm is proposed in Figure 3. Convolution operation is the core of CNNs. It utilizes sliding convolution filters, which move vertically and horizontally over the two-dimensional degraded data to refine them.

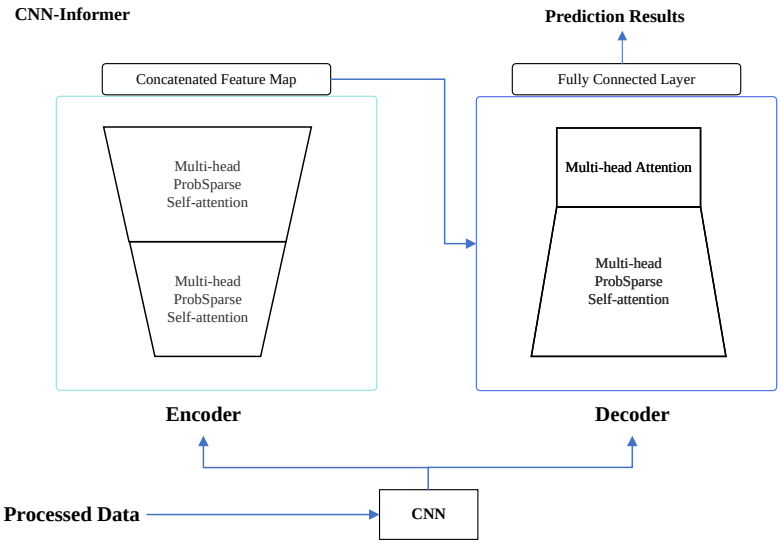


Figure 3. CNN-Informer architecture diagram.

The mathematical expressions for convolution and activation operations can be represented as follows:

$$\begin{aligned} V_{ij}^n &= \sum A^n \oplus xf_{ij} + B_n \\ Y^n &= \sigma(V^n) \end{aligned} \tag{7}$$

where V_{ij}^n is the output of convolution operation, n represents the n-th feature map, i and j correspond to the number of steps of convolution filter in vertical and horizontal directions, which A^n are convolution kernel matrix, xf_{ij} is the filter matrix, B_n is the bias term and σ is the activation function.

The Encoder captures the input sequence and maps it to a higher-dimensional vector, which is then fed into the Decoder to generate the output sequence. In this paper, the Encoder is used to learn long-term dependencies from the operational data records of circuit devices. Before the input sequence enters the Encoder, positional and temporal codes are incorporated into the input sequence, fully utilizing the positional and temporal information embedded within it. Sine and cosine functions are employed to derive positional encoding [27].

$$\begin{aligned} PE(t, 2l) &= \sin\left(t/10,000^{2l/m}\right) \\ PE(t, 2l+1) &= \cos\left(t/10,000^{2l/m}\right) \end{aligned} \tag{8}$$

where t is the position of the word in the sequence, l is the dimension of the word embedding vector, m is the size of the word embedding projection, and PE is the positional encoding.

The Probspare self-attention layer is designed to capture dependencies between features. It does not calculate the dot product of each vector in thw Q matrix and each vector in the K matrix. It calculates the sparsity measure of any Q vector according to the query sparsity measure Formula (9) and selects the top-u vectors which play a leading role in attention calculation, using these u vectors for calculations. The $PE(t, 2l+1)$ specific formula is as follows [29]:

$$M(q_i, K) = \ln L_K + \max_j \left\{ \frac{q_i k_j^T}{\sqrt{d}} \right\} - \frac{1}{L_K} \sum_{j=1}^{L_K} \left(\frac{q_i k_j^T}{\sqrt{d}} \right) \tag{9}$$

$$\text{Attention(Q,K,V)} = \text{softmax}\left(\frac{\bar{Q} K^T}{\sqrt{d_h}} \right) \text{V} \tag{10}$$

where Q is a sparse matrix of the same size of q, and it only contains the Top-u queries under the sparsity measurement $M(q, K)$.

The Decoder mainly includes masked multi-head probabilistic self-attention and multi-head attention. The former receives the following tag vectors [29]:

$$X_{j+1}^t = \text{MaxPool}\Big(\text{ELU}\Big(\text{Conv1d}\Big[X_j^t \Big]_{AB} \Big) \Big) \tag{11}$$

where $[\cdot]_{\text{AB}}$ represents the attention block. It contains the Multi-head ProbSparse self-attention and the essential operations, where Conv1d() performs a 1-D convolutional filters (kernel width = 3) on a time dimension with the ELU() activation function.

Multiple Probsphere self-attention is a stack of several Probsphere self-attention layers. Generally speaking, there are multiple multi-head Probspare self-attention layers in the Encoder. When the input data flows from the ith bullth Probsphere self-attention to the $i+1$ th bullth Probsphere self-attention, the data are processed according to the following equation [28]:

$$X_{\text{feed}}^t = \text{Concat}\big(X_{\text{token}}^t, X_o^t \big) \in R^{\left(L_{\text{token}} + L_y \right) \times d} \tag{12}$$

where X_{token}^t is the start token, and X_o^t is a placeholder for the target sequence.

3.3. Threshold Generation

After obtaining the prediction sequence, dynamic thresholds for each time step are generated based on the quartile threshold generation method. The quartile threshold generation method replaces the mean and standard deviation used in traditional statistical methods with the median and standard interquartile range [30], respectively. The statistical parameters in the quartile method include the data volume (N), the lower quartile (Q1), the upper quartile (Q3), and the interquartile range (IQR). The anomaly detection of the reaction wheel speed is performed by comparing the predicted data with the generated thresholds. The upper and lower limits of the threshold for the (i)-th step of the prediction sequence are set by the formulas represented as

$$up_i = Q\_75 + 3 * (Q\_75 - Q\_25) + output_i \tag{13}$$

$$low_i = Q\_25 - 3 * (Q\_75 - Q\_25) + output_i \tag{14}$$

where Q_25 and Q_75 are the lower and upper quartiles of the dataset, respectively, $output_i$ is the i-th output of the prediction sequence, and up_i and low_i are the upper and lower thresholds of the predicted value, respectively.

