A Generative Adversarial Nets Method for Monitoring Data Generation on Aircraft Engines

FU Qiang, *WANG Huawei**

College of Civil Aviation, Nanjing University of Aeronautics and Astronautics, Nanjing 211106, P. R. China

(Received 21 March 2018; revised 9 May 2018; accepted 1 June 2018)

Abstract: A sufficient sample size of monitoring data becomes a key factor for describing aircraft engines state. Generative adversarial nets (GAN) can be used to expand the sample size based on the existing state monitoring information. In the paper, a GAN model is introduced to design an algorithm for generating the monitoring data of aircraft engines. This feasibility of the method is illustrated by an example. The experimental results demonstrate that the probability density distribution of generated data after a large number of network training iterations is consistent with the probability density distribution of monitoring data. The proposed method also effectively demonstrates the generated monitoring data of aircraft engine are in a reasonable range. The method can effectively solve the problem of inaccurate performance degradation evaluation caused by the small amount of aero-engine condition monitoring data.

Key words: generative adversarial nets(GAN); aircraft engine; condition monitoring; monitoring data

CLC number: V231 **Document code**: A **Article ID**: 1005-1120(2019)04-0609-08

0 Introduction

The engine, as the heart of an aircraft, is a complex and highly integrated system. A major problem facing airlines and manufacturers is how to improve operation reliability of aircraft engines. Because of monitoring costs and environment restrictions, minimal monitoring data are available, leading to some uncertainties. It is difficult to meet the the requirements of accurate assessment for health management[1].

At present, condition monitoring is the key component of aircraft engine health management. Wang et al. conducted a reliability evaluation method to analyze competing failures for aircraft engines[2]. Sun et al. predicted future engine health conditions by using bias state estimation and prediction formulas in a comprehensive evaluation[3]. Jaklinski provided an analysis of the aircraft engine dual-control system during failure conditions. The design and control algorithms are insusceptible to a single sensor failure[4]. Yu et al. described the decision method with uncertainty interval information for the condition monitoring of aircraft engines[5]. Liu et al. discussed the correlation between multiple degenerate quantities through the copula function family and fused the edge distributions to obtain the joint distribution function of the remaining life. A lifetime prediction method based on the copula function with multiple degenerate quantities was proposed[6]. However, these methods are based on small sample data. Data regeneration technology is an effective way to solve the problem of insufficient data. Zhang et al. provided synthetic data generation for end-to-end thermal infrared tracking[7]. Wang et al. established the white light emitting diode(WLED) application of a simulation data generation platform of light sources' colour characteristics[8]. Huang et al. provided a multi-pseudo regularised label for generated data in person re-identification[9]. In this paper, for the limitation of monitoring data, a data regeneration technology based on the generative adversarial nets(GAN) model is proposed.

GAN, as one of the frontier research direc-

*Corresponding author, E-mail address: wang_hw66@163.com.

How to cite this article: FU Qiang, WANG Huawei. A Generative Adversarial Nets Method for Monitoring Data Generation on Aircraft Engines[J]. Transactions of Nanjing University of Aeronautics and Astronautics, 2019,36(4):609-616.
http://dx.doi.org/10.16356/j.1005-1120.2019.04.007

tions in the field of artificial intelligence, was proposed by Goodfellow et al.[10] in 2014. GAN is widely used in image and visual field research, which can generate digital faces, and other objects. GAN can form various realistic indoor or outdoor scenes, restore original images from image segmentation and add colour to black-white images[11-12]. GAN can generate object images from object contours and produce high-resolution images from low-resolution images[13]. Additionally, GAN has been applied to processing language speech[14-15], monitoring computer viruses[16], studying chess game programs[17], and other issues. For intelligent fault identification of mechanical failure modes, GAN has extracted the characteristics of mechanical fault vibration signals, expanded fault samples, and improved the accuracy of fault identification[18]. Using the GAN method has realized traffic data generation, traffic modelling, traffic prediction, and traffic control in parallel transportation systems, and it has provided specific algorithms for implementing parallel-transportation systems[19].

