Data-Driven Prognostics and Health Management Solutions in the Case of Multiple Degradations in Complex Systems

DATA-DRIVEN PROGNOSTICS AND HEALTH MANAGEMENT SOLUTIONS IN THE CASE OF MULTIPLE DEGRADATIONS IN COMPLEX SYSTEMS

Huxiao Shi[1], Shuo Yang[1], Jie Geng[2], Gabriele Baldissone[1] and Micaela Demichela[1]

[1] *Safety, Reliability and Risk centre (SAfeR), Department of Applied Science and Technology, Politecnico di Torino, Turin Italy*

[2] *China Research Institute of Regulation and Public Policy, China Institute of Regulation Research, Zhejiang University of Finance & Economics, Hangzhou China*

ABSTRACT

Reliability engineering benefits asset management from many aspects. Currently, the consideration of reliability becomes more critical but difficult since the interaction among physical, digital, and human dimensions becomes more complex. Prognostics and Health Management is regarded having more benefits at both component and system levels, and is able to provide dynamic and predictive maintenance strategies. In this study, reliability analysis tools are embedded in a data-driven PHM framework in the case that the turbofan engine is considered as a complex system. Multiple degradation modes are analyzed by exploring systems' fault conditions. The results show the benefit of combining reliability analysis tools with data-driven PHM approaches, as well as the advantages in the decision-making process.

Keywords: Asset Management; Dynamic Monitoring; Prognostics and Health Management; Multiple Degradation; Complex System.

1. INTRODUCTION

Reliability engineering is considered to be highly related to asset management (AM). As pointed out by Katina et al. (2021), concepts in reliability engineering and also risk management benefit making decisions on assets, and thus maximize asset values. Though the modern AM scheme is instilled with more meanings from the organization, finance, and other perspectives, assets' lives and health care based on them are still crucial and significant in industries. However, with the development of Industry 4.0, the interaction among physical, digital, and human dimensions becomes more complex and cannot be analyzed separately. To better conduct AM with the consideration of reliability, availability, maintainability, and safety (RAMS) of structures, systems and components (SSCs), Prognostics and Health Management (PHM) is burgeoning and being considered to have more benefits.

Being different from traditional passive maintenance approaches like reactive maintenance and preventive maintenance, PHM aims to provide dynamic or predictive maintenance strategies based on its capability of predicting Remaining Useful Life (RUL) of SSCs (Demichela et al. 2018; Compare et al. 2022). Currently, there are three types of PHM approaches have been developed, which are based on physical knowledge, data-driven, and even hybrid model with other technologies. Specifically, PHM could be based on physical approaches through using explicit mathematical equations to predict degradation/failure behaviors (Pecht and Jaai 2010). In some cases where precise physical models are not available, data-driven approaches are prone to be applied with the support of artificial intelligence techniques (Meng et al. 2019; Omri et al. 2021).

In this study, the data-driven PHM is adopted in the analysis of a complex system with multiple degradation modes with the support of monitoring activities. And reliability analysis tools are embedded in the proposed data-driven PHM framework to understand systems' fault condition. And thus facilitates designing the failure threshold in the prognostics module. How does results from the prognostics module benefit decision-making processes are discussed in the prognostics-supported decision-making module.

2. METHODOLOGY

The methodology in this study aligns with the proposed PHM framework in previous study, which is regarded as one data-driven solution with the advantages of combining the potential of records or measurements from industrial monitoring activities. Four parts are designed contributing to conduct data-driven PHM analysis on system's degradations, which are:

- Health Index (HI): Define the HI of target systems based on the monitored data. Monitored data usually are records from deployed sensors, operation parameters, or even external parameters.
- Failure Threshold Design: Design systems' failure threshold. When HI reaches the failure threshold, systems turn to fail and no longer work. The failure threshold could be either static or dynamic based on specific demands.
- Stochastic Process-based Degradation Modeling: Model the HI through stochastic process modeling. Stochastic process modeling is considered one prominent technique in degradation-related scenarios, with the advantage of capturing the stochastic dynamics that frequently occur in degradation processes.
- Prognostics-supported Decision-making Module: The results from the above parts in the framework support the decision-making process. To be specific, the prediction of Remaining Useful Life (RUL) helps to understand and predict the health status of systems, and the availability of conditional data (CM) facilitates suggestions of whether the maintenance action needed to be taken in real scenarios.