4. Experiments

4.1. Dataset Processing and SSA Procedure

In order to accurately predict the health status of a satellite's rotating mechanism, this paper utilize the telemetry data collected from a satellite's 30-day turntable operation as our

dataset. The data used in this study were sourced from the telemetry system of an in-orbit satellite. The data collection frequency varies depending on the operating mode, with a maximum frequency of 1 Hz and a minimum of 1/64 Hz. The dataset includes two primary variables: motor current and turntable temperature. The motor current reflects the current consumed by the motor during different phases of turntable operation, providing critical information about motor load and performance. The turntable temperature data are used to monitor temperature variations, helping to assess the thermal state of the equipment and its operational safety. Furthermore, due to the limited availability of satellite telemetry data, these two types of data are among the few that can directly reflect the state of the rotational mechanisms. The dataset consists of 200,000 records covering a 30-day period, providing a rich foundation for subsequent analysis and modeling. These data volumes are sufficient not only for basic statistical analysis but also for training and validating more complex time-series forecasting models and anomaly detection algorithms. In this study, the dataset will be used for an in-depth analysis to identify potential patterns, trends, and anomalies, thereby providing empirical support for the satellite turntable's health monitoring and fault prediction system. Figure 4 presents the variation curves of the turntable temperature and motor current data over a one-day period.

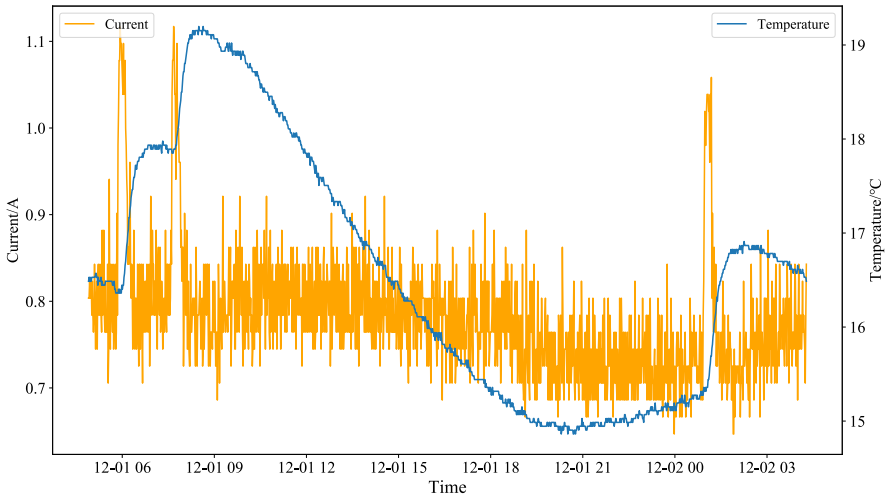


Figure 4. Diagram of motor current and turntable temperature data.

The raw telemetry data from the satellite turntable subsystem are normalized using min-max normalization to map the values into the [0,1] range. This step is crucial for ensuring that the data are on a comparable scale, which aids in improving the performance of the subsequent analysis. The normalization formula is

$$\text{Normalized\_data} = \frac{\text{data} - \min(\text{data})}{\max(\text{data}) - \min(\text{data})} \tag{15}$$

The dataset spans a considerable period and has a sampling frequency of 1/16 Hz, adequately capturing the state changes of the satellite turntable during its operation. After preliminary data analysis, a time series representing the normal operation of the satellite rotation mechanism was selected for modeling. A 7-day period of turntable motor current data and turntable temperature data under normal operating conditions was used to train and generate a health model for the satellite rotation mechanism's temperature. To evaluate the accuracy of predictions and to facilitate comparison, this paper applies three classical evaluation indexes, which are mean absolute error (MAE), mean square error (MSE) and R^2 score defined in Formulas (16)–(18), respectively.

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \bar{y}_i| \tag{16}$$

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \bar{y}_i)^2 \tag{17}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \tag{18}$$

where y_i is the actual value, $\hat{y}_i$ is the predicted value, $\bar{y}$ is the mean of the actual values, and n is the number of data points.

Since the SSA-CNN-Informer model combines the SSA algorithm with an improved informer model, the parameter settings for this model need to be analyzed in two steps. First, the parameters for the SSA algorithm must be configured. As shown in Table 1, to ensure the effectiveness of the SSA algorithm, the length of the time series used for training must be greater than the number of decomposed sequences. For the turntable motor current data and turntable temperature data used in this study, we chose to decompose the data into five sub-sequences. Through preliminary experiments, the window length was fixed at 120, and the reconstruction order was set to 3.

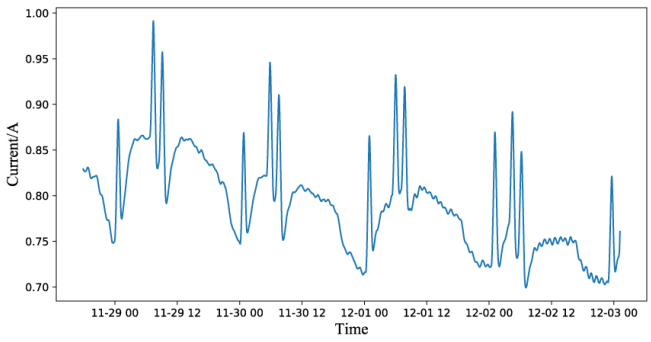
Table 1. SSA parameter setting.