The GAN principle is equally applicable to aero-engine condition monitoring. The previously mentioned characteristics of the GAN method are in line with the characteristics of aircraft engine condition monitoring. GAN is an effective way to generate enough data for deep analysis and learning. When the condition monitoring data of aircraft engines are too minimal to accurately predict performance degradation, enough monitoring data can be produced by using the GAN model. In this paper, the GAN model is designed for data regeneration, aiming at the health monitoring data of aircraft engines. The design algorithm is used to achieve data regeneration, and the accuracy of the data is verified.

1 Condition Monitoring for Aircraft Engines

1.1 Condition monitoring parameters

Engine performance degradation usually reflects monitoring parameters. Condition monitoring for these parameters is crucial. The gas path system is the core of the aircraft engine and includes a pressure machine, a chamber, and a turbine. Some thermodynamic parameters of the gas path system can effectively reflect changes in engine state performance. Lubrication oil monitoring data are used for lubrication system components and the condition of sealing systems, thereby playing an important role in machine wear monitoring and fault diagnosis. Vibration monitoring is observed through the rotation of the engine rotor including vibration signals from fault. Mechanical wear is assessed by vibration monitoring[20-22]. Fig.1 describes the condition monitoring of aircraft engines.

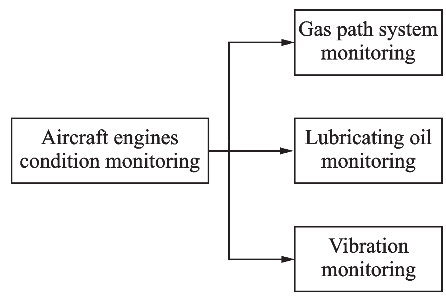

Fig.1 Main content of aircraft engine performance monitoring

1.2 Characteristics of performance parameters

Performance degradation can be comprehensively described by the exhaust gas temperature margin (EGTM), oil pressure deviation, high-pressure rotor speed deviation, fuel consumption deviation, low-pressure rotor vibration deviation, and high-pressure rotor deviation. Some correlations exist among these performance parameters. For instance, some parameters, such as EGTM[23], are more sensitive than others. These parameters are monitored to analyse the cumulative degradation of engines. Although aircraft engines involve multiple performance monitoring parameters, EGTM is one of the main reference parameters of engines in the actual operation process of airways. When EGTM declines to the threshold, the airline company replaces the engine. The other performance monitoring parameters are used as auxiliary references for engine health conditions.

1.3 Analysis of the sample size of performance parameters data

The health monitoring of aircraft engine is based on the precise monitoring model and enough reliable data. An aircraft engine degradation pattern includes normal performance degradation and fault-caused degradation. The performance is not the same degradation model of different engines. The rate of declining varies with the operating environment and time on wing. Therefore, it is difficult to precisely define the degradation model of an aircraft engine.

However, the condition monitoring data is from different collection environments. The accuracy of the collected data cannot be guaranteed, and the distributions of different condition monitoring parameters of aircraft engines are also different. Therefore, many condition monitoring data samples are needed to reduce errors. In the present work, the GAN model is used to generate a large amount of reliable condition monitoring data. This method can effectively address the insufficient sample size of the condition monitoring data.

2 GAN Model of Generating Monitoring Data for Aircraft Engines

2.1 GAN model for monitoring data generation

Generating condition monitoring data for aircraft engines requires the accuracy of the model. First, the data generated by GAN need to meet the sample size of the subsequent mining state feature. Second, the generated data should be accurate and capable of smoothing the noise in different acquisition environments. In data generation modelling, GAN can be constructed by two network models: generative (generator G) and discriminant (discriminator D) network models. The framework sizes of networks G and D in GAN are similar and based on multilayer perceptron neurons.

Fig.2 depicts the network structure of GAN. In general, G and D can be developed by a neural network algorithm. After repeating adversarial network training, GAN can generate condition monitoring data, the distribution of which is similar to the distribution of aircraft engine monitoring data. The G model learns an approximate distribution of monitoring data, and the monitoring data are distinguished from the data generated using the D model.

