Research Article

A LSTM-RNN-Based Fiber Optic Gyroscope Drift Compensation

Ning Mao¹, Jiangning Xu¹, Jingshu Li², and Hongyang He¹

¹Department of Navigation Engineering, PLA Naval University of Engineering, Wuhan 430033, China

²Department of Operation and Planning, PLA Naval University of Engineering, Wuhan 430033, China

Correspondence should be addressed to Hongyang He; xgdhehongyang@163.com

Received 14 May 2021; Accepted 2 July 2021; Published 15 July 2021

Academic Editor: Venkatesan Rajinikanth

Copyright © 2021 Ning Mao et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Fiber optic gyroscope (FOG) inertial measurement unit (IMU) containing a three-orthogonal gyroscope and three-orthogonal accelerometer has been widely utilized in positioning and navigation of military and aerospace fields, due to its simple structure, small size, and high accuracy. However, noise such as temperature drift will reduce the accuracy of FOG, which will affect the resolution accuracy of IMU. In order to reduce the FOG drift and improve the navigation accuracy, a long short-term memory recurrent neural network (LSTM-RNN) model is established, and a real-time acquisition method of the temperature change rate based on moving average is proposed. In addition, for comparative analysis, backpropagation (BP) neural network model, CART-Bagging, classification and regression tree (CART) model, and online support vector machine regression (Online-SVR) model are established to filter FOG outputs. Numerical simulation based on field test data in the range of -20°C to 50°C is employed to verify the effectiveness and superiority of the LSTM-RNN model. The results indicate that the LSTM-RNN model has better compensation accuracy and stability, which is suitable for online compensation. This proposed solution can be applied in military and aerospace fields.

1. Introduction

FOG is a kind of optical fiber sensor based on Sagnac effect, which is widely used in strapdown inertial navigation system (SINS) and engineering at present. In addition, FOG has been broadly used in military and aerospace fields because of its simple structure, small size, and high accuracy. The accuracy of FOG is greatly affected by the surrounding environment, especially the environment temperature change [1], which will bring errors to the FOG output. Therefore, there are two ways to compensate FOG drift. The first one is temperature control and compensation in hardware, which is efficient but high cost and complex. The second one is compensation FOG drift in software by analyzing the test data, which is easy implementation and utility draws lots of the researcher's attention. Studies [2, 3] have shown that the FOG drift has a strong nonlinearity, and the general approach of polynomial linear fitting methods cannot express the nonlinear characteristics of FOG exactly. Some optimization algorithms and artificial neural networks [4] can better approximate nonlinear problems. With this

consideration, this paper is devoted to compensating FOG drift by using several fine artificial intelligence algorithms.

The past few decades have seen the fruitful methods' development for the compensation of FOG drift such as the neural network model [5, 6], autoregressive integrated moving average (AIMA) model [7, 8], adaptive compensation model [9, 10], and deep learning model [11, 12]. Pan et al. [13] applied the neural network method to the fitting and compensation of the FOG temperature drift model. However, random white noise in FOG output data was not considered, which might lead to overfitting of the network model. Guo et al. [14] used a BP neural network algorithm based on optimized prediction data, which could effectively reduce the influence of random white noise in FOG data on the model compensation accuracy. However, the establishment and learning process of the neural network belong to the nondeterministic polynomial (NP) problem. It is difficult to find a suitable network structure, which may lead the generalization ability of the neural network insufficient. Therefore, some researchers have introduced some optimization algorithms to tackle the limitations. The genetic

algorithm (GA) was applied by Pan et al. [15], the artificial fish swarm algorithm was applied by Song et al. [16], and the particle swarm optimization (PSO) was applied by Wang, Cheng, and Tong et al. [17–19]. The experimental results indicated that the optimized models are effective and can improve the system's performance. In addition, the idea of ensemble learning [20] is used by Liu et al. [21]. The BP-Bagging model of FOG temperature drift was established, which had better compensation effect than the linear regression model and BP neural network model.