Decomposition and Reconstruction Model	
Decomposition sequence	5
Window length	120
Reconstruction order	3

As can be seen from Figure 5, the main components are determined by calculating the contribution rate of each sub-sequence of motor current data. The main components were determined by calculating the contribution rate of each sub-sequence; the contribution rates from the first sequence to the fifth sequence were 66%, 13.7%, 11.5%, 5.4% and 3.4%, respectively. By observing the waveform of each sub-sequence and using autocorrelation coefficient analysis, it can be determined that first sub-sequence is a trend component, the second sub-sequences and third sub-sequences are oscillation components, and the rest are noise components. Therefore, the reconstruction order is determined to be 3.

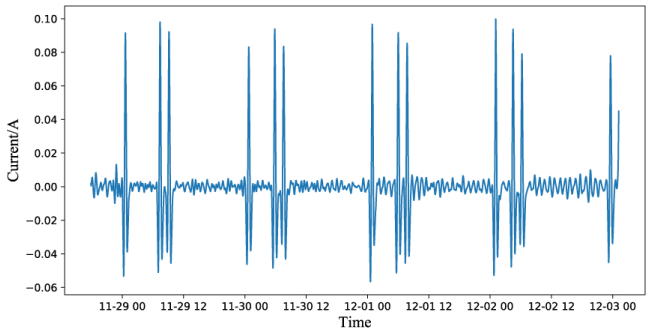
The sequences were reconstructed by regrouping them according to the first three decomposed sub-sequences. The reconstructed sequence was then compared with the original sequence, as shown in Figure 6.

The same method was used to decompose the turntable temperature data, as shown in Figure 7. The contribution rates of each sub-sequence were calculated to identify the primary components; by observing the waveform of each sub-sequence and using autocorrelation coefficient analysis, it was determined that a reconstruction order of 2 was required. The reconstructed sequence was then compared with the original sequence, as illustrated in Figure 8.

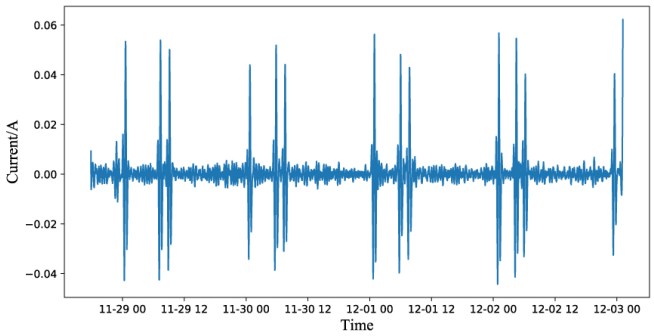


(**a**) Decomposition sequence 1

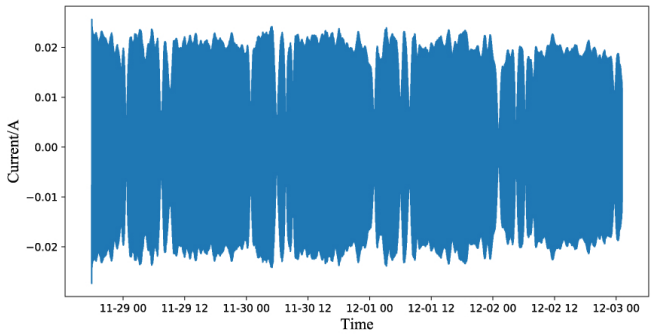
Figure 5. *Cont.*



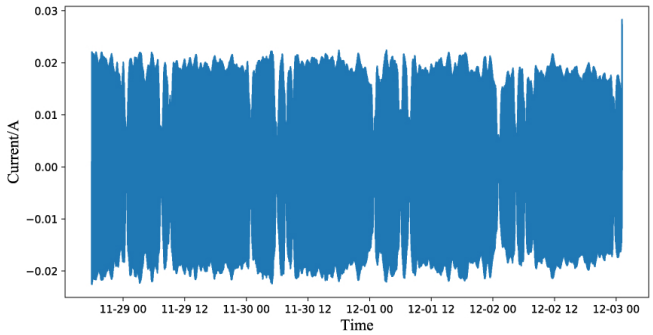
(**b**) Decomposition sequence 2



(**c**) Decomposition sequence 3



(**d**) Decomposition sequence 4



(**e**) Decomposition sequence 5

Figure 5. SSA decomposition sub-sequence of motor current.

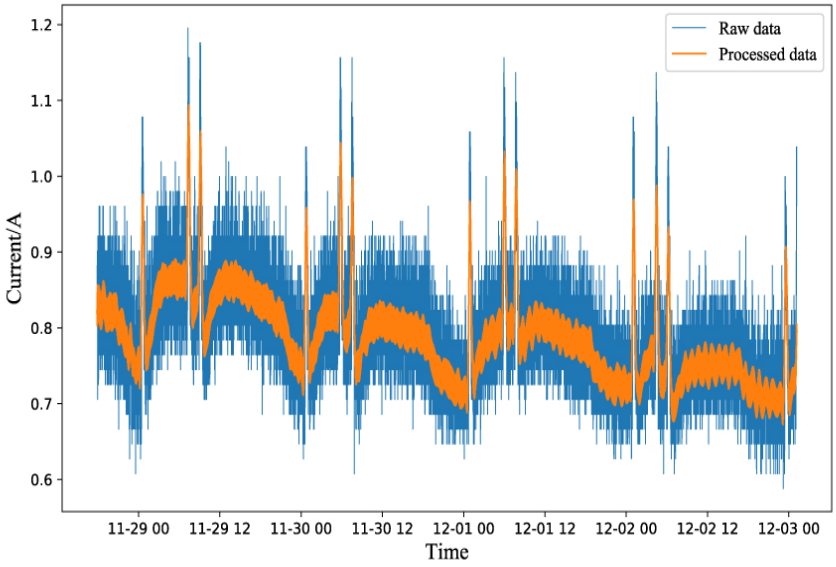
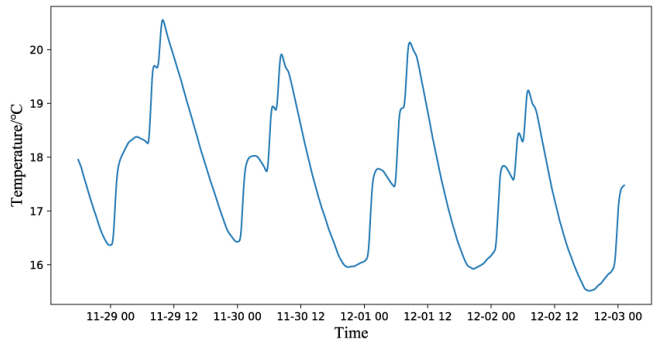
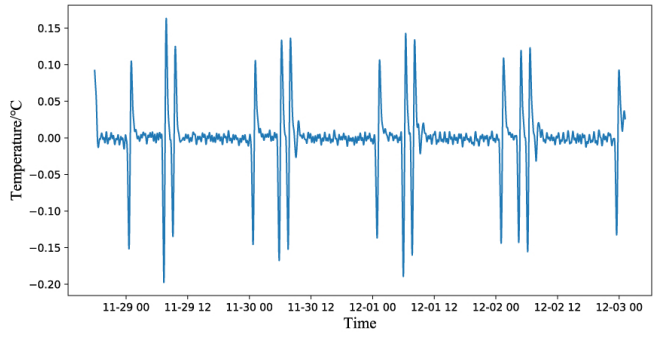


