Article

Unsupervised Noise Reductions for Gravitational Reference Sensors or Accelerometers Based on the Noise2Noise Method

Zhilan Yang [1,2,3], Haoyue Zhang [4], Peng Xu [3,4,5,*] and Ziren Luo [3,5]

1   National Space Science Center, Chinese Academy of Sciences, Beijing 100094, China;
    yangzhilan20@mails.ucas.ac.cn
2   University of Chinese Academy of Sciences, Beijing 100094, China
3   Hangzhou Institute for Advanced Study (UCAS), Hangzhou 310000, China
4   Lanzhou Center of Theoretical Physics, Lanzhou University, Lanzhou 730000, China
5   Institute of Mechanics, Chinese Academy of Sciences, Beijing 100094, China
*   Correspondence: xupeng@imech.ac.cn

Abstract: Onboard electrostatic suspension inertial sensors are important applications for gravity satellites and space gravitational-wave detection missions, and it is important to suppress noise in the measurement signal. Due to the complex coupling between the working space environment and the satellite platform, the process of noise generation is extremely complex, and traditional noise modeling and subtraction methods have certain limitations. With the development of deep learning, applying it to high-precision inertial sensors to improve the signal-to-noise ratio is a practically meaningful task. Since there is a single noise sample and unknown true value in the measured data in orbit, odd–even sub-samplers and periodic sub-samplers are designed to process general signals and periodic signals, and adds reconstruction layers consisting of fully connected layers to the model. Experimental analysis and comparison are conducted based on simulation data, GRACE-FO acceleration data, and Taiji-1 acceleration data. The results show that the deep learning method is superior to traditional data smoothing processing solutions.

Keywords: Noise2Noise; deep learning; denoising; accelerometer; inertial sensor

1. Introduction

Inertial sensors are able to detect acceleration, angular velocity, gravity vectors, and other motion factors. This paper focuses on the application of inertial sensors in the fields of aviation and gravitational wave detection. In space, star trackers can serve as the inertial reference for spacecraft, and are employed in space satellite missions related to precise space measurements, high-precision navigation, and mapping. High-precision inertial sensors play a crucial role as key scientific payloads in space missions, such as global gravity field inversion, space gravitational wave detection, and gravitational field measurements. Among modern advanced technologies, including superconducting magnetic levitation and SQUID technology, and cold atom interferometry technology, high-precision electrostatic suspension inertial sensors based on electrostatic servo-control technology are still the most advanced and reliable onboard inertial reference technology for in-orbit operation. Accomplishing a series of gravity satellite missions, including GRACE [1], GRACE-FO [2], and GOCE [3], high-precision inertial sensors that use electrostatic levitation have proven to be successful. The ultra-high-precision electrostatic suspension inertial sensor, verified by the LISA PathFinder [4] satellite, is expected to become a crucial payload for upcoming space gravitational wave detection missions, such as the LISA [4], Taiji [5], and Tianqin plans [6] In the Taiji mission, inertial sensors can not only track the motion of the test mass by measuring the capacitance change caused by its displacement but also measure the acceleration signal of the test mass with high accuracy. The accuracy of the inertial sensor has an important impact on the sensitivity of space gravitational wave detection

missions. The electrostatic suspension inertial sensor system mainly consists of the test mass as the inertial reference, the electrode housing with sensing and driving electrodes, the servo control electronic system, the vacuum system, and other charge management subsystems and isolation subsystems that depend on the specific space environment and satellite platform characteristics for a given mission. From a hardware perspective, it is challenging to minimize the effects of environmental interference and noise on the measurement system and achieve higher detection sensitivity through hardware improvements once a certain level is reached. Due to its extremely high sensitivity, the on-board electrostatically suspended inertial sensor is subject to complex physical environment coupling interference, including temperature gradient fluctuations, magnetic fields, electric field fluctuations, microvibrations, high-energy particle beam charging, and other environmental factors related to the specific space environment and satellite platform characteristics. Moreover, the measurement data of the sensor are extremely complex with interference and noise components, including readout noise and control noise. Therefore, traditional processing methods, such as noise modeling, subtraction, data smoothing, and trend fitting are used to suppress noise and improve signal-to-noise ratio. However, the noise signal often exhibits unfavorable fluctuations in frequency similar to potential scientific signals, making it difficult to achieve the best noise filter [7]. Intelligent computing methods are considered to assist in completing this task. This is a critical technology for further improving the detection sensitivity of inertial sensors and achieving accurate and efficient scientific applications of measurement data.

In recent years, deep learning has developed rapidly and made progress in many fields. For example, in the fields of Computer Vision, Natural Language Processing and Time-series, deep learning methods continue to evolve and reach advanced levels of noise reduction compared with traditional processing methods. Generally speaking, deep learning techniques can be divided into two categories, supervised and unsupervised learning. With supervised methods, we need true data to train the neural network model. Changhui Jiang et al. [8], based on supervised learning, used a combination of recursive neural network (RNN) and long and short time memory network (LSTM) to process the output of inertial sensor as a time series signal, finally improving the accuracy of the inertial sensor. Despite the wide application of supervised learning, its limitations are evident, as it is difficult to acquire data that are close to true values and the accuracy of the true value samples severely affects the effectiveness of denoising algorithms. In the case of signal measurements, such as signals from satellites like Taiji-1 and GRACE-FO, only noisy signals but no true signals are obtained. Although it is possible to obtain data from simulating dynamic equations constructed by sensors and observed objects, or from ground tests simulating orbit environmental conditions, the data obtained by these methods are often not accurate enough, thus affecting the final denoising effects. Therefore, considering unsupervised learning methods for signal denoising is recommended. For unsupervised or self-supervised learning denoising, Jaakko Lehtinen et al. [9] proposed the Noise2Noise (N2N) denoising framework, which can train a denoising network model without using clean images as training samples. The authors experimented with simple noise distributions (Gaussian, Poisson, Bernoulli) and complex, intractable synthetic noise from Monte Carlo images. The final results show that deep learning neural networks are able to denoise signals without the need for clean target data, and the final performance level is equal to or close to that of using clean target data. In order to solve the limitation of requiring multiple noise samples in the N2N method, Tao Huang et al. [10] proposed the Neighbor2Neighbor method based on the N2N method, and proposed a random neighbor subsampler to generate training image pairs based on a single noise image sample. The final experimental results show that the deep learning method can deal with the problem of only single noisy signal denoising, and has achieved higher effectiveness and superiority than the existing methods. In addition to the field of image, many people have applied the N2N method to the signal processing of time series. Qingchun Li et al. [11] proposed a single noise audio denoising framework (SNA-DF) based on N2N for processing single noise audio denoising,

and used the deep and complex U-net model to realize the denoising processing. Shirong Koh et al. [12] proposed the WaveN2N model to deal with the noise of acoustic signals in underwater areas without prior knowledge and clean signals from real data; Takayuki Takaai et al. [13] applied the Noise2Noise method to current waveform signals obtained from multi-stage narrow nanochannels, which are characterized by high noise and complex measurement principles. The CAE model and U-net model are used, respectively, and the final noise reduction effect is better than the traditional signal processing methods, such as frequency filter, wavelet transform, and Kalman filter, which can retain the signal details more accurately. Noise reduction framework is not only limited to Noise2Noise. Mario Morvan et al. [7]'s Transformer model based on the Noise2Self framework combined with time series realizes light curve signal noise reduction for damaged TESS models. This method has flexibility and better performance when dealing with large datasets. The main contributions of this work are as follows:

1. We applied the N2N method for the first time to suppress noise in inertial sensor data.
2. The signal was divided into periodic and general components, and we proposed the use of a periodic sub-sampler and an odd–even sub-sampler. For the periodic component, we proposed the addition of a reconstruction layer to the model.
3. We applied the N2N method to the Taiji-1 calibration task and GRACE-FO Level-1A data, effectively suppressing noise.

The structure of this paper is as follows. Section 2 introduces the working principle and noise analysis of electrostatic suspension inertial sensor. Section 3 introduces an overall denoising framework, the theoretical basis of the N2N method, and the custom neural network architecture. Section 4 introduces the simulation experiment and real experiment about the Taiji-1 satellite and GRACE-FO satellite, and also carries out the comparative research results and discussion with a variety of filters. Finally, Section 5 summarizes the results of the work as well as the potential range of applications.

2. Electrostatic Levitation Inertial Sensors

2.1. Overview

The electrostatic levitation inertial sensor system comprises several primary subsystems, including the test mass, electrode housing that surrounds the test mass, electronic measurement and control unit, and vacuum chamber, with the test mass serving as the fundamental inertial reference component. Depending on the mission environment and satellite platform characteristics, supplementary subsystems such as charge management and system isolation may also be incorporated. The data experiments conducted in this paper are based on available GRACE-FO inertial sensor (or accelerometer) Level-1A data and measured data from the Taiji-1 satellite inertial sensor. The basic principles of the inertial sensors used in both missions are similar. The distribution of the test mass and its surrounding measurement and control electrodes in the inertial sensor are illustrated in Figure 1.