This study combines the advantages of this framework with reliability analysis tools to expand its applications (Figure 1). Specifically, original framework faces the challenge of analyzing multiple degradation modes, which hinders the design of failure thresholds. In this study, reliability analysis tools are applied to support understanding systems' fault conditions and designing fault mechanisms, and hence helps to design the failure threshold. Failure Mode and Effects Analysis (FMEA), Fault Tree Analysis (FTA), Reliability Block Diagrams (RBD), and other tools are general reliability analysis tools, the selection of them depends on the practical demands. By embedding reliability analysis tools into the data-driven framework, a more reliable and explainable model is expected to be modelled with the help of expert knowledge.

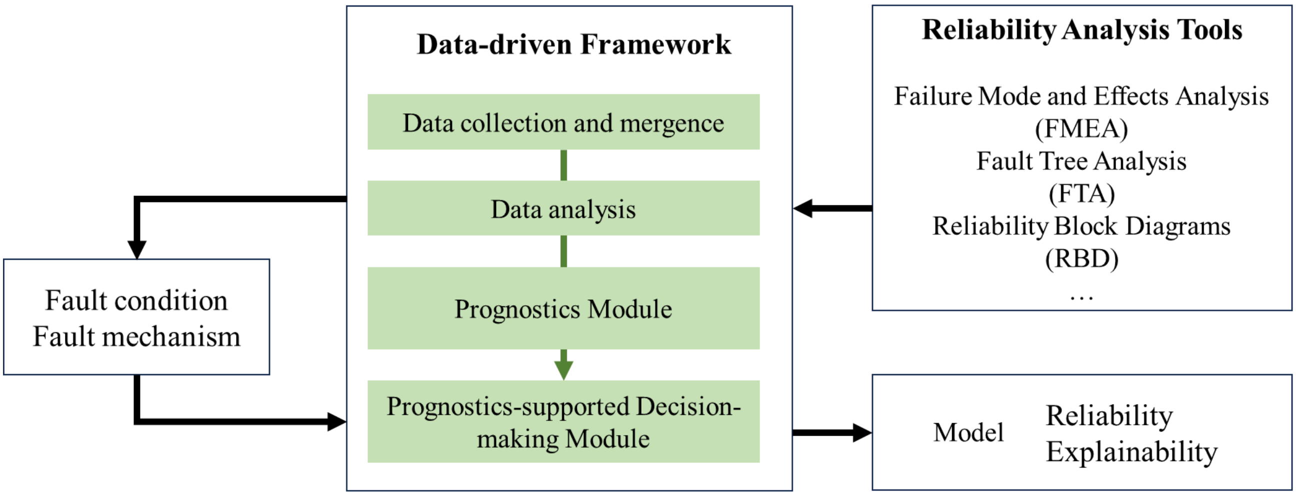

Figure 1. The framework of this study

3. CASE STUDY

3.1. Introduction of applied case study

In this study, turbofan engine datasets (CMAPSS) are adopted as the dataset to be analyzed. To be more specific, this dataset records the run-to-failure performance of one hundred engines (units) in the same batch through the deployment of 21 sensors and three preset operational settings (Frederick et al. 2007). One of the main challenges when analyzing CMAPSS is that multiple degradations are faced in this complex system. Because of the high interaction degree among each part in aircraft engine systems, it is challenging to not only conduct diagnosis through measurements from multiple sensors but also identify an appropriate model to reflect the health status of target systems. The dataset FD003 is chosen in this study, with High-Pressure Turbine (HPT) degradation and Fan degradation as the failure mode. FD003 involves both the training set and test set attached with their RUL, which helps to design the prognostics model.

3.2. Health Index (HI)

A basic preprocessing process is needed in this study, where records from sensors 1, 5, 6, 10, 16, 18, and 19 as well as operation settings 2 and 3 are removed since their constant values, which are considered to have less contribution to design HI. Z-Transformation (Z-score normalization) is also adopted to convert all remaining variables at the same scale, which supports the following analysis.