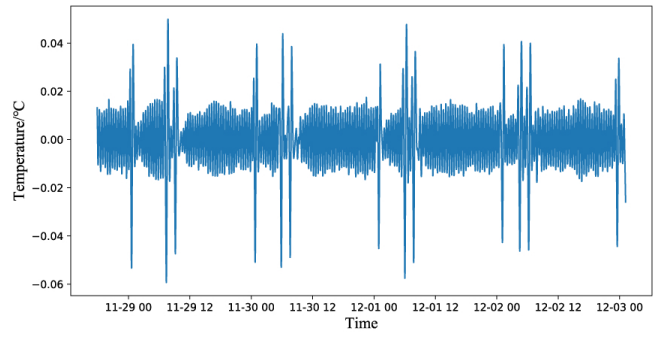
Figure 6. SSA reconstruction sequence of current and raw sequence.



(**a**) Decomposition sequence 1

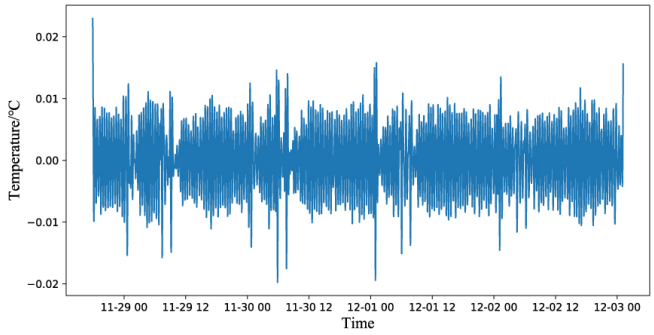


(**b**) Decomposition sequence 2

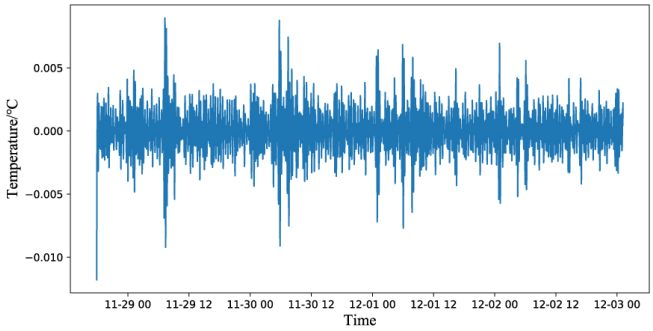


(**c**) Decomposition sequence 3

Figure 7. *Cont*.



(**d**) Decomposition sequence 4



(**e**) Decomposition sequence 5

Figure 7. SSA decomposition sub-sequence of turntable temperature.

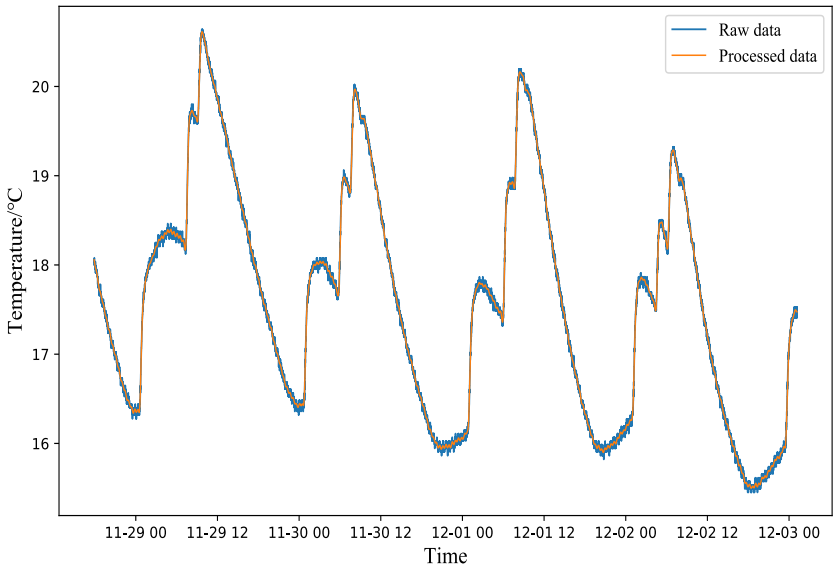


Figure 8. SSA reconstruction sequence of temperature and raw sequence.