The test mass is located at the nominal geometric center position inside the electrode housing. The sensitive structure consisting of the test mass and the electrode housing is situated in a vacuum environment with stable magnetic field fluctuations and temperature gradients. During scientific operations of the spacecraft, deviations from the inertial trajectory of free fall in the gravitational field occur due to non-gravitational disturbances from the space environment, resulting in relative motion between the satellite platform and the test mass. The resulting relative displacement is projected onto a measurement axis, causing changes in the distance between the test mass surface and the corresponding electrodes on either side, which is reflected in changes in capacitance. The differential capacitance detection circuit precisely measures the displacement of the test mass relative to its nominal center position. The control unit receives these data and adjusts the electrostatic force by changing the voltage, in order to real-time control the relative motion between the test mass and the electrode housing, and maintain the test mass in the vicinity of the nominal equilibrium position.
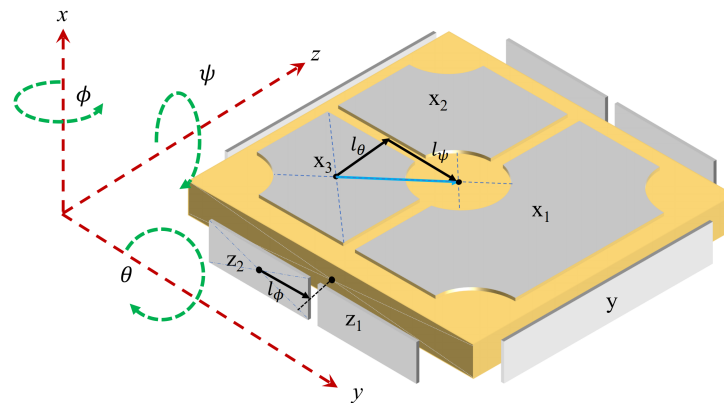
Figure 1. Arrangement of the primary mechanical components of the inertial sensor.

2.2. The Working Principle of Inertial Sensors

In actual operation, electrostatically suspended inertial sensors on-board spacecraft have two operating modes, namely, the accelerometer mode and the drag-free control mode. In missions such as LISA PathFinder and Taiji-1, the inertial sensors on the space gravitational wave detection technology experimental satellite can switch between these two modes depending on the mission requirements. For gravity satellite missions, such as GRACE and GRACE-FO, the inertial sensors operate in the accelerometer mode. In the accelerometer mode, the electrostatic force is applied to the test mass by changing the voltage on the electrodes surrounding the test mass to keep it in its nominal equilibrium position. In the drag-free mode, the object of control is changed, and the satellite is controlled by feedback control applied to the thrusters, so that the satellite follows the motion of the test mass, which is suspended near its equilibrium position. Taking the accelerometer measurement mode as an example, the residual relative acceleration is negligible. The resulting non-gravitational perturbations to the spacecraft (atmospheric drag, solar pressure, Earth reflection, etc.) are as follows

$$a^i_{para,SC}(t) = -p_1^{i\alpha}V_{\alpha}(t) - G^{ij}(t)d_j - a^i_{para,TM}(t) - k^i \tag{1}$$

where $a^i_{TM}(t)$ and $a^i_{SC}(t)$, respectively, denote the accelerations of the test mass and spacecraft relative to the inertial reference frame, k^i represents the acceleration bias and $p^{i\alpha}$ represents the linear acceleration scale factor. The subscript "para" signifies the parasitic disturbance acceleration. d is the deviation between the test mass centroid and the spacecraft centroid [14]. $V_{\alpha}(t)$ represents the control voltage at each electrode measured and read out, and $G^{ij}(t)$ represents the comprehensive terms such as gravitational gradient matrix and spacecraft angular momentum matrix.

After the center of mass deviation has been corrected and working parameters have been accurately calibrated, control voltage or control acceleration can be used to provide precise measurements of non-gravitational disturbances acting on the spacecraft. The level of parasitic acceleration disturbance to the test quality in the equation is a key factor that limits measurement accuracy.

2.3. Noise Analysis of Accelerometers

During the actual operation of the satellite-borne electrostatic inertial sensors, due to their extremely high sensitivity and the complex coupling relationship between the test mass and the surrounding physical fields of multiple physical fields, the components of the parasitic disturbance acceleration $a^i_{para,TM}(t)$ that the test mass in Equation (1) experiences are extremely complex. These mainly include acceleration disturbances caused by the measurement and control of the test mass, including displacement detection noise and control errors caused by it, and acceleration noise caused by unstable control voltage; accelera-

tion noise caused by the coupling of the relative displacement jitter and parasitic stiffness (mainly electrostatic stiffness and self-gravity stiffness) of the test mass with respect to the equilibrium position; acceleration noise caused by the coupling of the residual magnetic moment and induced magnetic moment of the test mass with magnetic field fluctuations; noise generated by residual gas molecule random collisions; residual gas molecule coupling temperature gradient fluctuations of thermal radiometer effect in vacuum environment, as well as asymmetric gas outflow effect; acceleration noise generated by the coupling of test mass charge fluctuations with electric field environment; and acceleration noise generated by dissipative effects, including dielectric loss.

The majority of the aforementioned noises cannot be precisely eliminated by modeling, combining with other cross-checking or platform environment data. In fact, most of the noises exhibit broad-spectrum colored noises in the amplitude spectral density of the measurement frequency band. Therefore, from the perspective of data analysis, effective suppression of unmodelable broad-spectrum random noises through learning from the noise and further improving the detection signal-to-noise ratio will have significant practical implications for the application of high-precision spaceborne electrostatically suspended inertial sensors.

3. Methodology

This section covers the following topics. Firstly, we introduce an overall noise reduction framework, including the theoretical foundation of the Noise2Noise algorithm and the sub-sampler. Then we introduce the U-net model and the CAE model with added reconstruction layers separately.

The overall framework of the algorithm is shown in Figures 2 and 3. The accelerometer measurement signals are divided into periodic signals and general signals, and the training set is constructed using a periodic sub-sampler and an odd–even sub-sampler, respectively. An appropriate network model is trained through the training set (adding reconstruction layers to the network model for periodic signals, and not adding them for general signals). The accelerometer measurement signals are then inputted into the trained network model to obtain the noise-suppressed signals.
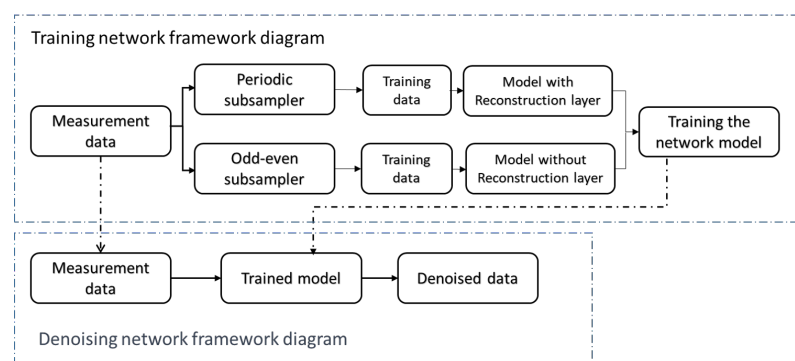

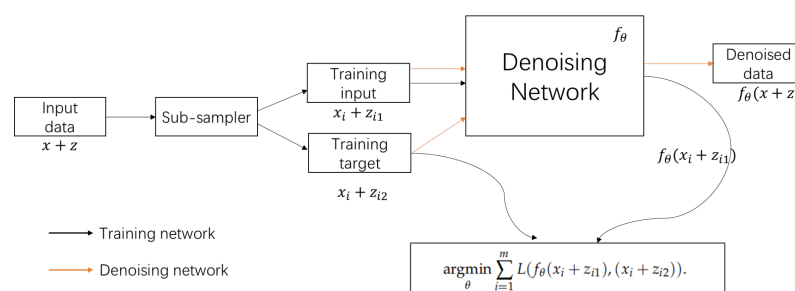**Figure 2.** Overall framework of noise reduction algorithm.

Figure 3. The overview of N2N framework.

3.1. Noise2Noise Revisit

Based on the content of Section 2, the working mode of the inertial sensor in accelerometer mode for on-orbit measurement data can be expressed as follows

$$y_i = x_i + z_i, \tag{2}$$

where x_i represents the actual non-gravitational disturbance acceleration signal $a_{para,SC}$ experienced by the spacecraft, z_i represents the accelerometer noise acceleration signal $a_{para,TM}$ due to test mass, and y_i represents the measured accelerometer signal a_c. The subscript i denotes the data label ($i = 1, 2, 3 \ldots k$). The basic idea of deep learning method to deal with noise is to establish the relationship between y and x, and finally the denoised accelerometer signal can be obtained through this relationship.