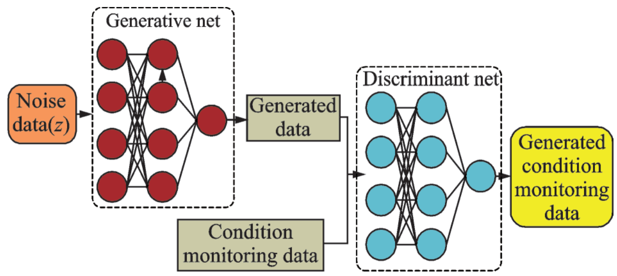

Fig.2 Network structure of GAN

In GAN, random noise data (z) are added to G, and big data are generated. The model G is a map of noise data to generated data. The output $G(z)$ is similar to the monitoring data distribution. The inputs of D are the generated and original condition monitoring data. The D model outputs probability values to determine the quality of the generated data.

2.2 Design of generated data

In the GAN model, G generates a flow of data samples, and D estimates the accuracy of the generated data from G. Both G and D are part of the adversarial process. In this process, G and D are constantly learning, and GAN ensures that the learning rates are consistent. If one of the learning rates of G or D is faster or slower than the others, the training process is no longer balanced. The loss function of the slower learning rate model cannot decline. The generated performance monitoring data are inaccurate. To take advantage of GAN for generating engine monitoring data, the feasibility of the method using the example is verified. Then, GAN generates a large volume of condition monitoring data, which are used for data analysis. The design idea is presented in Fig.3.

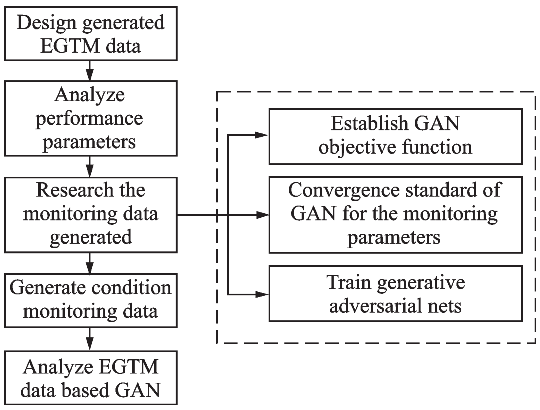

Fig.3 Monitoring data generation flow diagram of aircraft engines

3 GAN Model Algorithm of Generating Monitoring Data for Aircraft Engines

3.1 Establishment of GAN objective function

In this section, GAN is constructed by a neural network. The network training process of GAN is as follows: x represents the condition monitoring data, and z the noise data (generally Gauss or uniform noise data), which are the inputs of G. After the training of G, the output $G(z)$ is similar to the condition monitoring data. Then, x and $G(z)$ become the inputs of D. D provides the input data with the probability of the condition monitoring data. Furthermore, G outputs the final generated data samples.

To describe the objective function of GAN, The number of monitoring data is m. A sample training set is $S=\{x_1,\cdots,x_m\}$. In addition, the probability density function is $P_z(z)$. Using a random variable $z\sim P_z(z)$ can generate m noise samples $\{z_1,\cdots,z_m\}$. In conclusion, Eq.(1) presents the likelihood function, in which θ_G and θ_D are the network model parameters of G and D, respectively.

$$L(x^{(1)},\cdots,x^{(m)},\cdots,z^{(m)}|\theta_G,\theta_D)=\prod_{i=1}^{m}D(x^{(i)})\prod_{j=1}^{m}(1-D(G(x^{(j)}))) \tag{1}$$

In accordance with the law of large numbers, when $m\rightarrow\infty$, the experience loss approximates the expected loss. Eq.(2) presents the log likelihood function obtained by Eq.(1). In Eq.(2), $P_{\mathrm{data}}(x)$ denotes the probability density function of the condition monitoring data

$$\log L\approx E_{x\sim P_{\mathrm{data}}(x)}[\log D(x)]+E_{z\sim P_z(z)}[\log(1-D(G(z)))] \tag{2}$$

The adversarial modelling framework of the generated condition monitoring data is directly applied when the G and D models are multilayer perceptrons. Accordingly, G learns the distribution P_g of the generated condition monitoring data. D maximizes the probability of the generated condition monitoring data from G and the original condition monitoring data. As a result, G is trained to minimise $\log(1-D(G(z)))$. In other words, G and D play a two-player mini-max game with the value function $V(G,D)$, which is presented in Eq.(3)

$$\min_G\max_D V(G,D)=E_{x\sim P_{\mathrm{data}}(x)}[\log D(x)]+E_z\sim P_z(z)[\log(1-D(G(z)))] \tag{3}$$