In recent years, deep learning models, such as recurrent neural network (RNN) and convolutional neural network (CNN), have been rapidly developed and widely applied in the data science community. LSTM-RNN [22] is a time series prediction model, which can address the long-term memory compared with RNN. Jiang et al. [23] proposed a LSTM-RNN-based microelectromechanical system (MEMS) IMU errors' modelling method, which achieved better denoising effect than the ARMA model and performed excellently in time series signal processing. In addition, LSTM-RNN has performed excellently in speech recognition, Google translate, image recognition, and others [24–26]. Therefore, we establish the LSTM-RNN model to compensate FOG drift. Moreover, the BP model, CART-Bagging model, and Online-SVR model are established, respectively, for comparing and analyzing.

The structure of this paper is arranged as follows. The mechanism of FOG temperature drift is briefly reviewed in Section 2, with an emphasis to Shupe effect. In Section 3, the BP, CART-Bagging, Online-SVR, and LSTM-RNN are established, respectively. Simulation results based on field test data are reported in Section 4. Finally, Section 5 draws the conclusions.

2. Problem Statement

FOG drift is mainly caused by temperature change and vibration [27]. According to Shupe's study [28], the change of temperature will cause different refractive index of optical fiber in every sector of the fiber coil, which is called Shupe effect.

Denote that the optical wave of FOG propagates along the fiber coil with length L and refractive index n . When the environment temperature changes, the possible phase delay can be written as

$$\phi = \beta_0 nL + \beta_0 \left(\frac{\partial n}{\partial T} + n\alpha \right) \int_0^L \Delta T(z) dz, \quad (1)$$

where $\beta_0 = 2\pi/\lambda$ is the propagation constant of light in vacuum, λ is the light wavelength, n is the refractive index of the fiber coil, L is the length of the fiber coil, $\partial n/\partial T$ is the temperature coefficient of the refractive index, α is the temperature expansion coefficient of the optical fiber material, and $\Delta T(z)$ is the temperature gradient value of the fiber ring at position z .

According to (1), the thermally induced nonreciprocal phase delay produced by two beams along counterclockwise (CCW) and clockwise (CW) can be expressed as

$$\phi_{ccw}(t) = \beta_0 nL + \beta_0 \frac{\partial n}{\partial T} \times \int_0^L \Delta T\left(L - z, t - \frac{z}{c}\right) dz, \quad (2)$$

$$\phi_{cw}(t) = \beta_0 nL + \beta_0 \frac{\partial n}{\partial T} \times \int_0^L \Delta T\left(z, t - \frac{z}{c}\right) dz, \quad (3)$$

where $\phi_{ccw}(t)$ is the phase shift caused by the interference light propagating along the fiber ring counterclockwise, $\phi_{cw}(t)$ is the phase shift generated by the interference light propagating along the fiber ring clockwise, c is the propagation speed of the light, and z is the distance from any point to the end point.

According to (2) and (3), the Shupe nonreciprocal phase delay caused by temperature effects can be calculated as

$$\Delta\phi = \frac{\beta_0}{c_0} n \frac{\partial n}{\partial T} \int_0^{L/2} [\Delta T(z) - \Delta T(L - z)] \times (2z - L) dz, \quad (4)$$

where c_0 is the speed of light in vacuum.

The influence of temperature on FOG drift is mainly manifested in three aspects: temperature change, temperature gradient, and temperature change rate [6]. The common method in engineering is to approximate the variation of FOG drift in a linear relationship and establish a polynomial model. In fact, the temperature characteristics of FOG drift have complex nonlinear characteristics, and the general approach of polynomial linear fitting methods cannot express the nonlinear characteristics of FOG exactly.

3. Theory of Modelling for FOG

In this section, the conventional BP and CART-Bagging representing the postprocessing methods and the Online-SVR and LSTM-RNN representing real-time methods are presented. The principles, basic equations, and information flow are briefly introduced.

3.1. BP Model. Essentially, the BP neural network is a mapping from input to output, and a 3-layer forward network can approximate any continuous function at any expected accuracy. Therefore, a 3-layer neural network is established in this paper, and its structure is shown in Figure 1.

The output of the BP neural network can be given by

$$y = f_2(\mathbf{W}_2(f_1(\mathbf{W}_1 \times \mathbf{P} + \mathbf{b}_1)) + \mathbf{b}_2), \quad (5)$$

where $\mathbf{P}$ is the input vector, $\mathbf{W}_1$ and $\mathbf{W}_2$ are weight matrices, and $\mathbf{b}_1$ and $\mathbf{b}_2$ are bias vectors.

In this paper, the specific setting of important parameters of the BP model is shown in Table 1.