The supervised learning method uses clean signal and noisy signal pairs $(x_i, x_i + z_i)$ to train the parameters of the network model, also known as Noise2Clean(N2C) method, and its specific expression is as follows

$$\underset{\theta}{\mathrm{argmin}} \sum_{i=1}^{m} L(f_\theta(x_i + z_i), x_i), \tag{3}$$

where L represents a measure of the difference between the output of the network model and the true signal, m represents the number of training samples, f_θ represents the signal denoising network parameterized by θ. We can choose several p-norms as the loss function L. The N2N model is an unsupervised deep learning method, which does not require clean signals, requires only the noisy measurements $x_i + z_{i1}$ and $x_i + z_{i2}$ based on the true value x^i to train the network model. Its specific expression is as follows:

$$\underset{\theta}{\mathrm{argmin}} \sum_{i=1}^{m} L(f_\theta(x_i + z_{i1}), (x_i + z_{i2})). \tag{4}$$

Equation (3) is equivalent to Equation (4) when the two following conditions are satisfied:

- **Condition 1**: The noise of measurement in the input is independent from the noise in the target which is used to train the network;
- **Condition 2**: The expectation of noise added to the signal is zero.

The N2N model can learn to remove noise from a signal rather than learn from noise $x + z_1$ to noise $x + z_2$ mapping. For one pair of noise samples (y_1, y_2), in order to avoid the problem of insufficient samples, we can generate two pairs of noise samples (y_1, y_2) and (y_2, y_1) by using noise y_1 and noise y_2 as input and target, respectively [15]:

$$\underset{\theta}{\mathrm{argmin}} \sum_{i=1}^{m} \left(\frac{1}{2} L(f_\theta(x_i + z_{i1}), (x_i + z_{i2})) + \frac{1}{2} L(f_\theta(x_i + z_{i2}), (x_i + z_{i1})) \right). \tag{5}$$

The N2N method requires multiple noise samples, and we use a periodic signal sub-sampler to achieve the generation of noise sample pairs from a single noise sample with periodicity. We define a periodic noise sequence as y and its period is T, and use a hyper-parameter k to control the number of periods of the interval, $k \geq 2$. The input and output of the network model is a d dimensional time series vector. For example, if we choose i-th to $(i + d - 1)$-th as sample $s_1(y)$, correspondingly, if we choose $(i + kT)$-th to $(i + kT + d - 1)$-th as sample $s_2(y)$, we obtain two sample pairs $(s_1(y), s_2(y))$ and $(s_2(y), s_1(y))$. The periodic sub-sampler is shown in Figure 4. Based on the sample pairs, we obtain the following equations:

$$f_\theta(x_t + z_t) \rightarrow x_{t+kT} + z_{t+kT}, \tag{6}$$

$$f_\theta(x_t + z_t) \to x_t + z_{t+kT} + (x_{t+kT} - x_t). \tag{7}$$
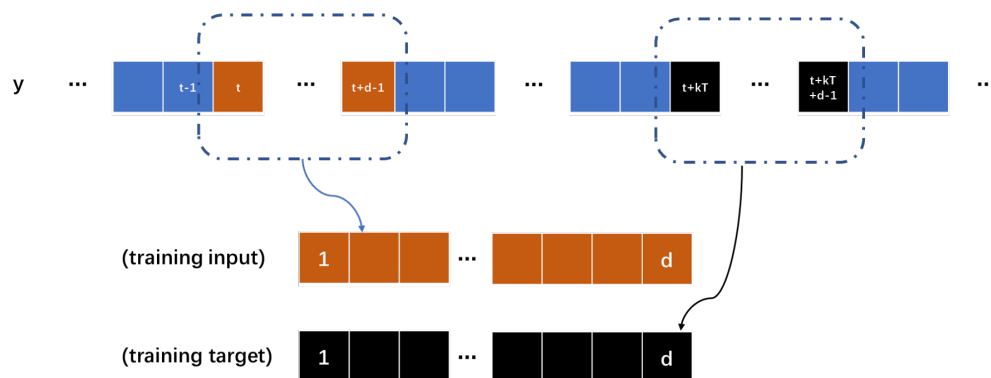
Figure 4. The period sub-sampler module in framework.

We can ignore the difference between the truth values of the signal between intervals n periods, and the map learned by the trained neural network model is $f_\theta(x_t + z_t) \to x_t + z_{t+kT}$, which satisfies the conditions of the N2N method. *nums* is the total number of cycles. The hyper-parameter $k \in [2, nums - 1]$ can be selected to obtain $(2nums - 4)$ samples, avoiding the problem of under-fitting of the model and poor noise reduction effect caused by insufficient samples. For a general signal without periodicity, we use odd–even samplers to generate sub-samples, dividing the noise signal y into y_{odd} and y_{even}. The odd–even sub-sampler is shown in Figure 5. We can obtain the following equation:

$$f_\theta(x_{odd} + z_{odd}) \to x_{even} + z_{even}, \tag{8}$$

$$f_\theta(x_{odd} + z_{odd}) \to x_{odd} + z_{even} + (x_{even} - x_{odd}). \tag{9}$$
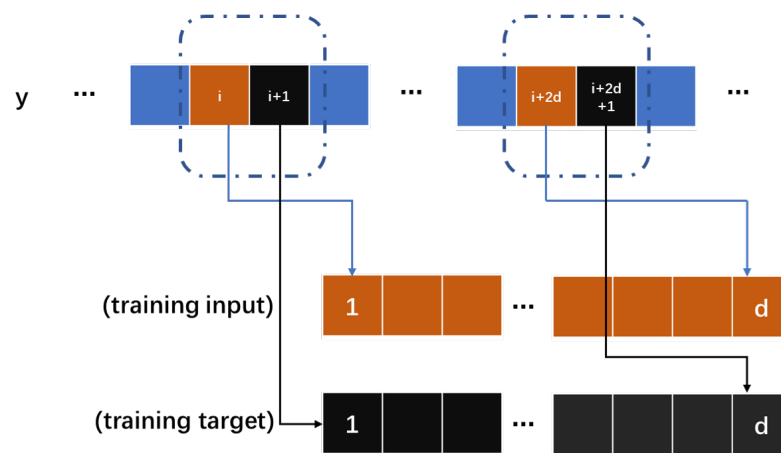
Figure 5. The odd-even sub-sampler module in framework.

We can assume that $(x_{even} - x_{odd})$ is approximately equal to zero, and the mapping is $f_\theta(x_{odd} + z_{odd}) \to x_{odd} + z_{even}$, which satisfies the condition of the N2N method.

3.2. Network Model Architecture

In the selection of deep learning models, we try the improved CAE model and U-net model used by Lehtinen in Noise2Noise, which both use one-dimensional convolution to adapt to the input and output of one-dimensional time series. In the model, we try to use the dropout layer which randomly sets the weight of some neurons to 0 with probability

p during the training process to prevent the neural network from overfitting [16]. The dimension reduction in the model uses a convolution layer with a step size of 2 instead of a pooling layer, which can avoid the loss of part of the signal information. Although using convolutional layers increases the computational complexity, it can achieve better results, and we do not have to process the data online, so it is not sensitive to the time it takes to train the network.

The difference between the two models mainly lies in the network model architecture, where the overall architecture of U-net is shown in Figure 6, which can be divided into three parts, encoder, decoder, and skip connection. The input $x1$ of U-net is 1500 dimensional, representing time series waveform, and goes into the encoder. Each layer of the encoder consists of a one-dimensional convolution layer, dropout layer, and pooling layer (or a convolution layer and the stride width be 2–3) to generate a feature layer, and then goes into the decoder, which consists of deconvolution and up-sampling layer. Skip connections concatenate the decoder and encoder parts of the same dimension. These connection channels enable the network to learn deep features and shallow features of the training data. Finally, the m-dimensional output $y1$ is obtained to achieve signal noise reduction.

(**a**) U-net Architecture   (**b**) CAE Architecture
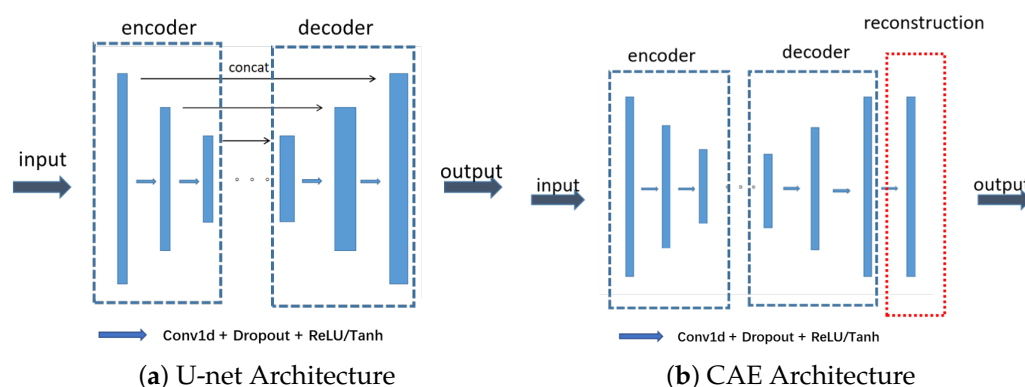
Figure 6. Network model under N2N framework for signal denosing.