3.2 Convergence standard of GAN for the monitoring parameters

The generator G implicitly defines a probability distribution, P_g, because G is a map from the noise $z\sim P_z(z)$ to the sample $G(z)$. The goal is to maximise $V(G,D)$, as demonstrated in Eq.(4). Both χ and Ω are the integral spaces of the condition monitoring and noise data, respectively.

$$V(G,D)=\int_\chi P_{\mathrm{data}}(x)\log(D(x))\mathrm{d}x+\int_\Omega P_z(z)\log(1-D(G(z)))\mathrm{d}z \tag{4}$$

In accordance with the second term of Eq.(4), Eq.(5) can be obtained by using the mapping relation of $x=G(z)$

$$\int_\Omega P_z(z)\log(1-D(G(z)))\mathrm{d}z=\int_\chi p_g(x)\log(1-D(x))\mathrm{d}x \tag{5}$$

Then, Eq.(6) is obtained through Eqs.(4), (5).

$$V(G,D)=\int_\chi[P_{\mathrm{data}}(x)\log(D(x))+p_g(x)\log(1-D(x))]\mathrm{d}x \tag{6}$$

The values P_{data} and P_g are nonzero functions because the data distribution has been determined. The final goal is to identify a function that causes $V(G,D)$ to arrive at the maximum. Therefore, Eq.(7) can be obtained. The maximum point $D_G^*(x)$ can be expressed by Eq.(7)

$$D_G^*(x)=\frac{P_{\text{data}}(x)}{P_{\text{data}}(x)+P_{\text{g}}(x)} \tag{7}$$

The process of GAN can be further formalised as a minimisation of $C(G)$ by Eq.(8)

$$C(G)=\max_D V(G,D)= E_{x\sim P_{\text{data}}}\left[\log\frac{P_{\text{data}}(x)}{P_{\text{data}}(x)+P_{\text{g}}(x)}\right]+ E_{x\sim P_{\text{g}}}\left[\log\frac{P_{\text{g}}(x)}{P_{\text{g}}(x)+P_{\text{g}}(x)}\right] \tag{8}$$

For the stability of training, the training target of GAN is different from that of the traditional neural network algorithm. In the network training process, P_{g} remains close to P_{data}, which is the convergence criterion of GAN for generating monitoring data. The necessary and sufficient condition of the global minimum in GAN is calculated in Eq.(9)

$$P_{\text{g}}(x)=P_{\text{data}}(x) \tag{9}$$

3.3 Training GAN for aircraft engine condition monitoring data

The net training process requires training G first and then D to ensure a stable training process. In the process of training GAN, G and D are dynamic. The ideal condition expressed by Eq.(3) to derive the optimal goal and train GAN is to optimise D to some steps first, followed by G. In this method, if G changes slowly enough, D can always be in the vicinity of the optimal solution. Both G and D have enough learning abilities and can achieve the optimal point to generate optimal condition monitoring data. In accordance with Eq.(10), the optimal target P_{g} can converge to P_{data}

$$E_{x\sim P_{\text{data}}}[\log D_G^*(x)]+E_{x\sim P_{\text{g}}}[\log D_G^*(x)] \tag{10}$$

4 Examples

For the performance degradation of aircraft engines, EGTM is the sensitive condition monitoring parameter. Additionally, EGTM, as a critical parameter, is commonly used in engineering practice. Therefore, EGTM is selected to generate data by using GAN. If GAN can generate valid EGTM data, the other parameters are equally valid. This example provides a type of engine for an airline and its 13 200 cycles of EGTM data[24]. The EGTM failure threshold is 30, and Fig.4 depicts the performance data.

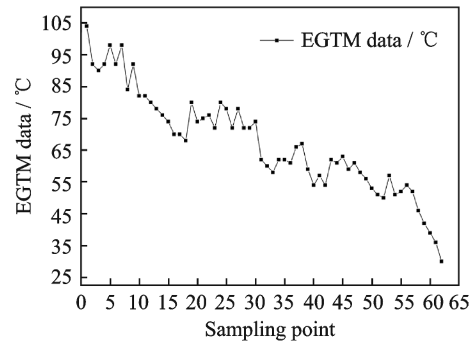

Fig.4 EGTM data

The sample data of EGTM follows a normal distribution. The generated data are estimated by the 10 000 cycles of EGTM data, and the generated data also follow a normal distribution. The probability density curve of the EGTM sample data generated by GAN is presented in Fig.5.