3.2. CART-Bagging Model. The decision tree model is a classification algorithm [29], which is similar with tree structure. Test the data from the root node; then, the data will be divided into different subsets by using different test results. And, each data sample subset constitutes a subnode.

FIGURE 1: Structure of the three-layer BP neural network.

TABLE 1: The parameters setting of the BP model.

Resolver	Hidden layer	Activation function	Initial learning rate	Learning rate
adam	20	relu	0.1	Adaptive

The CART regression tree is a kind of decision tree [29], which analyses continuously distributed data and supports the processing when the data feature is continuous distribution. The CART regression tree algorithm can be regarded as a collection of many if-then rules. Every path from the root node to the leaf node has a rule [30], and the leaf node corresponds to the conclusion of the rule. There is a termination condition in the process of training a model, which is the threshold condition from the root node to the leaf node.

At present, the most widely used ensemble learning algorithms are bagging and boosting [31]. The bagging algorithm uses the bootstrapping method to extract the training set and put them back after training. The prediction result is the average of the results of several base models, which are suitable for processing the temperature drift data of FOG. Therefore, as shown in Figure 2, we combine the CART regression tree with the bagging algorithm, which is recorded as the CART-Bagging model in this paper. The algorithm flow of CART-Bagging can be summarized as Algorithm 1.

In this paper, the maximum depth of CART is $d = 8$, the maximum number of leaf nodes is $m = 40$, and the number of decision trees is $L = 10$.

3.3. Online-SVR Model. The Online-SVR algorithm, an online regression modelling method, is a time series analysis method, which can realize online dynamic update [32, 33]. The main difference between the Online-SVR model and traditional SVR model is the way of data processing [34]. Compared with the traditional SVR method, the data in the Online-SVR model is gradually generated in an orderly manner, and the computer cost in each iteration does not increase.

Denote the initial data size as S . For any orthogonal basis $\rho_1, \rho_2, \dots, \rho_k$ in Hilbert space $\mathbf{H}$, $\sum_{i=1}^k \cos^2(\rho_i, \theta) = 1$ can be given by $\theta \in \mathbf{H}$, where $\cos(x, y)$ is the cosine function of vectors x and y and $\cos(x, y) = x^T y / (\|x\| \times \|y\|)$.

A vector sequence $\{\alpha_1, \alpha_2, \dots, \alpha_3\}$ is constructed firstly, where $\alpha_k = \Phi(x_i)y_j - \Phi(x_j)y_i$ and $k = 1, 2, l(l-1)/2$. Φ is a nonlinear mapping, which can map the input data to a high-

dimensional feature space. Then, an orthogonal vector sequence $\{\beta_1, \beta_2, \dots, \beta_d\}$ should be constructed, and

$$\beta_i = \frac{\alpha_i - \sum_{j=1}^d \beta_j (\beta_j^T \alpha_i)}{\|\alpha_i - \sum_{j=1}^d \beta_j (\beta_j^T \alpha_i)\|}, \quad (6)$$

where $d = \text{rank}\{\alpha_1, \alpha_2, \dots, \alpha_3\}$ can be obtained by Schmidt orthogonalization. $\{W, \beta_1, \dots, \beta_k\}$ is an orthogonal basis in space $\mathbf{H}$ and $\Phi(x_i) \in \mathbf{H}$, which yields

$$\sum_{j=1}^d \cos^2(\beta_j, \Phi(x_i)) + \cos^2(W, \Phi(x_i)) = 1, \quad (7)$$

where $\|W\|$ is the Euclidean norm of W , which can be formulated as

$$\|W\| = \frac{y_i}{\|x\| \sqrt{1 - \sum_{j=1}^d \cos^2(\beta_j, \Phi(x_i))}} \quad (8)$$

Therefore, the optimal kernel function parameters can be obtained by minimizing $\|W\|$. When a new sample $(l+1)$ th enters, we need to judge whether it is suitable for the current model. If $\alpha_{l+1} = \sum_{i=1}^d \beta_i \alpha_{l+1}^T \beta_i$, then the model remains unchanged. Otherwise, the optimal model needs to be updated. Calculate $A(d+1, i, j)$, $B(d+1, i, j)$, γ_{d+1} , $C(d+1, i, j)$, and β_{d+1} in turn for updating the model, which can be given by

$$\begin{aligned} C(l, i, j) &= B(l, i)y_j - B(l, j)y_i, \\ A(l, i, j) &= K(x_l, x_i)y_i - K(x_l, x_j)y_j, \\ B(l, i) &= \beta_l^T \Phi(x_i), \\ \gamma j &= \beta_j^T \Phi(x_{i^*}) = B(j, i^*), \end{aligned} \quad (9)$$

where i^* is any value in $\{1, 2, \dots, l\}$.

