De-Noising Algorithm for Flight Data Recording System Based on Modified Ensemble Empirical Mode Decomposition

DAI Shaowu, CHEN Qiangqiang* and DAI Hongde
Naval Aviation University, China, Yantai, 264000.

*1195275597@qq.com

Abstract. The Flight Data Recording System (FDRS) records a lot of parameters of the aircraft during flight, which can be used for the test-flying, training mission of aircraft and so on. Effected by the working environment, information interference and its non-stability, the outliers and noise often exists in the FDRS data. These noises and outliers have a great impact on the use of FDRS. The aim of this paper is to remove outliers and de-noising of navigation data in FDRS. The causes of outliers and noise in FDRS data are analyzed firstly, with a reference suggestion proposed. Then the Letts criterion is used to remove outliers and the Modified Ensemble Empirical Mode Decomposition (MEEMD) is applied to achieve de-noising for FDRS. Results demonstrate that outliers are removed and the navigation data are de-noised effectively.

1. Introduction
FDRS were first introduced in the 1950s, which also known as black box [1]. The critical mission of an airplane FDRS is to collect, record and save flight parameters from a variety of airplane units to survive an accident [2]. Depending on the service life of an airplane, the FDRS may consists of an analog or a digital data. With the development of FDRS, which can record up to thousands of flight parameters for 25 hours, such as acceleration, air speed, time, altitude and attitude of aero engine, et al [3]. These data provide abundant resources with SINS and engine capability analysis [4].

However, FDRS is disturbed by many factors such as working environment and vibration noise when recording data [5]. For example, the electromagnetic interferences and pulse signals generated by sensor faults will cause the recorded flight data to contain abnormal signals and noise signals inevitably, which severely affect the subsequent analysis and application of flight data [6]. Therefore, it is essential to de-noising for the FDRS before data analysis.

At present, various methods have been developed to remove outliers from the obtained data. The Auto Regressive Moving Average (ARMA) model is extensively utilized in constructing a de-noising model, which can be judge whether there are outliers in the obtained data [7]. In [8], the Observer / Kalman filter identification (OKID) method is applied to remove outliers in flight data, and the estimated values are used to replace them, but the experiment data is a simulation signal, which has no practical application in FDRS. In [9], a flight data novelty detection method based on improved support vector data description (SVDD) method is proposed to realize automatic flight data interpretation. Besides, this method has a repeated process of adjustment, and the choice of parameters of SVDD has a great influence on the results.

Wavelet analysis has the capability of multi resolution decomposition and reconstruction, extracting features from the upper and lower frequencies of the data. In [10], the simulation of data de-
noising to pitch angle of a type aircraft was carried out based on wavelet de-noising. In [11], wavelet neural network is applied to flight data pre-processing. However, the de-noising results of Wavelet analysis depend largely on the choice of wavelet basis and decomposition levels. In [12], combined Ensemble empirical mode decomposition (EEMD) with Hilbert Transform (HT) is used for quick access recorder data. Besides, the calculation amount grows and the completeness loses due to the white noise unneutralized completely of EEMD.

To further improve the de-noising effects, a new de-noising algorithm based on Letts criterion and MEEMD is studied in this paper. This method uses Letts criterion to eliminate the outliers of FDRS data. Finally, the MEEMD is used to eliminate or reduce the noise in FDRS data in this paper.

2. Remove outliers
Outliers are always deviation from most real values of original data. The definition of outliers is: one or more observation points in a set of observation data that violate the data variation rule are called outliers [13].

Affected by the working environment of aircraft, the FDRS data is effected by some impulse signals or vibration, which leads to the existence of outliers. Letts criterion (3σ criterion) is used for judging outliers [14]. Set the original FDRS data is \( S \) have \( n \) points. The average value of original FDRS data is \( \bar{S} \). The residual series is \( 
\begin{align*}
\text{n}S
\end{align*}
\hat{S}=S-\bar{S}
\end{align*}
, have \( n-1 \) points. The standard deviation of \( \hat{S} \) is:

\[
\sigma = \sqrt{\frac{\sum_{i=1}^{n} \hat{S}^2}{n-1}}
\] (1)

Time series follows the characteristics of Gauss distribution, the probability of the residual falling in the interval \([-3\sigma, 3\sigma]\) is about 99.7%. When the residual of a measured value is greater than \( 3\sigma \), the measured value is regarded as an outlier. According to the \( 3\sigma \) criterion, the outliers of FDRS can be removed effectively.

After remove outliers by \( 3\sigma \) criterion, the location of outliers should be reconstruct. For the detected abnormal value of \( x_i \), the first-order difference method is used to replace the abnormal information of the point [15]:

\[
\hat{x}_i = 2x_{i-1} - x_{i-2}
\] (2)

Where \( \hat{x} \) is the replacement value of outlier \( x_i \). With this substitution, we can reconstruct the data without gross errors, and make preparation for data de-noising.