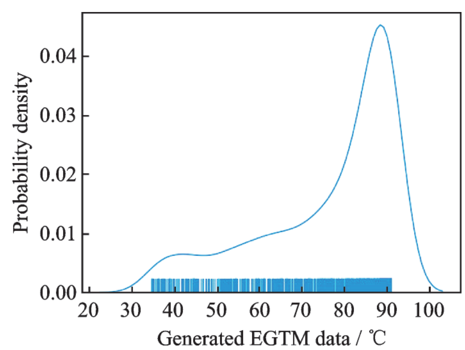

Fig.5 Generated EGTM data

To facilitate the analysis of the EGTM probability density function, the whole training process is illustrated in Fig.6.

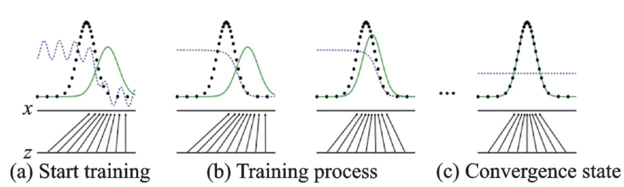

Fig.6 Network training process of GAN

In Fig. 6, the lower horizontal line z denotes the domain of noisy data, and x refers to the domain of sample data. The green curve shows P_{g}, which is the probability density function of the generated EGTM data. The black curve represents P_{data}, which is the probability density function of the original EGTM data. The blue curve is the probability function of the discriminant network.

In the training process, the discriminant network (blue line) can be trained and updated. Then, D can distinguish between the data distribution of EGTM data (black lines created by black points) and P_g (green line). In the final iteration, the discriminant network converges to Eq.(9). Then, the discriminant network is fixed, and the generative network is trained. After updating G, the gradient of D guides $G(z)$ to the direction that is likely to be classified as EGTM data. An equilibrium point is obtained after enough training times. At this point, P_g is equal to P_{data}, and the generated data can be considered similar to the original EGTM data. At the equilibrium point, D and G cannot be improved further.

The generated EGTM sample data is displayed in Fig.7. The volume of the generated EGTM data is equal to 1 000. The average is 70, and the standard deviation is 6. All of the generated EGTM data exceed the threshold value. Thus, no failure data are generated. In the analysis of the generated EGTM data and the consideration of the actual situation, abnormal data should be eliminated. However, in the results, all the generated EGTM data meet the Pauta criterion.

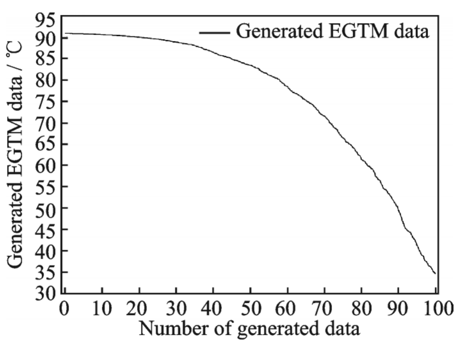

Fig.7 Generated EGTM sample data

The generated EGTM data are based on the first 10 000 cycles of the EGTM data. Drawing a comparison between the generated EGTM data and the 10 200—13 200 cycles of EGTM data is possible. Thus, GAN can predict the moves of EGTM. The back propagation (BP) method is utilized to predict the performance degradation trend of the small sample data of EGTM and the relative ratio between the data generated using GAN and the BP method. Fig.8 depicts the 10 200—13 200 cycles of EGTM data.