Therefore, the acquisition of the optimal model only needs to calculate the cost of the minimization optimization problem corresponding to the optimal kernel selection, and the whole calculation cost in each iteration is not increased compared with the traditional SVM training. The real-time compensation process is shown in Figure 3, and the algorithm flow of Online-SVR can be summarized as Algorithm 2.

FIGURE 2: CART-Bagging integrated model.

- (i) Initialization: the maximum depth of CART d ; the maximum number of leaf nodes m ; the number of decision trees L .
- (ii) Input: Training samples $S = \{(x_1, y_1), (x_2, y_2), \dots, (x_m, y_m)\}$
- (iii) for $M = 1, 2, 3, \dots$ do
- (iv) Extract $S(i)$ from S according to a set proportion.
- (v) Training CART-Bagging model with $S(i)$.
- (vi) If the termination condition is reached on a node
- (vii) The node is defined as a leaf node.
- (viii) Else
- (ix) Continue training.
- (x) $M = M + 1$
- (xi) end for
- (xii) Output: Average value of each basic model
- (xiii) M is the current time, which is an integer.

ALGORITHM 1: CART-Bagging. Algorithm procedure of the CART-Bagging model.

FIGURE 3: Real-time compensation process of Online-SVR.

3.4. LSTM-RNN Model. RNN is a deep learning algorithm, which is mainly used to deal with sequence prediction problems. As shown in Figure 4, the output of the last calculation “ $State_t$ ” and the input at the current time “ $Input_t$ ” are used in each cycle of RNN. The output of RNN contains all the previous information, which is the time series prediction.

However, in RNN, the more recent time has a greater impact on the current time, which may cause gradient disappearance or gradient explosion [35]. Compared with RNN, LSTM-RNN adds a structure that can transmit data at multiple time steps, which can save long-ago information and effectively solve the problems of gradient disappearance and gradient explosion. As shown in Figure 5, LSTM-RNN is

- (i) Initialization: $\text{Ker}(\Sigma) = \{K_1(\Sigma), K_2(\Sigma), \dots, K_p(\Sigma)\}$, where $K_i(\Sigma)$ is the kernel function of type i with continuously adjusting kernel parameter Σ .
- (ii) Input: Training samples $S = \{(x_1, y_1), (x_2, y_2), \dots, (x_m, y_m)\}$
- (iii) for $M = 1, 2, 3, \dots$ do
- (iv) Optimal kernel function selection.
- (v) For each $K_{\Sigma} = K(\Sigma) \in \text{Ker}(\Sigma)$, calculate the optimization problem $\Lambda^* = \arg\min_{\Lambda \in \Sigma} \{F(K_{\Lambda})\}$, where $F(K) = \|W\|^2 = (y_{i^*}^2 / (K(x_{i^*}, x_{i^*}) - \sum_{j=1}^d \gamma_j^2))$.
- (vi) The optimal kernel is $K^*(\Sigma^*) = K_o^*(\Lambda^*)$, where $o^* = \arg\min_{1 \leq i \leq p} (K_i(\Lambda^*))$.
- (vii) $M = M + 1$
- (viii) end for
- (ix) M is the current time, which is an integer.

ALGORITHM 2: Online-SVR. Algorithm procedure of the Online-SVR model.

FIGURE 4: The basic structure of RNN.

FIGURE 5: The basic structure of LSTM-RNN.

affected by the input of the previous layer, the pervious time state of the current layer, and the internal state of the LSTM-RNN.

A LSTM unit is composed of an “output gate,” a “forget gate,” an “input gate,” and a “memory cell,” and the state vector of LSTM can be set by i_t , f_t , o_t , and c_t . The algorithm flow of LSTM-RNN can be summarized as Algorithm 3.

It should be noted that BP and CART-Bagging belong to postprocessing algorithms, which are only used as controlled experimental. Online-SVR and LSTM-RNN belong to time series analysis methods, which can realize real-time compensation. This paper aims to verify the effect of LSTM-RNN.