3. Modified Ensemble Empirical mode decomposition

3.1. Empirical mode decomposition
EMD, invented by Huang in 1998, which is a newly developed useful algorithm for non-linear and non-stationary time series analysis [16]. Based on EMD, Any complicated signal can be decomposed into several number of intrinsic mode functions (IMF) and a residual series. In physics, the IMFs represent the dynamic characteristics of the original signal, and the residual series is the trend item of the original signal. Compared with the original signal, the IMFs and the residual series have less complexity. Obviously, it can reduce the non-linear and non-stationary of original signal by EMD. The process of decomposing an IMF as follows is called “sifting” [17]:

1. Determine all the local extrema of the original signal \( x(t) \).
2. Determine the upper envelope \( e_n(t) \) with the local maxima by a cubic spline.
3. Determine the lower envelope \( e_n(t) \) with the local minima by a cubic spline.
4. Calculate the average of the upper and lower envelope

\[
m(t) = \frac{(e_{n}(t) + e_{n}(t))}{2}
\] (3)
(5) Calculate the subtract of the original signal \( x(t) \) and \( m(t) \).

\[
h = x(t) - m(t)
\] (4)

(6) Use the IMF criteria to detect \( h \) is the IMF or not. If not, make the \( x(t)=h \) and repeats the above steps; If \( h \) is the IMF, mark the IMF \( c(n) \).

\[
c(n) = h
\] (5)

(7) Calculate the residual series \( r \), and detect whether \( r \) meets the stopping criteria.

From the above steps, the original signal \( x(t) \) can induce that:

\[
x(t) = \sum_{j=1}^{n} h_n(t) + r
\] (6)

3.2. Modified Ensemble Empirical mode decomposition

In the EMD algorithm, a major problem in the decomposition process is the mode mixing, that is, the signals of different scales and frequencies appear in the same IMF component, or the signals of the same scale and frequency are decomposed into multiple IMF components [18]. Multiple modes are mixed in several IMF components, which affects the physical meaning of IMF components and is not conducive to the establishment of subsequent time series prediction models. Using MEEMD method for reference in dealing with mode mixing problem. The steps are as follows [19]:

(1) For the original signal \( s_i \), white noise signal sequences \( n_i(t) \) and \(-n_i(t)\) with 0 mean and opposite numbers are added respectively. \( a_i \) denotes the noise amplitude and \( i \) denotes the added logarithm of white noise:

\[
S_i^+ = S_i + a_i n_i(t)
\]

\[
S_i^- = S_i - a_i n_i(t)
\] (7)

(2) The sequence in formula (7) is decomposed by EMD, and the first order IMF components are \( I_i^+(t) \) and \( I_i^-(t) \). Integrating this pair of IMF components:

\[
I_1(t) = \frac{1}{2N} \sum_{i=1}^{N_e} [I_i^+(t) + I_i^-(t)]
\] (8)

(3) The decomposed IMF components in step (2) are separated from \( s_i \) and the residual signals are obtained:

\[
r(t) = S_i - \sum_{j=1}^{p-1} I_j(t)
\] (9)

(4) The residual signal \( r(t) \) is decomposed by EMD and all IMF components are obtained.

In MEEMD, Permutation Entropy (PE) is introduced to evaluate the complexity of time series. By setting the PE value, the PE of each component obtained in the decomposition process is judged[20]. The sequence whose PE is greater than the set value is regarded as an abnormal signal, and the sequence whose permutation entropy is larger than the set value is filtered out. According to the literature, the threshold value of PE is 0.55-0.6. That is to say, if the PE value exceeds 0.6, it can be regarded as noise.

4. Results and Analysis

In this paper, we choose a set of intact flight data from FDRS. The position data output from airborne SINS were subtracted by GPS position data recorded by FDRS, Take latitude error as an example. Record latitude error as \( \delta L \).
4.1. Remove outliers
Before data processing of flight data, outliers need to be removed first. The latitude error $\delta L$ between the SINS and GPS as shown in Figure 1.

![Figure 1. The original flight data](image)

As shown in Figure 1, there are many outliers in the original flight data, which bring great errors to the subsequent flight data interpretation. In order to ensure the accuracy of flight data, it is necessary to remove the outliers.

Using the method of Chapter 2 to process the data, the results are shown in Figure 2.

![Figure 2. Flight data after removing outliers.](image)

As shown in Figure 2, after outliers are removed, the outliers in the original flight data are effectively reduced. After eliminating outliers, We should take the second step of de-noising. Firstly, EMD is used to de-noise the original data after eliminating outliers. As shown in Figure 3.

![Figure 3. The IMF after the EMD](image)
As shown in Figure 3, after EMD decomposition, 12 IMF components are obtained. There exists mode mixing in EMD decomposition, and there is no clear criterion for eliminating noise IMF components. So on this basis, the original flight parameter data after removing outliers are de-noised by MEEMD method. As shown in Figure 4.