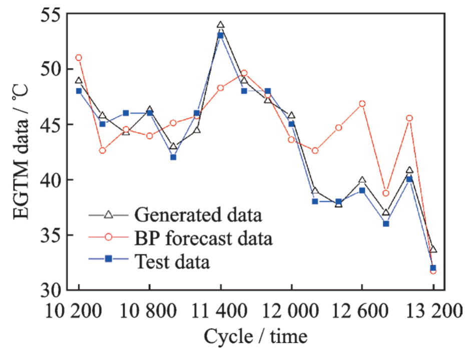

Fig.8 Generated EGTM sample data prediction

The generated data predict the trend of the changes in the small sample data of EGTM. In comparison with the traditional BP neural network, the GAN method predicts a more accurate degradation trend of EGTM. In Table 1, the range of the deviation percentage is 0.7%—5.0%. In conclusion, the prediction data are reasonable. Hence, the quality of the generated EGTM data is reliable.

Table 1 Devotion of generated EGTM data and original EGTM data

Cycle / time	Generated EGTM data / ℃	EGTM data / ℃	Deviation	Deviation percentage / %
10 200	48.8	48	0.8	1.7
10 400	45.7	45	0.7	1.6
10 600	44.2	46	1.8	3.9
10 800	46.3	46	0.3	0.7
11 000	42.9	42	0.9	2.1
11 200	44.4	46	1.6	3.5
11 400	53.9	53	0.9	1.7
11 600	48.6	48	0.6	1.3
11 800	47.1	48	0.9	1.9
12 000	45.7	45	0.7	1.6
12 200	38.9	38	0.9	2.4
12 400	37.7	38	0.3	0.8
12 600	39.9	39	0.9	2.3
12 800	36.9	36	0.9	2.5
13 000	40.8	40	0.8	2.0
13 200	33.6	32	1.6	5.0

5 Conclusions

This study presents a method for generating condition monitoring data for aircraft engines based on the theory of GAN. As a typical aircraft engine condition monitoring parameter, EGTM is used to verify the effectiveness of the proposed method. The example demonstrates that the probability density distributions of the EGTM and generated EGTM are consistent. All the generated EGTM data meet the Pauta criterion. For EGTM prediction, the generated prediction data are in a reasonable range, and the data generated using the GAN method reflect the characteristics of small sample monitoring data. This method can determine the inherent variations of small sample data. Further, GAN is an effective tool for solving the problem of insufficient samples, and it can provide enough and precise data to support performance degradation.

References

[1] LI T L, HE L M, ZOU S J, et al. Condition monitoring for aero-engine based on chaos exponents of dynamic system[J]. Journal of Aerospace Power, 2008, 23(11): 2133-2136.

[2] WANG H W, GAO J. A reliability evaluation study based on competing failures for aircraft engines [J]. Eksploatacja i Niezawodnosc-Maintenance and Reliability, 2014, 16 (2): 171-178.

[3] SUN J, ZUO H, WANG W, et al. Application of a state space modeling technique to system prognostics based on a health index for condition-based maintenance[J]. Mechanical Systems & Signal Processing, 2012, 28(2): 585-596.

[4] JAKLINSKI P. Analysis of the dual control system operation during failure conditions[J]. Eksploatacja i Niezawodnosc-Maintenance and Reliability, 2013, 15 (3): 266-272.

[5] YU H T, WANG H W, LI Q. A study of health evaluation for aviation engine [J]. Systems Engineering and Electronics, 2011, 30(6): 996-1000.

[6] LIU S N, LU N Y, CHENG Y H, et al. Remaining lifetime prediction for momentum wheel based on multiple degradation parameters [J]. Journal of Nanjing University of Aeronautics & Astronautics, 2015, 47 (3): 360-366.

[7] ZHANG L C, GONZALEZ-GARCIA A, WEIJER J, et al. Synthetic data generation for end-to-end thermal infrared tracking[J]. IEEE Transactions on Image Processing, 2019, 28(4): 1837-1850.

[8] WANG Z Y, ZHENG R L , YU K H, et al. The establishment and YAG:Ce-based WLED application of simulation data generation platform of light sources' color characteristics[J]. Optics Communications, 2019, 434(1): 230-238.

[9] HUANG Y, XU J S, WU Q, et al. Multi-pseudo regularized label for generated data in person re-identification[J]. IEEE Transactions on Image Processing, 2019, 28(3): 1391-1403.

[10] GOODFELLOW I, POUGETABADIE J, MIRZA M. Generative adversarial networks[EB/OL]. (2014-06-10)[2018-03-15]. http://arxiv.org/abs/1406.2661.