The overall architecture of CAE is shown in Figure 6, which also has an encoder and decoder. The encoder compresses the m-dimensional input into a feature layer through a 1-D convolutional layer with a step size of 1 or 2, and then enters a decoder consisting of deconvolution proportional to the encoder. The output of the decoder additionally goes into a reconstructor composed of fully connected layers, which enables the model to converge faster under the influence of low-frequency noise.

4. Experiments and Results

In the following sections, we show the effect of the N2N method for synthetic periodic noise signals and the specific noise reduction effect of real accelerometer sensor signals from Taiji-1 and GRACE-FO.

4.1. Simulation Data Experiments

In the calibration scheme of Taiji-1, we need to perform a specific maneuver scheme for the satellite. Considering the waveform of the accelerometer signal during the satellite maneuver experiment and adding more details in the period of the signal, We add the sinusoidal components y_1, y_2, and y_3 to form a square wave signal y. The synthetic noise consists of the following two parts, (1) Gaussian noise with a fixed level $\sigma = 0.2$; (2) colored noise generated by Gaussian noise ($\sigma = 0.1$) through filtering. We add synthetic noise to the square wave signal and the waveform diagram is shown in Figure 7:

$$y_1 = 4\pi \sin(x), \tag{10}$$

$$y_2 = \frac{4}{3}\pi\sin(3x), \tag{11}$$

$$y_3 = \frac{4}{5}\pi\sin(5x), \tag{12}$$

$$y = \frac{y_1 + y_2 + y_3}{25}. \tag{13}$$

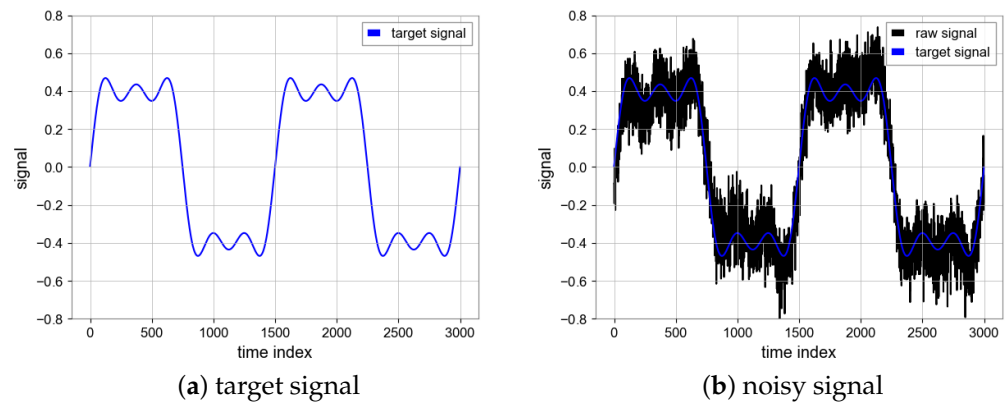**(a)** target signal   **(b)** noisy signal

Figure 7. Ground truth signal and mixed noise signal.

The total length of the simulation signal sequence is 150,000 s, the sampling frequency is 1Hz, and the signal period is 1500 s. The dataset is divided into a training set and a test set, which are 80% and 20% of the total sequence length, respectively. In total, 80% of the data are periodically sampled to generate a paired training set, and are used to train the neural network model, while 20% of the data are used to test the model. We choose mean square error (MSE) and signal noise ratio (SNR) to evaluate the noise reduction result of the N2N algorithm.

$$MSE = \frac{1}{m}\sum_{i=1}^{m}(x_i - f(y_i))^2, \tag{14}$$

$$SNR = 10\log(\frac{P_s}{P_n}), \tag{15}$$

where P_s represents the power of the signal and P_n represents the power of the noise. We also try to denoise the data by low-pass filter, wavelet decomposition denoising and Kalman filter to compare the advantages and disadvantages of N2N method and discuss whether it can be combined with N2N.

4.1.1. Wavelet Denoising Filter

In the process of wavelet denoising, the noise signal is decomposed by wavelet transform by selecting the appropriate wavelet, and the decomposed signal is divided into the high-frequency part and low-frequency approximate part. Low frequency approximate part can be further decomposed, and the threshold denoising method is used to deal with the noise in high frequency part [17]. The method of grid search is used to obtain the optimal parameters, and the specific results are shown in the Table 1. In this study we tried different wavelet types and thresholds and finally chose the "db36" wave with a threshold of 0.3. The SNR is 12.64 and MSE is 0.0078 which are achieved by the wavelet filter and the results can be seen in Figure 8.

Table 1. Noise reduction results of the wavelet filter.

Name/Threshold	0.1	0.3	0.5	0.7	0.9	1.0	1.5
Sym8	0.013	0.034	0.071	0.106	0.133	0.139	0.140
Coif4	0.013	0.034	0.071	0.107	0.134	0.140	0.140
Db3	0.013	0.034	0.067	0.103	0.131	0.144	0.144
Db36	0.010	0.009	0.013	0.016	0.018	0.018	0.018

(**a**) Denoised signal
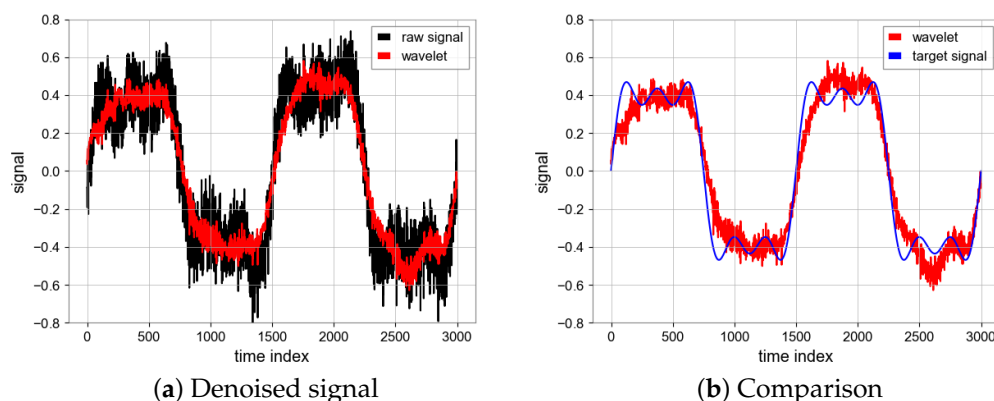
(**b**) Comparison

Figure 8. Denoised signal using wavelet transform and comparison between denoised signal (Red) and target signal (Blue).

4.1.2. Kalman Filter

The Kalman filter is divided into the prediction process and measurement process, which iterates continuously to obtain more accurate state estimation, and, finally, can effectively deal with noise. We tried different process variance matrices (Q) and measurement variance matrices (R), as shown in Table 2. Finally, we chose the values $Q = 10$ and $R = 1000$ [18]. The SNR is 16.54 and MSE is 0.0033 which are achieved by the Kalman filter and the results can be seen in Figure 9.

Table 2. Noise reduction results of the Kalman filter.

(Q, R)	(10, 1000)	(1, 100)	(1, 10)	(1, 1)	(10, 1)
MSE	0.0052	0.0058	0.0077	0.0123	0.0123

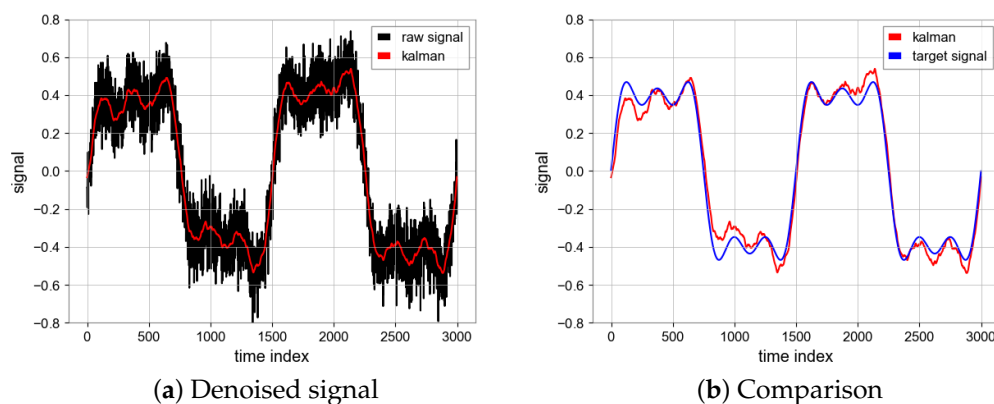(**a**) Denoised signal

(**b**) Comparison

Figure 9. Denoised signal using Kalman filter and comparison between denoised signal (Red) and target signal (Blue).