4.2. Model Learning Process

The rationale for employing a hybrid model is to leverage the strengths of multiple models, thereby enhancing the robustness of the model against diverse data distributions. This is particularly crucial in low-dimensional feature spaces, where a single model may fail to capture all underlying patterns. The CNN-Informer model obtains correlations within and between data windows through overall training of the window. The prediction length can be set to different values but must be less than the training window length. A single forward prediction is performed through the window. First, by applying the previously described data normalization method, we standardize the newly generated motor current data (denoised using singular spectrum analysis, SSA) along with the turntable temperature data. This ensures that both datasets are distributed within the reasonable range of [0,1]

and converted into data tensors. The next step is to add positional information within the attention window, which captures data correlations in the dataset. Finally, the Adam optimizer is used to optimize the model parameters. The Adam optimizer is a gradient descent algorithm with adaptive learning rates that can handle sparse gradients and mitigate gradient oscillation issues. The model parameter settings are shown in Table 2. As illustrated in Figure 9, with continued training and parameter adjustments, the training error rapidly decreases and stabilizes.

Table 2. The value of model parameter.

Model Parameter Setting	
num of convolution kernel	8
input sequence length of encoder	512
prediction sequence length	120
num of encoder layers	2
num of decoder layers	1
dropout	0.05
optimizer learning rate	0.001
activation function	relu
optimizer	Adam
loss function	MSE
batch size of train input data	22
epoch	80

The model trained with the parameters listed in Table 2 was used to make predictions on the test set, focusing on forecasting turntable temperature data for 600 time steps. The visualized results of the model's predictions are shown in Figure 10. To further validate the performance of the proposed model in this study, we selected several popular models for multi-step time-series prediction in recent years:

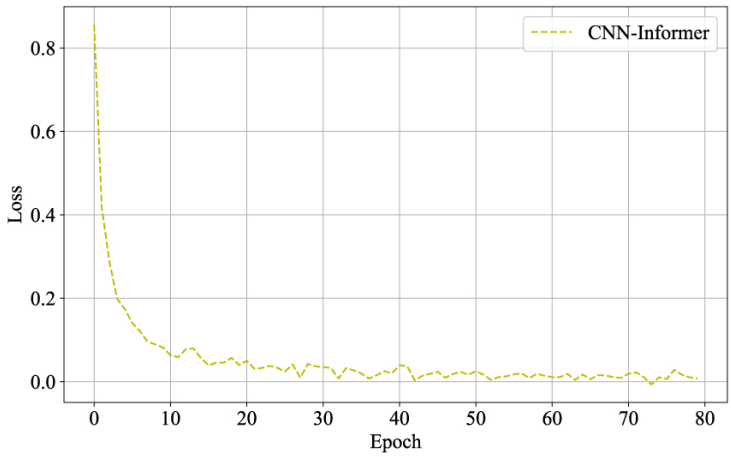


Figure 9. Model loss curve.

LightGBM: Light Gradient Boosting Machine (LightGBM) is an efficient gradient boosting framework, particularly suitable for large-scale data and high-dimensional features. Compared to traditional gradient boosting algorithms, LightGBM demonstrates significant advantages in speed and memory usage, making it highly suitable for tasks such as fault prediction. The complexity of the model can be controlled by adjusting parameters such as tree depth, number of leaves, and learning rate. In comparison to deep learning models, it offers higher training efficiency [31].

LSTM: Long Short-Term Memory is a type of recurrent neural network that can learn long-term dependencies.

CNN-LSTM: The CNN-LSTM model combines Convolutional Neural Networks (CNNs) and Long Short-Term Memory (LSTM) networks, making it suitable for handling sequence data that integrate spatial and temporal information. This model can extract spatial features and temporal dependencies in time-series data, making it applicable to various time-series prediction tasks [32].

DA-RNN: Dual-Stage Attention-Based Recurrent Neural Network (DA-RNN) is a time-series prediction model based on an encoder–decoder structure that utilizes an input attention mechanism to select input features and a temporal attention mechanism to select hidden layer features. By incorporating these dual attention mechanisms, DA-RNN effectively selects feature vectors and establishes long-term historical information dependencies [33].

TCN-SE: TCN-SE is a Time Convolutional Network (TCN) with an added attention mechanism. It utilizes the TCN network for feature extraction and employs the SE attention mechanism to allocate weights between feature vectors. Its inherent dilated convolution mechanism allows it to use long-term historical information for prediction by increasing the receptive field without significantly increasing model depth [34].

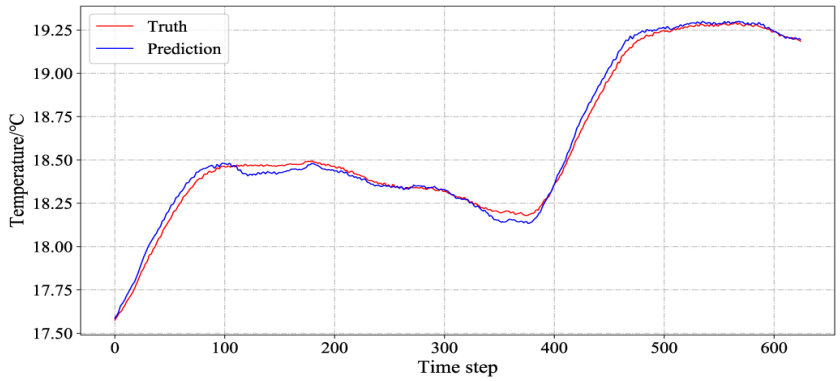
Transformer: The transformer consists of two parts: the encoder and the decoder. The encoder embeds the input sequence, performs feature extraction through multiple layers of self-attention mechanisms, and finally outputs the encoded representation. The decoder generates the output sequence based on the encoded representation using another layer of self-attention and a feedforward neural network [35].