The application of PCA is expected to extract the long-term trend of system performance from multiple variables, which is also significant variance directions from a mathematical point of view. Table 1 shows the results of PCA, there are a total of 15 PCs extracted from the analysis. According to the defined selection criterion, the Principal Component (PC) whose standard deviation is larger or equal to 1 is selected which are PC1, PC2, and PC3, and those three PCs are able to explain 80.6% variance of the original data. The trend performance of these three PCs helps to find out the PC related to degradation (DePC). Figure 2 shows the plot results of PC1 and PC2 for all 100 units, where PC1 shows increasing trends as the time cycle goes by. The trend of PC2 is relatively complex, where PC2 keeps decreasing for some units and stays relatively constant for other units. The trend performance of PC3 is not shown here since it does not show either an increasing or decreasing trend as the time cycle goes by. Hence, the system's HI is regarded being highly related to two defined DePCs.

Table 1: Results of PCA

Principal component	Standard deviation	Explained variance %	Cumulative variance %
1	2.999	0.6	0.6
2	1.448	0.14	0.739
3	1	0.067	0.806
4	0.643	0.028	0.834
5	0.607	0.025	0.858
6	0.59	0.023	0.881
7	0.553	0.02	0.902
8	0.538	0.019	0.921
9	0.505	0.017	0.938
10	0.454	0.014	0.952
11	0.443	0.013	0.965
12	0.42	0.012	0.977
13	0.415	0.012	0.988
14	0.382	0.01	0.998
15	0.171	0.002	1

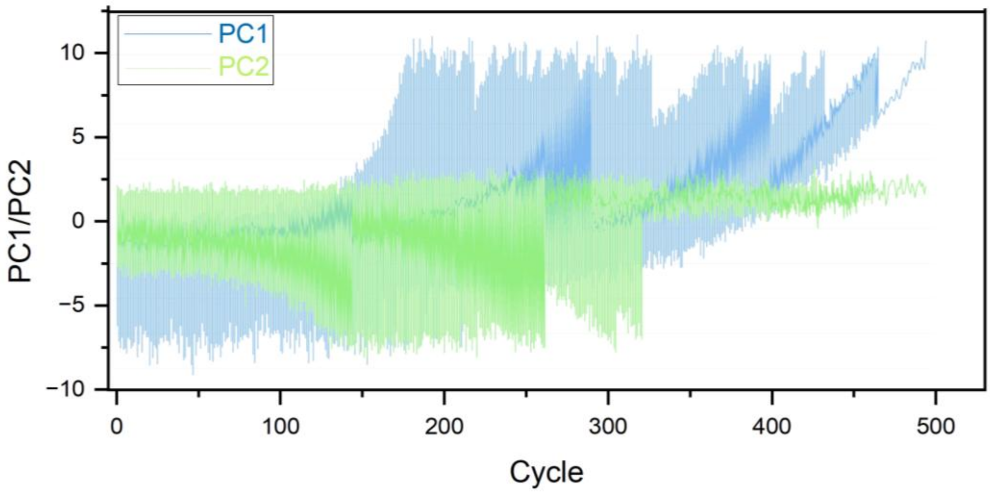

Figure 2. The trend performance of PC1 and PC2

The loading of each variable on defined DePCs are shown in the Table 2. Each variable may similarly or differently contribute to DePC1 and DePC2. Generally, four types of loading are defined as:

- Variables with similar loading magnitude and same loading direction (operational setting 1, sensor 7, 12).
- Variables with similar loading magnitude but opposite loading direction (sensor 2, 3, 4, 11, 17).
- Variables with different loading magnitude but same loading direction (sensor 9, 14, 15, 20, 21).
- Variables with different loading magnitude and different loading direction (sensor 8, 13).

Table 2: The loading of each variable

Attribute	DePC1 (PC1)	DePC2 (PC2)
operational_setting_1	-0.002	-0.002
sensor_2	0.228	-0.286
sensor_3	0.272	-0.215
sensor_4	0.258	-0.289
sensor_7	0.295	0.288
sensor_8	0.332	-0.05
sensor_9	0.312	0.002
sensor_11	0.283	-0.262
sensor_12	0.3	0.28
sensor_13	0.332	-0.049
sensor_14	0.294	0.054
sensor_15	-0.154	-0.408
sensor_17	0.28	-0.225
sensor_20	0.146	0.408
sensor_21	0.148	0.408

The loading of variables facilitates conducting detailed Root Cause Analysis (RCA) on failure modes of systems. It is expected to support component-level analysis with the potential of combining expert knowledge by knowing the physical meaning of each variable.