4.1.3. Butterworth Filter

The Butterworth filter makes the signal frequency flat after the passband, while the signal between the cut-off frequencies is rejected. For the signal of the experiment. Based

on the row search, we select a low-pass filter with a cut-off frequency of 0.005 Hz, as shown in Table 3.

The SNR is 16.58 and MSE is 0.0033 which are achieved by the Butterworth Filter and the results can be seen in Figure 10.

Table 3. Noise reduction results of Low-pass Filter.

Cut-Off Frequency	0.0005	0.001	0.0025	0.005	0.01	0.015	0.02	0.025	0.05	0.1
MSE	0.1336	0.0223	0.0097	0.0052	0.0052	0.0053	0.0054	0.0055	0.006	0.0069

(**a**) Denoised signal
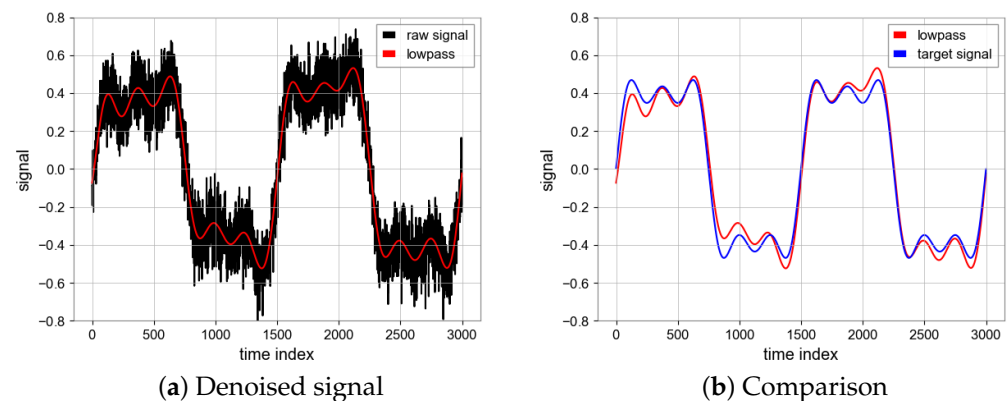
(**b**) Comparison

Figure 10. Denoised signal using butterworth filter and comparison between denoised signal (Red) and target signal (Blue).

4.1.4. N2N Algorithm

All of the above experiments are run on NVDIA RTX2060 GPU, the running environment of the program is Python3 under a Window system, and libraries, such as TensorFlow and Numpy, are used. During the process of network training, we utilized the Adam optimizer [19] with a learning rate of 0.0005 and a batch size of 16, resulting in improved performance. The choice of epoch is of great importance. When the number of epochs is large, we find that the loss value of the model loss function is very low and the convergence of the model is good, but the ultimate noise reduction effect becomes worse. Taking MSE as an example, the MSE of the test set will be lower than the MSE of the noise signal and the true value, because the MSE of the test set is obtained by calculating the input and output noise sequence signal, and the model learns from one noise distribution to another noise distribution, and loses the effect of noise reduction. Therefore, we use the early stopping mechanism, the epoch is 25, then the overfitting phenomenon is avoided, the model can learn the noise reduction ability, the performance of N2N algorithm is shown in Figure 8 [20]. The SNR is 23.56 and MSE is 0.0006 which are achieved by the N2N algorithm with CAE model while the SNR is 17.62 and MSE is 0.0024 based on the U-net model. Unlike the CAE model, the U-net model has a similar noise reduction effect as low-pass filtering (Figures 11 and 12).
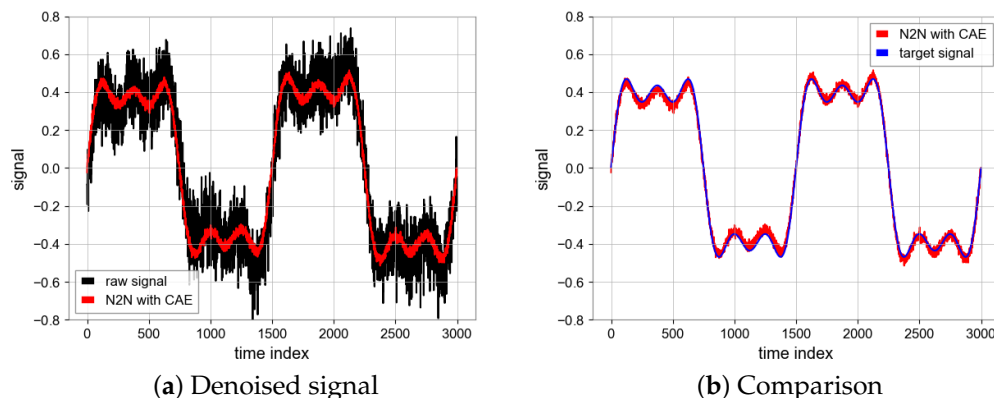
(**a**) Denoised signal   (**b**) Comparison

Figure 11. Denoised signal using CAE and comparison between denoised signal (Red) and target signal (Blue).

(**a**) Denoised signal   (**b**) Comparison
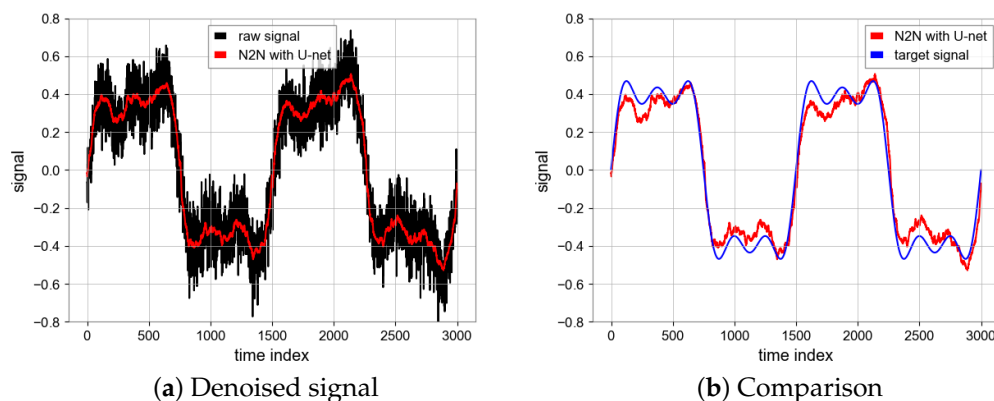
Figure 12. Denoised signal using U-net and comparison between denoised signal (Red) and target signal (Blue).

4.1.5. Comparison between Filters and N2N

Since the reconstruction layer of the model introduces high-frequency noise, we try to combine N2N and low-pass filter. It can be observed from Table 4 that N2N algorithm perform better than other filters in terms of noise reduction. As shown in Figure 13, the combination of N2N and the low-pass filter can further smooth the curve and achieve a higher SNR.
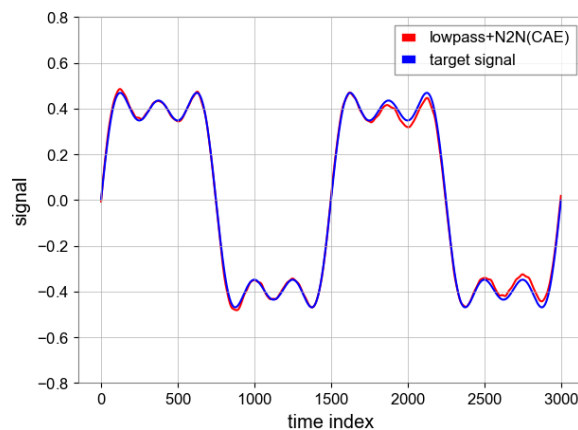

Figure 13. The result of combining N2N and low-pass filter.

Table 4. Comparison between Filters and N2N.

Filters	Wavelet Transform	Kalman Filter	Butterworth Filter	N2N	N2N+Lowpass Filter
SNR	12.64	16.54	16.58	23.56	25.28
MSE	0.0078	0.0033	0.0033	0.0006	0.0004

Impulse noise and non-stationary noise whose intensity changes with time are added to the simulation signal, and the final noise suppression results are shown in Figure 14.