A comparison of the complexities of various models indicates that the computational complexity of CNN-LSTM is primarily constrained by the sequence length and the hidden state dimensions of the LSTM component. This results in a model with moderate complexity, making it suitable for medium-scale data and sequences. DA-RNN, which integrates attention mechanisms with RNNs, significantly increases the number of model parameters. Its computational complexity is relatively high, especially for longer sequences, due to the substantial computational overhead of the attention mechanism. The transformer model's number of parameters depends on the number of layers, attention heads, and hidden dimensions, leading to a large model size. It also exhibits high computational complexity, making it suitable for medium-length sequences and large datasets. Informer has a comparable number of parameters to the standard transformer; however, by utilizing a sparse attention mechanism, it reduces actual computational load and memory consumption, particularly for long sequences and large-scale data. Therefore, overall, the method proposed in this paper has lower complexity compared to other models and can more efficiently handle large-scale data tasks.

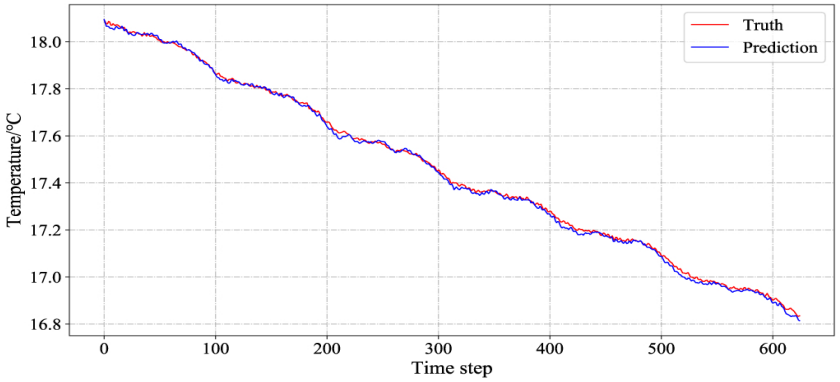
Additionally, this paper compared the prediction results of our proposed method with those of LightGBM, LSTM, CNN-LSTM, DA-RNN, TCN-SE, transformer, and informer. The MSE, MAE and R^2 for the two datasets obtained using different methods are shown in Table 3, with the best results highlighted in bold. The results in the table indicate that our proposed model performs exceptionally well in both evaluation metrics, with accuracy improvements exceeding 10% compared to other algorithms.

Table 3. Performance comparison of different models.

Model	MAE	MSE	R^2
LightGBM	1.225	0.296	0.704
LSTM	0.763	0.187	0.793
CNN-LSTM	0.657	0.176	0.816
DA-RNN	0.781	0.202	0.815
TCN-SE	0.658	0.159	0.834
Transformer	0.486	0.127	0.882
Informer	0.442	0.108	0.898
Proposed model	**0.405**	**0.086**	**0.904**



(**a**) Prediction result 1



(**b**) Prediction result 2

Figure 10. Comparison of predicted data and true data.

4.3. Fault Prediction

After training the satellite rotational mechanism temperature health model using motor current and temperature data from 22 November to 28 November, we used normal telemetry data from the rotational mechanism spanning 72 h, from December 1 to 3 December, as input for the model. The model's output data represent the ideal data under normal operation of the rotational mechanism. Further, using the threshold generation formula described in this paper, we established the upper and lower bounds for the data thresholds as a reference. This study conducted experiments using fault data over an 8 h time span to predict rotational mechanism faults. Figure 11 illustrates the fault detection process determined by the proposed model. The data shown in the figure start at 12:53 on 3 December, with each time step corresponding to 16 s, matching the sampling frequency of our signals. Anomaly detection on subsequent data begins from the 750th time step, which corresponds to 16:13 on 3 December, using the threshold generation method and predictive data approach we proposed. The pink region in the figure represents the normal data range after the 750th time step, while the red curve indicates abnormal temperature data exceeding the threshold. As time progresses, the rotational mechanism temperature data gradually exceeds the threshold range, signaling an impending fault.

Further data collection on the satellite's rotation mechanism confirmed that the turntable indeed experienced a temperature anomaly, consistent with our experimental predictions. Overall, the experimental results demonstrate that the temperature of the satellite's rotation mechanism is not only affected by factors such as orbital position, satellite attitude, and solar irradiation but that it also has a strong correlation with the motor current of the rotation mechanism. The results also show that the proposed method of predicting turntable temperature anomalies using motor current data is highly reliable, providing an effective approach for fault prediction in satellite rotation mechanisms like the turntable.

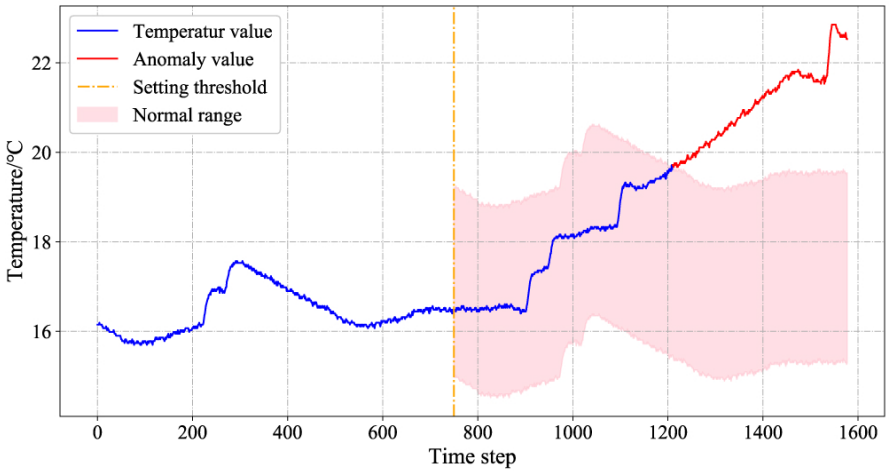


Figure 11. Schematic diagram of turntable temperature threshold generation and fault prediction.