3.3. Failure threshold design

Concepts in Fault Tree Analysis (FTA) as a reliability analysis tool are adopted to help define the fault conditions of the target system. By checking the values of DePC1 and DePC2 at the fault time and their relative failure modes, it is found that either DePC1 or DePC2's value reaches a certain level result in systems' fault. In this case, the system fault which is regarded to be related with DePC1 and DePC2 is designed as Figure 3, where failure 1, 2 are respectively linked with DePC1 and DePC2.

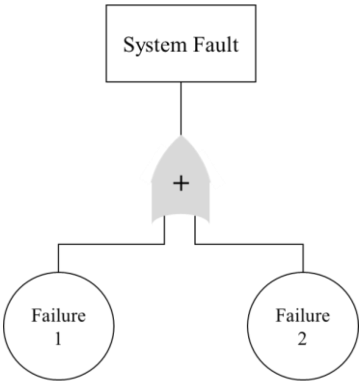

Figure 3. The fault tree defined by DePC1 and DePC2

Two failure modes explored according to two DePCs correspond to the systems' two degradation modes. Notably, no matter which degradation mode causes system fault, the value of DePC1 keeps increasing. As for the DePC2, it may either stay relatively constant or decrease for different units, only when the DePC2 shows decreasing trends may cause system fault with Failure 2 mode. Figure 4 conducts the cluster analysis on the last DePC1/2 value of each unit, two clusters extracted related to failure 1 and failure 2 help to design the failure threshold. According to the results in Table 3, Gaussian distributions can be fitted as two fault conditions, each one of the fault conditions has been met is considered causing system faults.

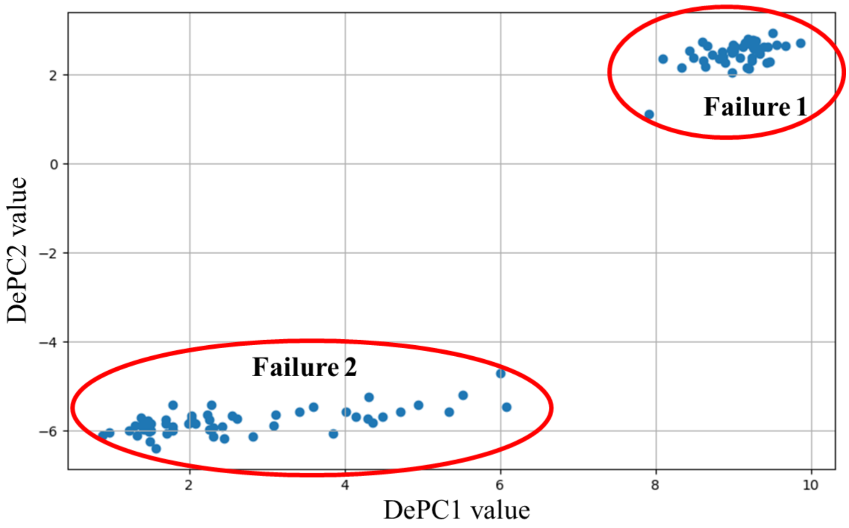

Figure 4. Cluster analysis for system fault conditions

Table 3: Statistics on fault conditions

Failure mode	Min/Max	Mean	S.D.	P value from Anderson Darling Test
1 (DePC1)	7.9114/9.8581	9.042	0.4036	0.110
2 (DePC2)	-4.713/-6.395	-5.799	0.2871	0.105

3.4. Stochastic process-based degradation modeling and simulation

The second order of polynomial regression is adopted to model two degradation modes, and Maximum Likelihood Estimation (MAE) is applied to form the parameter repository. Figure 5 shows a modeling example for DePC1 of Sample 1.