(**a**) N2N (**b**) Combination of N2N and low-pass filter
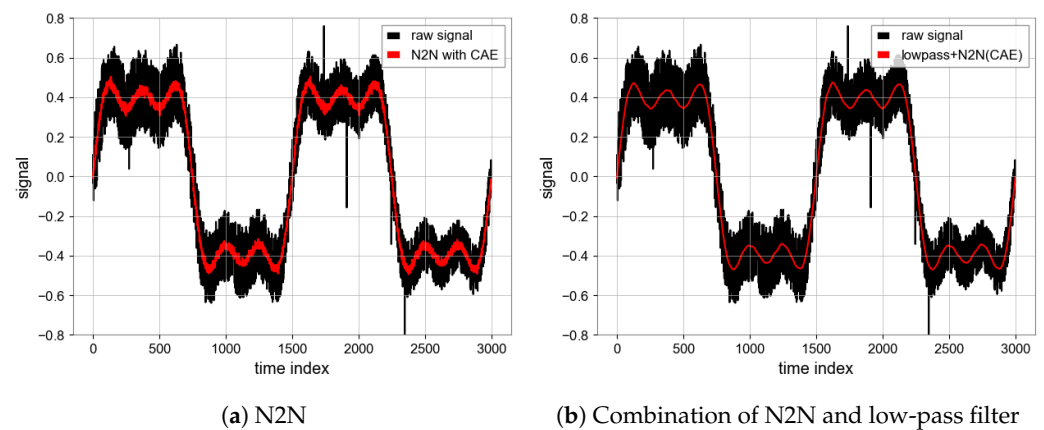
Figure 14. Noise suppression results of non-stationary noise and impulse noise.

4.2. Real Data Experiments

4.2.1. Taiji-1 Data

We take the Taiji-1 satellite as the experimental platform and upload maneuver commands to it, on 18 May 2022. Based on the data from the readout systems of AOS (attitude and orbit control subsystem) and inertial sensor in Taiji-1 satellite and the corresponding algorithm, where we can obtain the acceleration of TM and angular velocity of the satellite platform, we can calibrate the deviation between the COM of spacecraft and Test Mass [14].

Firstly, we choose the maneuver section, which is periodic and highly noisy. Secondly, the data have 7 cycles and there are obvious outliers in the beginning segment of each cycle. We remove them and fill them with non-null previous values. Then, the data are scaled to $(-1, 1)$ to fit the input and output size of the neural network model, which has 1500 dimensions. Finally, for the parameters of the network model, the 12 sets of data make it difficult for the neural network to converge, and we use the following method for data augmentation; when using the periodic sampler, we can choose to generate a pair of training samples at an interval of n periods, where $k = 2, 3, 4 \ldots$. In this way we generated 44 samples. We choose the CAE model mentioned above, and the specific training details are as follows; similar to the simulation experiment, early stopping mechanism is adopted, epoch is 25, dropout is 0.1, and batchsize is 16. We first analyze the acceleration data from the *y*-axis of the inertial sensor and the specific noise reduction results are as follows.

As shown in Figure 15, the N2N method can improve the signal-to-noise ratio of the target signal. We can use a low-pass filter and CRN filter to remove high-frequency noise which has little impact on target signal noise. After processing, the periodicity of data seems better. In the calibration scheme for COM calibration, we need the angular velocity measured by star tracker and the linear acceleration measured by the capacitive sensor in inertial sensor. Before calibration, the N2N algorithm was used for data processing and the periodic types of data are square and triangular wave signals.

In Figure 16, there are different period types, but N2N all achieve similar results, effectively improving the signal to noise ratio and making the data periodicity better and

signal-to-noise ratio higher. As shown in Figures 17 and 18, we will show the results of COM calibration with and without N2N algorithm.
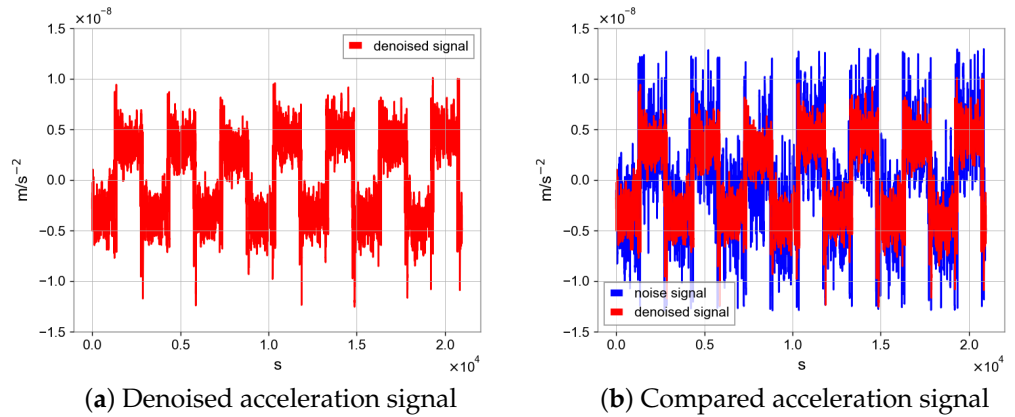

(**a**) Denoised acceleration signal   (**b**) Compared acceleration signal

Figure 15. Measured accelerations of the inertial sensor's *y*-axis with N2N and without N2N.

(**a**) Square wave signal   (**b**) Triangle wave signal
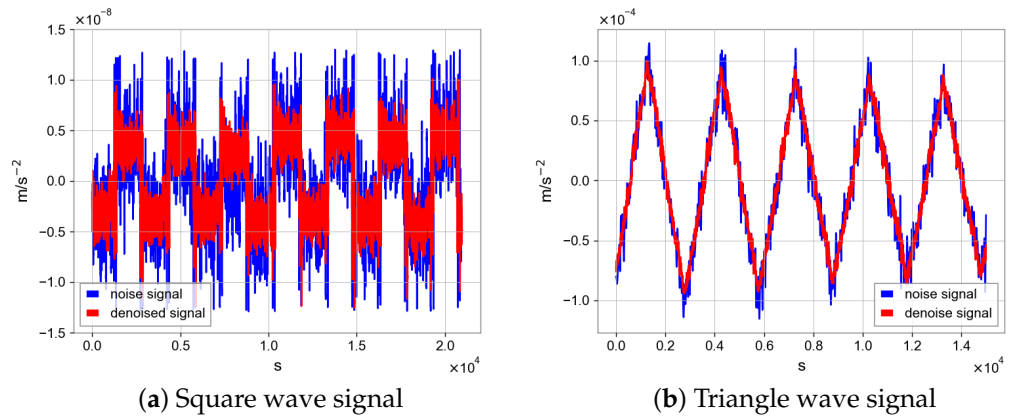
Figure 16. Different periodic signal processed by N2N algorithm.
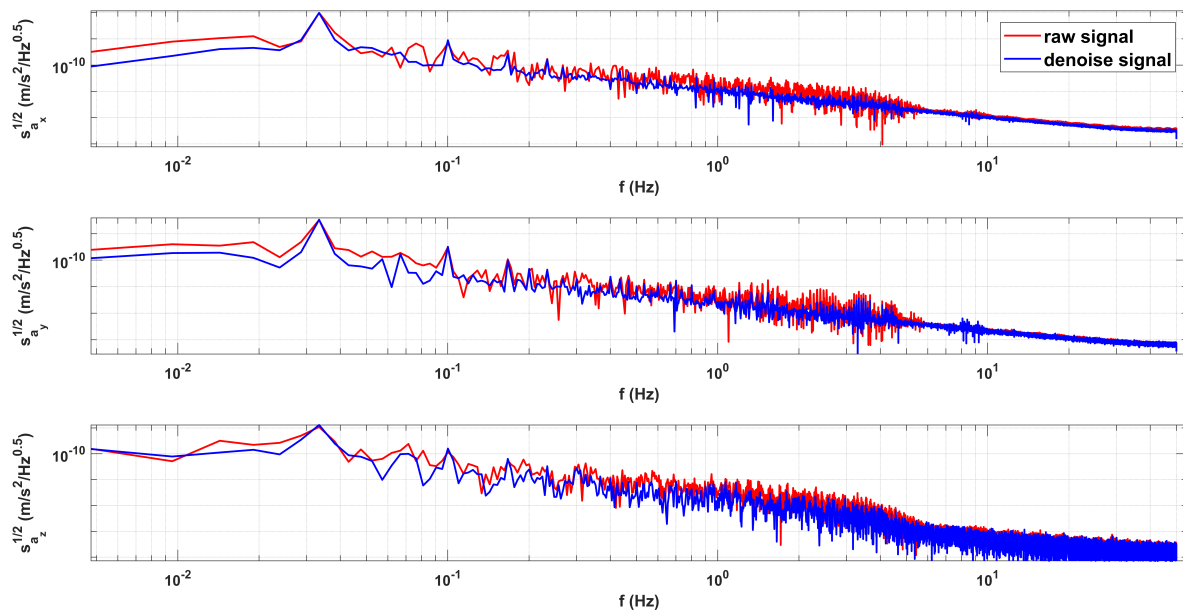
Figure 17. The comparison of Amplitude Spectral Density (ASD) for inertial sensor acceleration with and without N2N algorithm.