5. Conclusions

Based on SSA and an improved version of the informer, this paper proposes a method for predicting satellite rotating mechanism faults using telemetry data and validates this method through experimental tests. The main contributions of this study are as follows. Firstly, experimental validation was performed using rotational mechanism data from on-orbit satellite telemetry. By studying the working principles and background knowledge of the rotational mechanisms, we propose a method for fault prediction in satellite rotational mechanisms based on the relationship between the turntable motor current and the turntable temperature. Secondly, during data processing, to address the issue of significant background noise interference in telemetry signals, we propose the use of SSA to decompose and reconstruct the telemetry data. Using this method, the original telemetry data were decomposed into five sub-sequences, which were then reconstructed based on their eigenvalue contribution rates. This approach retains the trend and oscillatory components while removing the noise components, thereby enhancing the accuracy of subsequent models in predicting turntable temperature data. Thirdly, for time-series signal prediction, a fault prediction method for satellite turntables based on SSA and an improved informer model was proposed. This method combines the exceptional signal decomposition and reconstruction capabilities of SSA with the outstanding time-series forecasting capabilities of CNN-Informer, thereby further enhancing the practicality and reliability of the prediction results. Lastly, multiple comparative experiments were conducted to validate the proposed method for predicting satellite telemetry data. The experimental results indicate that the proposed method achieved a MAE of 0.405 and a MSE of 0.086, representing at least a 10% improvement over other methods.

In summary, this paper focused on fault prediction for satellite rotational mechanisms. Experimental results on the satellite telemetry dataset demonstrated that the proposed model achieves higher and more stable prediction accuracy compared to the benchmark models. This indicates that the model is effective in performing fault prediction tasks for satellite rotational mechanisms. Our future work will explore the application of this model to other related satellite subsystems and even industrial equipment for fault prediction.

Author Contributions: Conceptualization, Q.L. and Y.Z. (Ye Zhu); methodology, Q.L. and B.L.; software, Y.Z. (Yizheng Zuo); validation, Q.L. and Y.L.; formal analysis, Q.L.; investigation, Y.Z. (Ye Zhu); resources, Y.Z. (Ye Zhu); data curation, Y.L.; writing—original draft preparation, Q.L.; writing—review and editing, Q.L.; visualization, Y.Z. (Yizheng Zuo); project administration, Y.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Excellent Fund of Youth Innovation Promotion Association of Chinese Academy of Sciences, grant number Y2022078.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available on request from the corresponding author. The data are not publicly available due to privacy.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Muthusamy, V.; Kumar, K.D. A novel data-driven method for fault detection and isolation of control moment gyroscopes onboard satellites. *Acta Astronaut.* **2021**, *180*, 604–621. [CrossRef]
2. Chen, Z. Satellite Reaction Wheel Fault Detection Based on Adaptive Threshold Observer. In Proceedings of the 2021 Global Reliability and Prognostics and Health Management (PHM-Nanjing), Nanjing, China, 15–17 October 2021; pp. 1–6.
3. Shi, Z.; Gu, F.; Lennox, B.; Ball, A.D. The development of an adaptive threshold for model-based fault detection of a nonlinear electro-hydraulic system. *Control Eng. Pract.* **2005**, *13*, 1357–1367. [CrossRef]
4. Yin, S.; Gao, H.; Qiu, J.; Kaynak, O. Fault detection for nonlinear process with deterministic disturbances: A just-in-time learning based data driven method. *IEEE Trans. Cybern.* **2016**, *47*, 3649–3657. [CrossRef] [PubMed]
5. Zhang, J.; Chen, H.; Chen, S.; Hong, X. An improved mixture of probabilistic PCA for nonlinear data-driven process monitoring. *IEEE Trans. Cybern.* **2017**, *49*, 198–210. [CrossRef]
6. Maurya, M.R.; Paritosh, P.K.; Rengaswamy, R.; Venkatasubramanian, V. A framework for on-line trend extraction and fault diagnosis. *Eng. Appl. Artif. Intell.* **2010**, *23*, 950–960. [CrossRef]
7. Stief, A.; Ottewill, J.R.; Baranowski, J.; Orkisz, M. A PCA and two-stage Bayesian sensor fusion approach for diagnosing electrical and mechanical faults in induction motors. *IEEE Trans. Ind. Electron.* **2019**, *66*, 9510–9520. [CrossRef]
8. Lee, G.; Han, C.; Yoon, E.S. Multiple-fault diagnosis of the Tennessee Eastman process based on system decomposition and dynamic PLS. *Ind. Eng. Chem. Res.* **2004**, *43*, 8037–8048. [CrossRef]
9. Patel, R.A.; Bhalja, B.R. Condition monitoring and fault diagnosis of induction motor using support vector machine. *Electr. Power Compon. Syst.* **2016**, *44*, 683–692. [CrossRef]
10. Zhou, Z.; Wang, J.; Yang, C.; Wen, C.; Li, Z. Fault detection and isolation of non-Gaussian and nonlinear processes based on statistics pattern analysis and the k-nearest neighbor method. *ACS Omega* **2022**, *7*, 18623–18637. [CrossRef]
11. Wang, Z.; Zarader, J.L.; Argentieri, S. A novel aircraft fault diagnosis and prognosis system based on Gaussian mixture models. In Proceedings of the 2012 12th International Conference on Control Automation Robotics & Vision (ICARCV), Guangzhou, China, 5–7 December 2012; pp. 1794–1799.
12. Wang, N.; Yang, F.; Zhang, R.; Gao, F. Intelligent fault diagnosis for chemical processes using deep learning multimodel fusion. *IEEE Trans. Cybern.* **2020**, *52*, 7121–7135. [CrossRef]
13. Liu, D.; Pang, J.; Song, G.; Xie, W.; Peng, Y.; Peng, X. Fragment anomaly detection with prediction and statistical analysis for satellite telemetry. *IEEE Access* **2017**, *5*, 19269–19281. [CrossRef]
14. Liu, L.; Liu, D.; Zhang, Y.; Peng, Y. Effective sensor selection and data anomaly detection for condition monitoring of aircraft engines. *Sensors* **2016**, *16*, 623. [CrossRef]
15. Ghasemi, S.; Khorasani, K. Fault detection and isolation of the attitude control subsystem of spacecraft formation flying using extended Kalman filters. *Int. J. Control* **2015**, *88*, 2154–2179. [CrossRef]
16. Lu, W.; Ghorbani, A.A. Network anomaly detection based on wavelet analysis. *EURASIP J. Adv. Signal Process.* **2008**, *2009*, 837601. [CrossRef]
17. Chen, H.; Cheng, H.; Jiang, G.; Yoshihira, K. Exploiting local and global invariants for the management of large scale information systems. In Proceedings of the 2008 Eighth IEEE International Conference on Data Mining, Pisa, Italy, 15–19 December 2008; pp. 113–122.
18. Ouadine, A.Y.; Mjahed, M.; Ayad, H.; El Kari, A. UAV quadrotor fault detection and isolation using artificial neural network and Hammerstein-Wiener model. *Stud. Inform. Control* **2020**, *29*, 317–328. [CrossRef]
19. Zhou, C.; Paffenroth, R.C. Anomaly detection with robust deep autoencoders. In Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Halifax, NS, Canada, 13–17 August 2017; pp. 665–674.
20. Yuan, X.; Qiao, Y. Diffusion-ts: Interpretable diffusion for general time series generation. *arXiv* **2024**, arXiv:2403.01742.
21. Meng, H.; Zhang, Y.; Li, Y.; Zhao, H. Spacecraft anomaly detection via transformer reconstruction error. In *Proceedings of the International Conference on Aerospace System Science and Engineering 2019*; Springer: Singapore, 2020; pp. 351–362.
22. Zhang, X.; Rane, K.P.; Kakaravada, I.; Shabaz, M. Research on vibration monitoring and fault diagnosis of rotating machinery based on internet of things technology. *Nonlinear Eng.* **2021**, *10*, 245–254. [CrossRef]
23. Anowar, F.; Sadaoui, S.; Selim, B. Conceptual and empirical comparison of dimensionality reduction algorithms (pca, kpca, lda, mds, svd, lle, isomap, le, ica, t-sne). *Comput. Sci. Rev.* **2021**, *40*, 100378. [CrossRef]
24. Li, H.; Cui, L.; Guo, S. A Hybrid Short-Term Power Load Forecasting Model Based on the Singular Spectrum Analysis and Autoregressive Model. *Adv. Electr. Eng.* **2014**, *2014*, 424781. [CrossRef]
25. Vahabie, A.H.; Yousefi, M.M.R.; Araabi, B.N.; Lucas, C.; Barghinia, S. Combination of singular spectrum analysis and autoregressive model for short term load forecasting. In Proceedings of the 2007 IEEE Lausanne Power Tech, Lausanne, Switzerland, 1–5 July 2007; pp. 1090–1093.