![Figure 4. The IMF after the MEEMD](image)

As shown in Figure 4, after EMD decomposition, 9 IMF components are obtained. The MEEMD method effectively alleviates the mode mixing problem in EMD decomposition, and the decomposition yields fewer IMF components. Compared with EMD, MEEMD has more advantages in mode mixing and de-noising. By setting the PE value threshold, MEEMD can effectively detect the abnormal IMF components obtained in decomposition. Therefore, the de-noising process of MEEMD is more adaptive than EMD.

The de-noising results of MEEMD are compared with the original flight data as shown in Figure 5.

![Figure 5. Original flight data after de-noising](image)

Comparing original flight data with MEEMD de-noising results, the results show that the MEEMD method can effectively reduce the noise contained in the data.

5. Conclusions

Influenced by the working environment of the aircraft and the factors of FDRS, the flight data which recorded by FDRS contain a large number of outliers and noises. These outliers and noises bring great interference to the subsequent flight parameter data processing, and even lead to erroneous flight state estimation. In order to eliminate these effects, flight data processing is effectively realized by using $3\sigma$ criterion and MEEMD algorithm. Firstly, the outliers of flight data can be effectively removed by $3\sigma$ criterion. Then, MEEMD algorithm is used to de-noise the original flight parameter data which has been eliminated by outliers. The results show that:
(1) Through $3\sigma$ criterion, the outliers in the original flight data can be detected and removed effectively.

(2) MEEMD method can effectively solve the mode mixing problem, and has less IMF than EMD.

(3) By setting PE, abnormal IMF detection in MEEMD decomposition can be realized, and the original flight parameter data can be de-noised effectively.

References

[1] D’Angelo G, Ferretti S, Marzolla M. A Blockchain-based Flight Data Recorder for Cloud Accountability[J]. 2018.

[2] Wiseman Y. Unlimited and protected memory for flight data recorders[J]. Aircraft Engineering and Aerospace Technology. 2016, 88(6): 866-872.

[3] Timi P H, Shermin S, Rahman A. Study of flight data recorder, underwater locator beacon, data logger and flarm collision avoidance system[C]. 2017.

[4] Aborizka M, Kavi K M. An Intelligent Flight Data Recorder And Real-Time Monitoring System[J]. Journal of Public Health Management & Practice Jphmp. 2007, 9(9): 157-164.

[5] Khadilkar H, Balakrishnan H. Estimation of aircraft taxi fuel burn using flight data recorder archives[J]. Transportation Research Part D: Transport and Environment. 2012, 17(7): 532-537.

[6] Onoz B, Oguz B. Assessment of Outliers in Statistical Data Analysis[M]. 2003.

[7] Barnett V, Lewis T, Abeles F. Outliers in Statistical Data[M]. 1978.

[8] Zhu Feifei, Gao Yanhui, Xiao Qiangui, et al. Application of observer/Kalman filter in flight data processing[J]. Electronic Design Engineering. 2016(20): 91-93.

[9] Sun Wenzhu, Qu Jianling, Yuan Tao, et al. “Flight Data Novelty Detection Method Based on Improved SVDD,” Chinese Journal of Scientific Instrument, 2014, 35(4): 932-939.

[10] Di Yazhou, Li Furong, Yu Jianli, et al. Application of Wavelet Analysis on Flight Data De-noising[J]. Computer Simulation. 2010(10): 1-4.

[11] Tang Chongkai, Qu Jianling, Gao Feng, et al. Flight data preprocessing based on wavelet neural network[J]. Journal of Data Acquisition & Processing. 2012, 27(1): 172-174.

[12] Wang Bangfeng, Lin Jianxiang, Lu Jiyun. Detection of Small Mutation Signals in Flight Data Based on EEMD-H T[J]. Journal of Vibration, Measurement & Diagnosis, 2013(03): 388-392.

[13] Peng Y, Sun H, Zu L, et al. Adaptively target tracking method based on double-Kalman filter in existence of outliers[C]. 2010.

[14] Freeman J. Outliers in Statistical Data (3rd edition)[J]. Journal of the Operational Research Society. 2017, 46(8): 1034-1035.

[15] Sun Yongshuai, Wang Shaoyun. The processing method of telemetry data based on letts criterion and wavelet transform[J]. Electronic Design Engineering. 2016, 24(20): 82-84.

[16] Huang N E, Shen Z, Long S R, et al. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis[J]. Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences. 1998, 454(1971): 903.

[17] Duan W, Huang L. A hybrid EMD-AR model for nonlinear and non-stationary wave forecasting[J]. Journal of Zhejiang University Science A. 2016, 17(2): 115-129.

[18] Ding Y. A novel decompose-ensemble methodology with AIC-ANN approach for crude oil forecasting[J]. Energy. 2018, 154: 328-336.

[19] ZHENG J D, CHEN J S, YANG Y. Modified EEMD algorithm and its applications[J]. Journal of Vibration and Shock 2013(21): 21-26.

[20] Bandt C. Permutation Entropy and Order Patterns in Long Time Series[M]. 2016.