In Figure 17, the N2N method can better retain the peak value, effectively suppress the data on both sides of the peak, and improve the signal-to-noise ratio. At the same time, we can find that the resonance peak identification becomes clearer and the resonance peak with larger uncertainty becomes sharper and more obvious after processing. In Figure 18, we can find that after using the N2N algorithm, the data for centroid calibration task are better. Finally, the calibrated COM offset values with and without N2N are shown in Table 5.
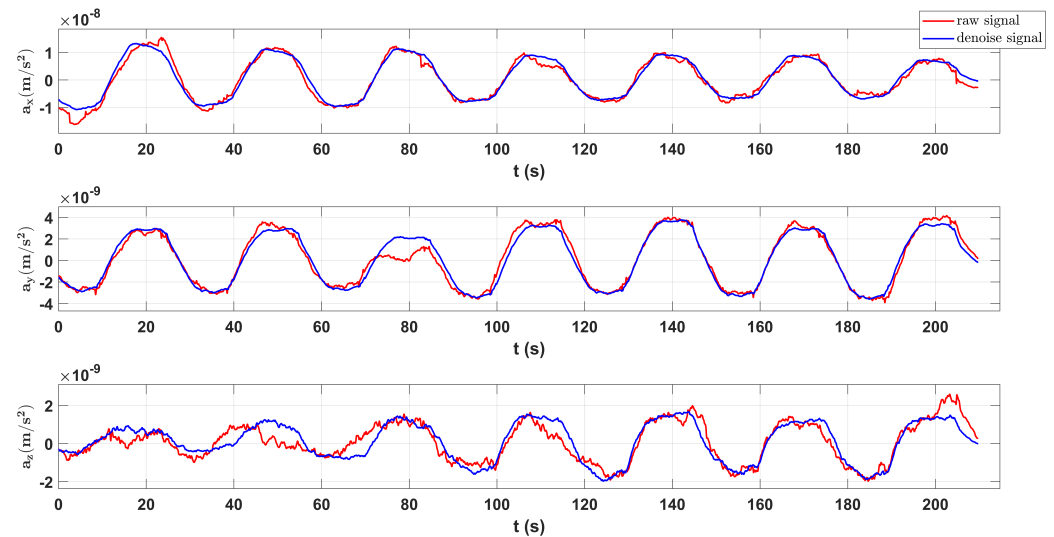
Figure 18. The comparison of raw signal and denoised signal.

Table 5. COM offset calibration results for Taiji-1 inertial sensor system with and without N2N.

COM Offset	Calibrated Value with N2N (mm)	Calibrated Value without N2N (mm)
x-axis	-0.0793	-0.1400
y-axis	0.3707	0.6270
z-axis	-0.8343	-0.8520

4.2.2. GRACE-FO Data

We consider the accelerometer data from GRACE-FO's dataset to analyze the noise reduction effect of the N2N method for general acceleration signals. The sampling frequency of Level-1A acceleration data is 10 Hz, and we choose 5 days of data for the experiment. The length of data per day is 864,000, the data are limited to $(-1,1)$ after preprocessing. By observing the data, it can be found that there are large high-frequency noise and obvious impulse noise in the signal. We did not remove the impulse noise in the preprocessing process in order to observe the denoising effect of the N2N method. In the specific experiment setup, there are some differences from the above text. We generate training samples by sampling noise signals. Our sampler principle is as follows. Every 10 points are divided into two parts, each part is sampled once and added to the input part and label part of the dataset, and finally 80 sets of data can be obtained. The input and output dimensions of the network models are 4800 and for the CAE model we have removed the final reconstruction layer. The time sequence diagram and ASD diagram of accelerometer signal after noise reduction are shown as follows.

As shown in Figure 19, the N2N method is effective in dealing with impulsive and high-frequency noise, even if the noise does not fully satisfy the zero-mean condition. The analysis of the ASD spectrum shows that the accelerometer signal processed by the N2N method (yellow line) can significantly suppress noise in the frequency range above 1mHz. Compared with traditional methods such as low-pass filtering, the N2N method can preserve the high-frequency part of the signal that may exist, rather than completely

filtering it out. It is observed in Figures 20 and 21 that there may be variable peaks that are recognized as derived features by deep learning. The CAE model has a better noise suppression effect than the U-net model.
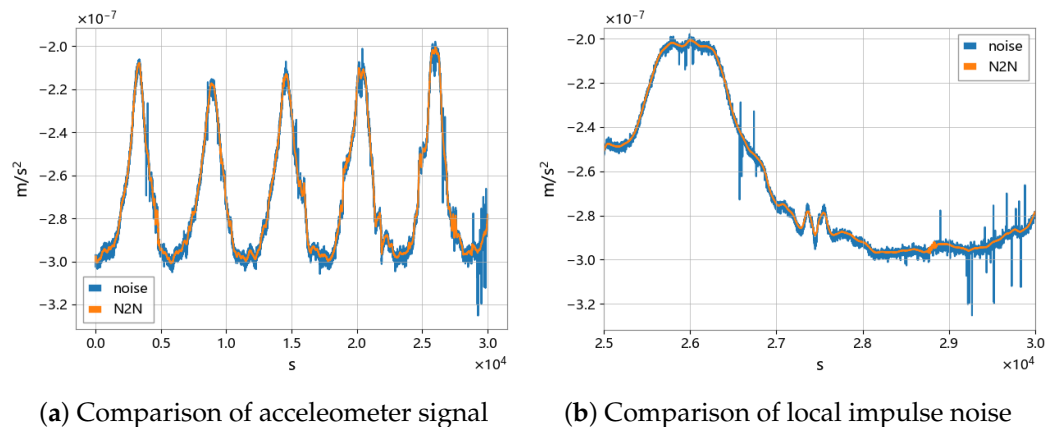
(**a**) Comparison of acceleometer signal   (**b**) Comparison of local impulse noise

Figure 19. Comparison of accelerometer signals before and after N2N noise reduction.
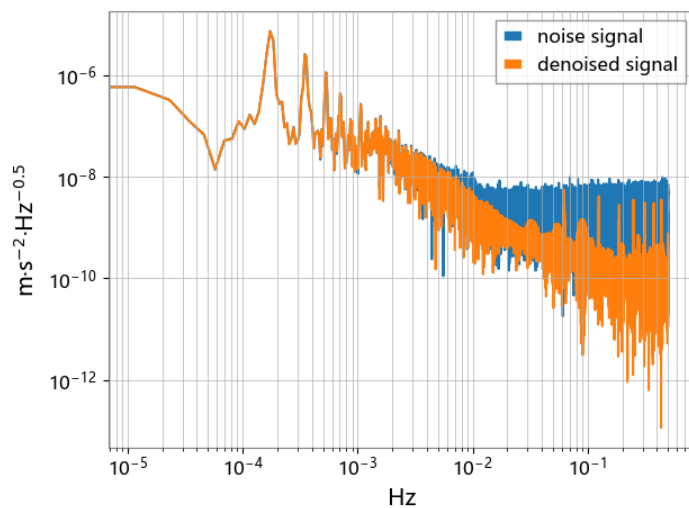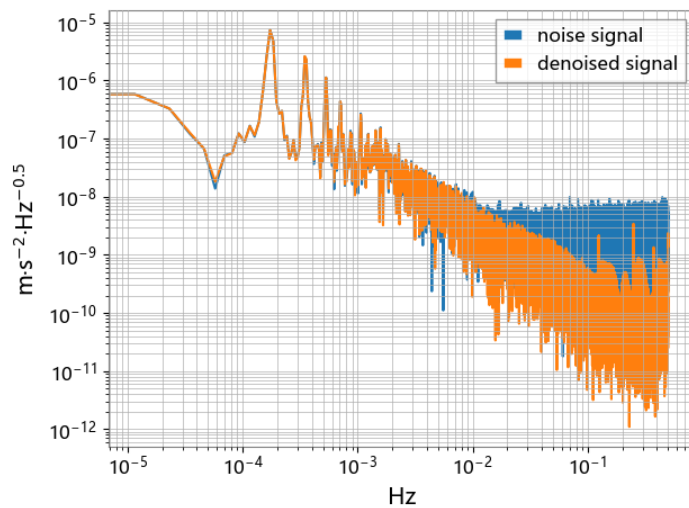
Figure 20. Comparison of ASD using the CAE model.

Figure 21. Comparison of ASD using the U-net model.