4. Experimental Study

In this section, a real-time acquisition method of the temperature change rate based on moving average is proposed. And, we design a temperature test. Three FOGs, which are

denoted by Gyro 1 to Gyro 3 in this paper, are used to validate the compensation methods abovementioned.

4.1. Temperature Experiment for FOGs. In order to evaluate the performance of the LSTM-RNN method, a temperature control experiment was carried out. In the experiment test, the gyro unit is fixed in a horizontal static temperature chamber, and gyroscopes are only affected by the rotation of the Earth. The temperature range of this experiment is -20°C to 50°C . The raw datasets of FOGs are logged and saved, which are applied to verify our established model. The update rate of FOGs is 200 Hz. In Figure 6, the smoothed Gyro 1 original output, temperature change, and temperature change rate with time are shown, and the average value for 1 minute data is used for 1 cluster. Gyro 1 drift is greatly affected by the temperature change rate.

Although the effectiveness of BP, CART-Bagging, and Online-SVR to compensate FOG drift has been proved in our previous work, it is necessary to experiment again.

- (i) Initialization: Time steps H
- (ii) Input: Training samples $S = \{(x_1, y_1), (x_2, y_2), \dots, (x_m, y_m)\}$
- (iii) for $M = 1, 2, 3, \dots$ do
- (iv) $i_t = \sigma(W_3 x_t + W_{\text{rec}3} y_{(t-1)} + W_{p3} c_{(t-1)})$
- (v) $f_t = \sigma(W_2 x_t + W_{\text{rec}2} y_{(t-1)} + W_{p2} c_{(t-1)})$
- (vi) $o_t = \sigma(W_1 x_t + W_{\text{rec}1} y_{(t-1)} + W_{p1} c_{(t-1)})$
- (vii) $l_t = \tanh(W_4 x_t + W_{\text{rec}4} y_{(t-1)})$
- (viii) $c_t = f_t c_{(t-1)} + i_t l_t$
- (ix) $y_t = \tanh(c_t) o_t$
- (x) $M = M + 1$
- (xi) end for
- (xii) W_i is input connection, $W_{\text{rec}i}$ is recurrent connection, and W_{pi} is peephole connection, $i = 1, 2, 3$, and 4. σ is the sigmoid function.

ALGORITHM 3: LSTM-RNN. Algorithm procedure of the LSTM-RNN model.

FIGURE 6: Gyro 1 drift, temperature change, and temperature change rate with time (blue line indicates Gyro 1 drift; black line indicates temperature gradient; red line indicates temperature; min = minute; °C = centigrade).

Actually, the superiority of compensating FOG drift can also be illustrated through the comparison of BP, CART-Bagging, Online-SVR, and LSTM-RNN, which will be shown in the following study.

4.2. Real-Time Acquisition of Temperature Change Rate. Different from the offline model, in order to realize the online compensation, the issue of real-time and stable acquisition of the temperature change rate needs to be solved first. Temperature change is a time series problem. The temperature at time i is only related to the environmental temperature at time i and before time i . Based on this characteristic, as shown in Figure 7, this paper establishes a real-time acquisition method of the temperature change rate based on moving average.

A moving window with length K is constructed. Firstly, a section of temperature data with length K is output; then, the

temperature T_1 at $K/2$ is the average value of the previous K temperature data. Therefore, the temperature change rate at any time i can be given by

$$\dot{T}(i) = \left(\frac{\sum_{m=i-K/2+1}^{i+K/2+1} T(m)}{K} - \frac{\sum_{j=i-K/2}^{i+K/2} T(j)}{K} \right) \times f, \quad (10)$$

where $T(i)$ is the temperature at time i , $\dot{T}(i)$ is the temperature change rate at time i , and f is the output frequency of temperature data.

According to (10), the noise pollution caused by temperature sensor measurement and data transmission can be effectively reduced by choosing an appropriate moving average window length K . In addition, this method is effective to avoid the temperature change rate to be zero, which is caused by high temperature measurement frequency and is harmful for compensation models.

FIGURE 7: Real-time acquisition of the temperature change rate.

FIGURE 8: FOG drift before and after compensation of Gyro 1.

FIGURE 9: FOG drift before and after compensation of Gyro 2.