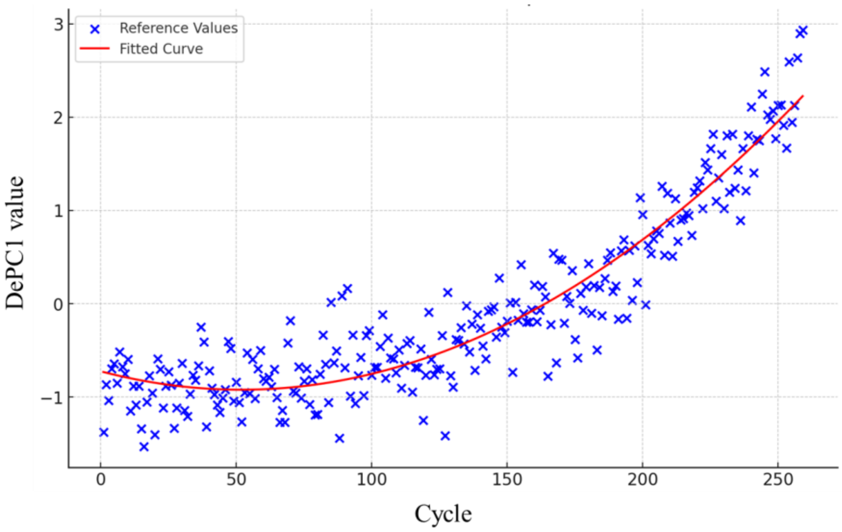

Figure 5. Modeling example for DePC1 of Sample 1

In order to design the RUL predictor, Monte Carlo (MC) simulations have been conducted based on two calibrated degradation models (Equations 1), where t denotes cycles and $B(t)$ denotes the standard Brownian Motion.

$$DePC1(t) = -0.897 + 0.0002t^2 - 0.02315t + 0.562066B(t) \qquad (t \geq 0) \tag{1}$$

$$DePC2(t) = -0.68 - 0.00023t^2 + 0.0183t + 0.469662B(t) \qquad (t \geq 0)$$

The proposed model and trained predictor are validated on the test set provided by FD003. Based on the CM at previous cycles for new systems, MC simulations keep on simulating systems' degradation processes, once DePC1(t) or DePC2(t) reaches the pre-defined failure threshold, the system is regarded to be reached the fault conditions, where its time cycle is taken as system's lifetime, and the difference between lifetime and the cycle of the last CM is considered as RUL. Figure 6 shows an example of the prediction result from MC simulations, 100 times MC simulations have been conducted for one sample, each simulation results in one potential degradation process as well as one potential RUL, the expect value of all 100 potential RULs is regarded as the final RUL, which is 53.94 shown in the figure while the actual RUL is 51 cycles. The performance of the trained predictor is summarized in the Table 4 (The simulation is conducted on MATLAB R2023a with the computer parameters provided as: Intel (R) Core (TM) i7-6700 CPU @ 3.40 GHZ 3.41GHZ; RAM: 32.0 GB). Except 100 simulations for each sample, this study also tests the prediction performance by setting the simulation times as 500 and 1000, and it is found that 100 simulation times is the best choice considering the accuracy and time assumption. Notably, this study focuses more about combining data-driven PHM framework and reliability analysis tools to analyze multiple degradation modes for complex systems, the performance of the proposed framework does not directly be compared with other methods. There is still a potential of optimizing the prediction accuracy by applying several optimization techniques, which are not focused in this study.