As shown in Figure 22, the difference between the data processed by the N2N method and the Level-1B data of GRACE-FO was analyzed. Due to the loss of edge information by the convolutional layer, the residual in the middle part is smaller. To solve this problem, a sliding window approach was employed with 50% overlap between adjacent samples, so that the edge parts can be excluded. The magnitude of the residual in the middle part is at the level of 10^{-10}.
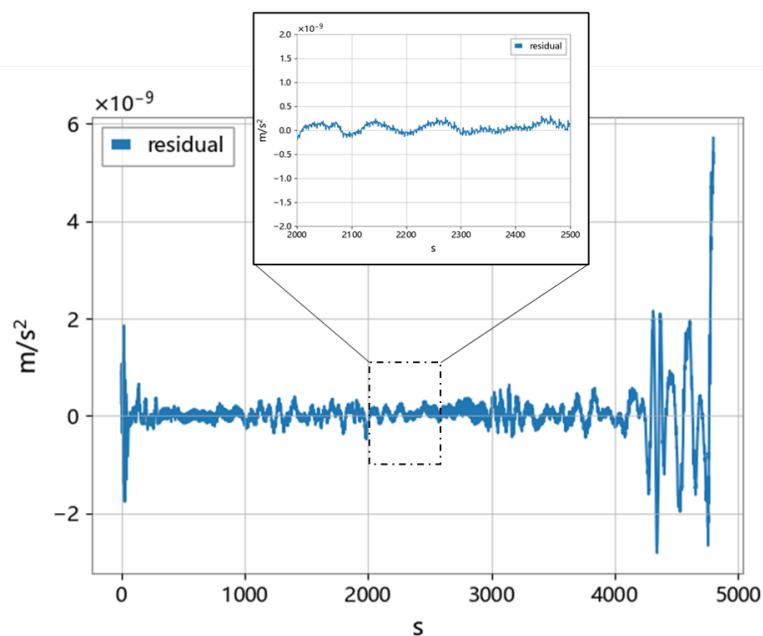

Figure 22. Residual between N2N method and Level$-$1B.

5. Conclusions

Due to the complex noise mechanism of inertial sensors, it is difficult to effectively suppress noise using traditional methods. Considering the characteristics of unknown true values and single signals in on-orbit measurement signals of inertial sensors, this paper proposes a broad-spectrum noise suppression method based on N2N unsupervised learning. The measurement signal is divided into periodic signals and general signals, and different sub-sequences and network model structures are obtained using different samplers. Using a simulated dataset as an example, better results were obtained compared to traditional signal processing methods, which can effectively remove the influence of mixed noise, and preliminarily demonstrate the feasibility of applying the N2N method to high-precision inertial data analysis and processing. Real data experiments were conducted on Taiji-1 and GRACE-FO, where the noise in Taiji-1 data was suppressed, and the high-frequency noise in GRACE-FO data was effectively suppressed, with data residuals at the 10^{-10} level compared to Level-1B data. In the future, we will study more complete N2N processing schemes and pipelines for the data processing needs of inertial sensors in gravity satellites and gravitational wave detection missions. Meanwhile, we will consider applying the N2N method more widely to high-precision sensor noise suppression.

Author Contributions: Z.Y. proposed the methodology, designed the experiments, and wrote the manuscript; H.Z. provided data and analyzed data; P.X. and Z.L. supervised the study and revised the manuscript. All authors have read and agreed to the published version of the manuscript.

Data Availability Statement: Not applicable.

References

1. Tapley, B.D.; Bettadpur, S.; Ries, J.C.; Thompson, P.F.; Watkins, M.M. GRACE measurements of mass variability in the Earth system. *Science* **2004**, *305*, 503–505. [CrossRef] [PubMed]
2. Bandikova, T.; McCullough, C.; Kruizinga, G.L.; Save, H.; Christophe, B. GRACE accelerometer data transplant. *Adv. Space Res.* **2019**, *64*, 623–644. [CrossRef]
3. Drinkwater, M.; Floberghagen, R.; Haagmans, R.; Muzi, D.; Popescu, A. GOCE: ESA's first Earth Explorer Core mission. In Proceedings of the Earth Gravity Field from Space—From Sensors to Earth Sciences: Proceedings of an ISSI Workshop, Bern, Switzerland, 11–15 March 2002; Springer: Berlin/Heidelberg, Germany, 2003; pp. 419–432.
4. Armano, M.; Audley, H.; Auger, G.; Baird, J.; Binetruy, P.; Born, M.; Bortoluzzi, D.; Brandt, N.; Bursi, A.; Caleno, M.; et al. The LISA pathfinder mission. *J. Phys. Conf. Ser.* **2015** *610*, 012005.
5. Luo, Z.; Guo, Z.; Jin, G.; Wu, Y.; Hu, W. A brief analysis to Taiji: Science and technology. *Results Phys.* **2020**, *16*, 102918. [CrossRef]
6. Luo, J.; Chen, L.S.; Duan, H.Z.; Gong, Y.G.; Hu, S.; Ji, J.; Liu, Q.; Mei, J.; Milyukov, V.; Sazhin, M.; et al. TianQin: A space-borne gravitational wave detector. *Class. Quantum Gravity* **2016**, *33*, 035010. [CrossRef]
7. Morvan, M.; Nikolaou, N.; Yip, K.H.; Waldmann, I. Don't Pay Attention to the Noise: Learning Self-supervised Representations of Light Curves with a Denoising Time Series Transformer. *arXiv* **2022**, arXiv:2207.02777.
8. Jiang, C.; Chen, S.; Chen, Y.; Zhang, B.; Feng, Z.; Zhou, H.; Bo, Y. A MEMS IMU de-noising method using long short term memory recurrent neural networks (LSTM-RNN). *Sensors* **2018**, *18*, 3470. [CrossRef] [PubMed]
9. Lehtinen, J.; Munkberg, J.; Hasselgren, J.; Laine, S.; Karras, T.; Aittala, M.; Aila, T. Noise2Noise: Learning Image Restoration without Clean Data. *arXiv* **2018**, arXiv:1803.04189.
10. Huang, T.; Li, S.; Jia, X.; Lu, H.; Liu, J. Neighbor2neighbor: Self-supervised denoising from single noisy images. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Nashville, TN, USA, 20–25 June 2021; pp. 14781–14790.
11. Li, Q.; Wu, J.; Kong, Y.; Yang, C.; Kong, Y.; Yang, G.; Senhadji, L.; Shu, H. Speech Denoising Using Only Single Noisy Audio Samples. *arXiv* **2021**, arXiv:2111.00242.
12. Koh, S.; Chia, C.S.; Tan, B.A. Underwater Signal Denoising Using Deep Learning Approach. In Proceedings of the Global Oceans 2020: Singapore–US Gulf Coast, Biloxi, MS, USA, 5–30 October 2020; pp. 1–6.
13. Takaai, T.; Tsutsui, M. Unsupervised Noise Reduction for Nanochannel Measurement Using Noise2Noise Deep Learning. In Proceedings of the Pacific-Asia Conference on Knowledge Discovery and Data Mining, Chengdu, China, 16–19 May 2021; Springer: Berlin/Heidelberg, Germany, 2021; pp. 44–56.
14. Zhang, H.; Xu, P.; Ye, Z.; Ye, D.; Qiang, L.E.; Luo, Z.; Qi, Q.; Cai, Z. A Systematic Approach for Calibrating the Inertial Sensor of Taiji-1. *arXiv* **2023**, arXiv:2304.00675.
15. Calvarons, A.F. Improved noise2noise denoising with limited data. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Nashville, TN, USA, 20–25 June 2021; pp. 796–805.
16. Srivastava, N.; Hinton, G.; Krizhevsky, A.; Sutskever, I.; Salakhutdinov, R. Dropout: A simple way to prevent neural networks from overfitting. *J. Mach. Learn. Res.* **2014**, *15*, 1929–1958.
17. Ebadi, L.; Shafri, H.Z.; Mansor, S.B.; Ashurov, R. A review of applying second-generation wavelets for noise removal from remote sensing data. *Environ. Earth Sci.* **2013**, *70*, 2679–2690. [CrossRef]
18. Alfian, R.I.; Ma'arif, A.; Sunardi, S. Noise reduction in the accelerometer and gyroscope sensor with the Kalman filter algorithm. *J. Robot. Control (JRC)* **2021**, *2*, 180–189. [CrossRef]
19. Kingma, D.P.; Ba, J. Adam: A Method for Stochastic Optimization. In Proceedings of the 3rd International Conference on Learning Representations (ICLR 2015), San Diego, CA, USA, 7–9 May 2015.
20. Duvenaud, D.; Maclaurin, D.; Adams, R. Early stopping as nonparametric variational inference. In Proceedings of the 19th International Conference on Artificial Intelligence and Statistics, Cadiz, Spain, 9–11 May 2016; pp. 1070–1077.