In this paper, the moving average window length is $k = 12000$, the length of each movement is 1, and the real-time temperature change rate $\dot{T}(i)$ is calculated by (10).

4.3. Result Analysis. The LSTM-RNN model is built for real-time compensation, and the compensation effect is compared with the BP model, CART-Bagging model and

Online-SVR model. The experimental results are shown in Figures 8–10. The image on the left is FOG drift before compensation, and the image on the right is residual drift after compensation.

The results show that the BP model and CART-Bagging model have poor performance in compensation experiments of Gyro 1 and Gyro 2. The Online-SVR model has poor performance in the compensation experiment of Gyro 3.

FIGURE 10: FOG drift before and after compensation of Gyro 3.

TABLE 2: Comparison of the statistical results before and after compensation with four models of Gyro 1 ($^{\circ}/h$).

Schemes	Average	Standard deviation	MAE	RMSE
Before compensation	0.19256	0.00351	—	—
BP	-0.04144	0.03380	0.04156	0.05347
CART-Bagging	0.01222	0.02737	0.01997	0.02998
Online-SVR	-0.01329	0.00906	0.01381	0.01609
LSTM-RNN	-0.00775	0.00276	0.00776	0.00823

$^{\circ}$ = degree; h = hour.

TABLE 3: Comparison of the statistical results before and after compensation with four models of Gyro 2 ($^{\circ}/h$).

Schemes	Average	Standard deviation	MAE	RMSE
Before compensation	0.19291	0.0039	—	—
BP	-0.02751	0.02519	0.02751	0.03724
CART-Bagging	0.00977	0.01917	0.01552	0.02151
Online-SVR	-0.01565	0.00694	0.01568	0.01712
LSTM-RNN	-0.01225	0.0029	0.01225	0.01259

$^{\circ}$ = degree; h = hour.

TABLE 4: Comparison of the statistical results before and after compensation with four models of Gyro 3 ($^{\circ}/h$).

Schemes	Average	Standard deviation	MAE	RMSE
Before compensation	0.19355	0.00131	—	—
BP	-0.00746	0.00567	0.00874	0.00937
CART-Bagging	-0.00542	0.00294	0.00547	0.00616
Online-SVR	-0.00366	0.01001	0.00848	0.01065
LSTM-RNN	-0.00358	0.00277	0.00387	0.00453

$^{\circ}$ = degree; h = hour.

And, the LSTM-RNN model has stable performance and good effect in the three experiments.

In this paper, the mean and standard deviation of FOG drift before and after compensation are used to compare the compensation effect of the models. In addition, the mean

absolute error (MAE) and root mean square error (RMSE) are used to compare the stability and effect of the models. MAE and RMSE are calculated as (11) and (12), and the smaller the value of RMSE and MAE, the better the model compensation effect:

$$\text{RMSE}(X, h) = \sqrt{\frac{1}{m} \sum_{i=1}^m (h(x^{(i)}) - y^{(i)})^2}, \quad (11)$$

$$\text{MAE}(X, h) = \frac{1}{m} \sum_{i=1}^m |h(x^{(i)}) - y^{(i)}|. \quad (12)$$

From Tables 2 to 4, the LSTM-RNN model has better compensation accuracy and stability than the BP model, CART-Bagging model, and Online-SVR model. In addition, the model has high operation efficiency, which can meet the real-time requirements of compensation.

5. Conclusion

This paper discussed a FOG drift modelling method based on LSTM-RNN. Three FOGs were employed for a temperature control experiment to test the proposed method. By analyzing and preprocessing the FOG outputs from -20°C to 50°C , we proposed a real-time temperature change rate acquisition method based on moving average. Through comparing the compensation effect of the BP neural network, CART-Bagging, and Online-SVR, the results show that LSTM-RNN has better compensation accuracy and stability, and the model has high operation efficiency, which can meet the real-time requirements of compensation.

In addition, there are two details that are worthy to investigate in the future: (1) a well-trained LSTM-RNN should be trained with massive data and optimized hyperparameters; more experiments are needed in future; (2) some variants of RNN, such as BRNN [36] and DRNN [37], have been proposed in recent years. It is meaningful to compare the effects of different RNN on compensating FOG drift.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was supported by the National Key Research and Development Plan under Grant 2016YFB0501700 and 2016YFB0501701 and National Nature Science Foundation of China under Grant 41804076.