26. Xu, K.; Liang, X.; Tian, X. Water Level Prediction Based on SSA-LSTM Model. In Proceedings of the 2022 7th International Conference on Computational Intelligence and Applications (ICCIA), Nanjing, China, 24–26 June 2022; pp. 108–112.
27. Zhou, H.; Zhang, S.; Peng, J.; Zhang, S.; Li, J.; Xiong, H.; Zhang, W. Informer: Beyond efficient transformer for long sequence time-series forecasting. In Proceedings of the AAAI Conference on Artificial Intelligence, Virtual Event, 2–9 February 2021; Volume 35, pp. 11106–11115.
28. Chen, S.; Lu, N. CNN-Informer-Based Remaining Useful Life Prediction for Electrical Devices. In Proceedings of the 2022 4th International Conference on Industrial Artificial Intelligence (IAI), Shenyang, China, 24–27 August 2022; pp. 1–6.
29. Ma, J.; Dan, J. Long-Term Structural State Trend Forecasting Based on an FFT–Informer Model. *Appl. Sci.* **2023**, *13*, 2553. [CrossRef]
30. Withers, C.S.; Nadarajah, S. The distribution and quantiles of the range of a Wiener process. *Appl. Math. Comput.* **2014**, *232*, 766–770. [CrossRef]
31. Guo, J.; Yun, S.; Meng, Y.; He, N.; Ye, D.; Zhao, Z.; Jia, L.; Yang, L. Prediction of heating and cooling loads based on light gradient boosting machine algorithms. *Build. Environ.* **2023**, *236*, 110252. [CrossRef]
32. Livieris, I.E.; Pintelas, E.; Pintelas, P. A CNN–LSTM model for gold price time-series forecasting. *Neural Comput. Appl.* **2020**, *32*, 17351–17360. [CrossRef]
33. Xu, L.; Zhang, L.; Sun, R.; Zhang, N.; Liu, P.; Guan, P. Multi-step Probabilistic Load Forecasting for University Buildings Based on DA-RNN-MDN. In *Proceedings of the International Conference on Intelligent Computing*; Springer: Singapore, 2023; pp. 662–673.
34. Chen, W.; Zhang, R.; Shi, C.; Zhu, Y.; Lin, X. A complement method for magnetic data based on TCN-SE model. *Sensors* **2022**, *22*, 8277. [CrossRef]
35. Han, K.; Xiao, A.; Wu, E.; Guo, J.; Xu, C.; Wang, Y. Transformer in transformer. *Adv. Neural Inf. Process. Syst.* **2021**, *34*, 15908–15919.