[11] REINHARD E, ADHIKHMIN M, GOOCH B, et al. Color transfer between images[J]. IEEE Computer Graphics and Applications, 2001, 21(5): 34-41.

[12] LARSSON G, MAIRE M, SHAKHNAROVICH G. Learning representations for automatic colorization [C]//European Conference on Computer Vision. Las Vegas: Springer, 2016: 577-593.

[13] LEDIG C, THEIS L, HUSZAR F, et al. Photo-realistic single image super-resolution using a generative adversarial network[EB/OL]. (2017-05-25)[2018-03-15]. https://arxiv.org/abs/1609.04802.

[14] LI J W, MONROE W, SHI T, et al. Adversarial learning for neural dialogue generation [EB/OL]. (2017-09-24)[2018-03-15]. https://arxiv.org/abs/1701.06547.

[15] YU L T, ZHANG W N, WANG Jet al. Seq GAN: Sequence generative adversarial nets with policy gradient[EB/OL]. (2017-08-25)[2018-03-15]. https://arxiv.org/abs/1609.05473.

[16] HU W W, TAN Y. Generating adversarial malware examples for black-box attacks based on GAN[EB/OL]. (2017-02-20)[2018-03-15]. https://arxiv.org/abs/1702.05983.

[17] CHIDAMBARAM M, QI Y J. Style transfer generative adversarial networks: learning to play chess differently[EB/OL]. (2017-02-22)[2018-03-15]. https://arxiv.org/abs/1702.06762.

[18] WANG Z R, WANG J, WANG Y R. An intelligent diagnosis scheme based on generative adversarial learning deep neural networks and its application to planetary gearbox fault pattern recognition[J]. Neurocomputing, 2018, 310: 213-222.

[19] LÜ Y S, CHEN Y Y, LI L, et al. Generative adversarial networks for parallel transportation systems[J]. IEEE Intelligent Transportation Systems Magazine, 2018, 10(3): 4-10.

[20] WANG H W, WU H Q. Residual useful life prediction for aircraft engine based on information fusion[J]. Journal of Aerospace Power, 2012, 27(12): 2749-2755.

[21] REN S H, ZUO H F. Real-time performance reliability prediction for civil aviation aircraft engines based on multiple performance measures[J]. Journal of Aerospace Power, 2010, 25(12): 2811-2815.

[22] WANG H W, GAO J, WU H Q. Reliability analysis on aero-engine using Bayesian model averaging[J]. Journal of Aerospace Power, 2014, 29(2): 305-313.

[23] CAI J, LIU X, ZHU B L. Study on remaining useful life prediction for aero-engines combining sate space model and KF algorithm[J]. Transactions of Nanjing University of Aeronautics & Astronautics, 2017, 34(3): 265-271.

[24] REN S H. Research on methods of performance reliability assessments and life on wing prediction for civil aeroengine[D]. Nanjing: Nanjing University of Aeronautics and Astronautics, 2010.

Acknowledgements The work was supported by the National Science Foundation for Young Scientists of China (No. 71401073). The authors would like to express the gratitude to all those who helped us during this study: Sun Zhongdong, Ni Xiaomei, Che Changchang, and all with the Laboratory of Civil Aviation Safety, Nanjing University of Aeronautics and Astronautics.

Authors Mr. **FU Qiang** received his B.S. degree in mechanical engineering from Shengyang University in 2015. He is a Ph.D. candidate majoring vehicle operation engineering at Nanjing University of Aeronautics and Astronautics, Nanjing, China. His research is focused on aeroengine condition monitoring and relevant fields.

Prof. **WANG Huawei** graduated from National University of Defense Technology in 2003 with a doctoral degree in management science and engineering. In 2006, she completed the post-doctorate work at Nanjing University of Aeronautics and Astronautics. Since 2006, she has been working at College of Civil Aviation, Nanjing University of Aeronautics and Astronautics. She is an outstanding young-backbone teacher in Jiangsu Blue project.

Author contributions Mr. **FU Qiang** designed the study, complied the models, conducted the analysis, interpreted the results and wrote the manuscript. Prof. **WANG Huawei** performed the model analysis with constructive discussions. Both authors commented on the manuscript draft and approved the submission.

Competing interests The authors declare no competing interests.

(Production Editor: Zhang Bei)