References

- [1] Y. Zhang, Z. Gao, G. Wang, and W. Gao, "Modeling of thermal-induced rate error for FOG with temperature ranging from -40°C to 60°C ," *IEEE Photonics Technology Letters*, vol. 26, no. 1, pp. 18–21, 2014.
- [2] X. Chen and C. Shen, "Study on temperature error processing technique for fiber optic gyroscope," *Optik*, vol. 124, no. 9, pp. 784–792, 2013.
- [3] L. Wang, G. Wei, Y. Zhu, J. Liu, and Z. Tian, "Real-time modeling and online filtering of the stochastic error in a fiber optic current transducer," *Measurement Science and Technology*, vol. 27, no. 10, Article ID 105103, 2016.
- [4] N. Ning Wang, M. J. Meng Joo Er, and M. Min Han, "Generalized single-hidden layer feedforward networks for regression problems," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 26, no. 6, pp. 1161–1176, 2015.
- [5] A. El-Rabbany and M. El-Diasty, "An efficient neural network model for de-noising of MEMS-based inertial data," *Journal of Navigation*, vol. 57, no. 3, pp. 407–415, 2004.
- [6] R. Rong Zhu, Y. Yanhua Zhang, and Q. Qilian Bao, "A novel intelligent strategy for improving measurement precision of FOG," *IEEE Transactions on Instrumentation and Measurement*, vol. 49, no. 6, pp. 1183–1188, 2000.
- [7] M. Narasimhappa, J. Nayak, M. H. Terra, and S. L. Sabat, "ARMA model based adaptive unscented fading Kalman filter for reducing drift of fiber optic gyroscope," *Sensors and Actuators A: Physical*, vol. 251, pp. 42–51, 2016.
- [8] J. Sun, X. Xu, Y. Liu, T. Zhang, and Y. Li, "FOG random drift signal denoising based on the improved AR model and modified sage-husa adaptive kalman filter," *Sensors*, vol. 16, no. 7, p. 1073, 2016.
- [9] L. Huang, Z. Li, F. Xie, and K. Feng, "Novel time series modeling methods for gyro random noise used in Internet of Things," *IEEE Access*, vol. 6, pp. 47911–47921, 2018.
- [10] Y. Feng, X. Li, and X. Zhang, "An adaptive compensation algorithm for temperature drift of micro-electro-mechanical systems gyroscopes using a strong tracking kalman filter," *Sensors*, vol. 15, no. 5, pp. 11222–11238, 2015.
- [11] Z. Tian, X. Gao, and K. Li, "Time-delay prediction method of networked control system based on EMD and LS-SVM," *Acta Electronica Sinica*, vol. 42, no. 5, pp. 868–874, 2014.
- [12] S. Dai, Q. Chen, Z. Liu, and H. Dai, "Time series prediction based on EMD-LSTM model," *Journal of Shenzhen University Science and Engineering*, vol. 37, no. 3, pp. 221–230, 2020.
- [13] H. Pan, A. Li, and B. Hu, "Application of chaos-BP combined artificial neural network in predicting FOG temperature drift," *Journal of the Chinese Institute of Engineers*, vol. 6, pp. 71–75, 2006.
- [14] S. Guo, J. Xu, and F. Li, "Modeling and compensation algorithm of FOG temperature drift with optimized BP neural network," *Journal of the Chinese Institute of Engineers*, vol. 24, no. 01, pp. 93–97, 2016.
- [15] C. Gu, J. Lu, Y. Wang, and J. Wang, "Temperature compensation of FOG based on GA-BP neural network," *Instrumentation Technology and Sensor*, vol. 03, pp. 113–116, 2018.
- [16] R. Song, X. Chen, C. Shen, H. Zhang, and X. Dong, "Modeling FOG drift using back-propagation neural network optimized by artificial fish swarm algorithm," *Journal of Sensors*, vol. 2014, Article ID 273043, 6 pages, 2014.
- [17] W. Wang and X. Chen, "Temperature drift modeling and compensation of fiber optical gyroscope based on improved support vector machine and particle swarm optimization algorithms," *Applied Optics*, vol. 55, no. 23, pp. 6243–6250, 2016.
- [18] J. Cheng, J. Fang, W. Wu, and J. Li, "Temperature drift modeling and compensation of RLG based on PSO tuning SVM," *Measurement*, vol. 55, pp. 246–254, 2014.
- [19] L. Tong, F. Qin, and K. Feng, "Segmentation compensation method for FOG temperature error based on particle swarm optimization algorithm," *Journal of the Chinese Institute of Engineers*, vol. 27, no. 04, pp. 505–509, 2019.

- [20] J. Mendes-Moreira, C. Soares, A. M. Jorge, and J. F. D. Sousa, "Ensemble approaches for regression," *ACM Computing Surveys*, vol. 45, no. 1, pp. 1–40, 2012.
- [21] Y. Liu, G. Yang, and S. Li, "Application of BP-Bagging model in temperature compensation of fiber optic gyro," *Journal of the Chinese Institute of Engineers*, vol. 22, no. 2, pp. 254–259, 2014.
- [22] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, 1997.
- [23] C. Jiang, S. Chen, Y. Chen et al., "A MEMS IMU de-noising method using long short term memory recurrent neural networks (LSTM-RNN)," *Sensors*, vol. 18, no. 10, p. 3470, 2018.
- [24] A. Graves, N. Jaitly, and A. R. Mohamed, "Hybrid speech recognition with deep bidirectional LSTM," *IEEE, in Proceedings of the 2013 IEEE Workshop on Automatic Speech Recognition and Understanding*, pp. 273–278, Olomouc, Czech Republic, December 2013.
- [25] M. Johnson, M. Schuster, Q. V. Le et al., "Google's multi-lingual neural machine translation system: enabling zero-shot translation," *Transactions of the Association for Computational Linguistics*, vol. 5, pp. 339–351, 2017.
- [26] G. Zhu, L. Zhang, P. Shen, and J. Song, "Multimodal gesture recognition using 3-D convolution and convolutional LSTM," *IEEE Access*, vol. 5, pp. 4517–4524, 2017.
- [27] X. Li, C. Zhang, Z. He, and Z. Zhong, "Temperature errors of IFOG and its compensation in engineering application," in *Proceedings of the 9th International Conference on Electronic Measurement and Instruments (ICEMI '09)*, pp. 2230–2234, Beijing, China, August 2009.
- [28] D. M. Shupe, "Thermally induced nonreciprocity in the fiber-optic interferometer," *Applied Optics*, vol. 19, no. 5, pp. 654–655, 1980.
- [29] R. J. Lewis, *An Introduction to Classification and Regression Tree (CART) Analysis* Annual Meeting of the Society for Academic Emergency Medicine, San Francisco, CA, USA, 2000.
- [30] R. L. Lawrence and A. Wright, "Rule-based classification systems using classification and regression tree (CART) analysis," *Photogrammetric Engineering and Remote Sensing*, vol. 67, no. 10, pp. 1137–1142, 2001.
- [31] M. A. Hassan, A. Khalil, S. Kaseb, and M. A. Kassem, "Exploring the potential of tree-based ensemble methods in solar radiation modeling," *Applied Energy*, vol. 203, pp. 897–916, 2017.
- [32] F. Mei, Y. Pan, K. Zhu, and J. Zheng, "A hybrid online forecasting model for ultrashort-term photovoltaic power generation," *Sustainability*, vol. 10, no. 3, p. 820, 2018.
- [33] J. Liu and E. Zio, "An adaptive online learning approach for Support Vector Regression: online-SVR-FID," *Mechanical Systems and Signal Processing*, vol. 76–77, pp. 796–809, 2016.
- [34] G. Esposito and M. Martin, "Bellman residuals minimization using online support vector machines," *Applied Intelligence*, vol. 47, no. 3, pp. 670–704, 2017.
- [35] Y. Liu, L. Guan, C. Hou et al., "Wind power short-term prediction based on LSTM and discrete wavelet transform," *Applied Sciences*, vol. 9, no. 6, p. 1108, 2019.
- [36] J. Chen, N. S. Chaudhari, "Cascaded bidirectional recurrent neural networks for protein secondary structure prediction," *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, vol. 4, no. 4, pp. 572–582, 2007.
- [37] C. C. Chao-Chee Ku and K. Y. Lee, "Diagonal recurrent neural networks for dynamic systems control," *IEEE Transactions on Neural Networks*, vol. 6, no. 1, pp. 144–156, 1995.