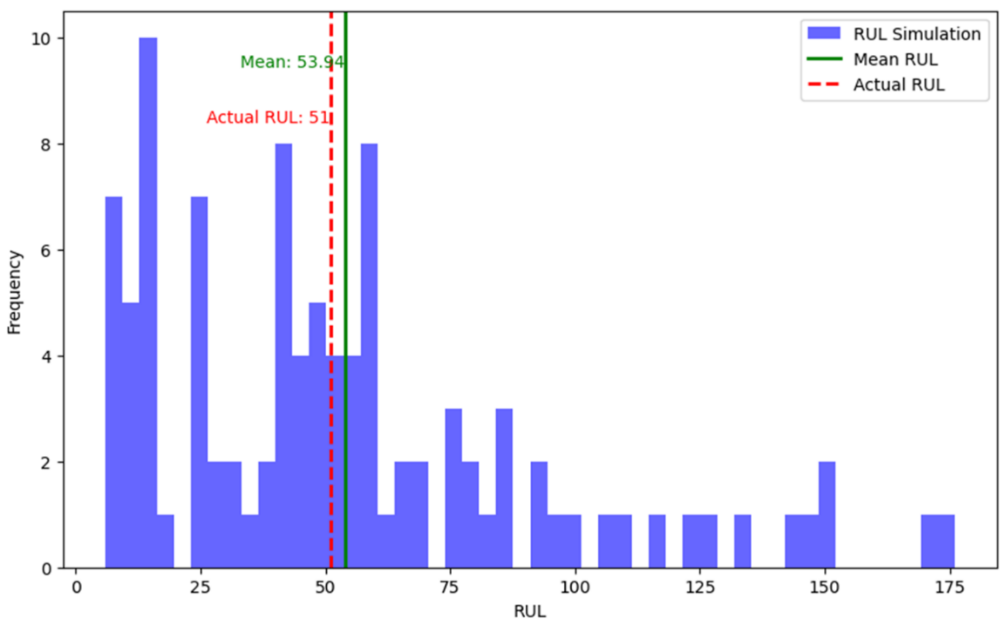

Figure 6. Prediction results of one sample from MC simulations

Table 4: Prediction performance

Simulation times	MAE	RMSE	Time consumed
100	25.54	33.17831	1.206h/4341.5111268s
500	25.4	32.77377	5.8386h/21018.9702064s
1000	25.46	32.90957	12.0815h/43493.4905502s

3.5. Prognostics-supported decision-making

The results from the Prognostics module (Part 3.2 to 3.4) support the decision-making process from two aspects. Firstly, with the built parameter repository, a new system's lifetime could be predicted before deployment. Managers or operators may build an initial understanding of the new system's degradation process according to historical data like the distribution of lifetimes, the earliest fault cycle, etc. Once the system starts operating, decisions could be made based on the availability of CM from data collection activities, which is dynamic. The prediction of RUL will be updated once a new CM is available, and thus a more agile maintenance strategy will be designed. It is possible to design maintenance actions at different levels according to the specific demands. For example, maintenance actions could be designed in two dimensions: Magnitude of maintenance actions, and Time of maintenance actions. In the dimension of Magnitude of maintenance actions, three periods could be defined as low-risk, high-risk, to extreme-risk, where different periods shall consider different types of actions. On the other hand, the dimension of Time of maintenance actions could also be defined as three levels: no actions, planning, and immediate intervention level. Considering both the two dimensions is expected to conduct more reliable and effective maintenance actions.

4. CONCLUSION

In this study, reliability analysis tools are embedded into one data-driven PHM framework in order to support the analysis of complex systems with multiple degradation modes. By applying reliability analysis tools, expert knowledge facilitates understanding systems' health status, and fault conditions, and even conducting Root Cause Analysis. Besides, the analysis according to the previous framework stays at the system level, and reliability analysis tools support diving into the component level, which increases the reliability and explainability of the proposed framework. The case study validates the prediction performance and feasibility of the proposed framework.

The future work will explore the potential of other reliability analysis tools in other multiple degradation mode scenarios. Optimization techniques will also be considered to improve the prediction accuracy without consuming much calculation time.

5. ACKNOWLEDGMENTS

This research is supported by the program of China Scholarship Council (No. CSC202207930021).

6. CITATIONS AND REFERENCES

Compare, M., Antonello, F., Pinciroli, L., Zio, E. (2022). “A general model for life-cycle cost analysis of Condition-Based Maintenance enabled by PHM capabilities.” Reliability Engineering & System Safety, 224, 108499.

Demichela, M., Baldissone, G., Darabnia, B. (2018). “Using Field Data for Energy Efficiency Based on Maintenance and Operational Optimisation. A Step towards PHM in Process Plants.” Processes, 6, 25.

Frederick, D. K., DeCastro, J. A., Litt, J. S. (2007). User’s Guide for the Commercial Modular Aero-Propulsion System Simulation (C-MAPSS) (No. E-16205).

Katina, P. F., Pyne, J. C., Keating, C. B., Komljenovic, D. (2021). “Complex System Governance as a Framework for Asset Management.” Sustainability, 13, 8502.

Meng, H., Li, Y. F. (2019). “A review on prognostics and health management (PHM) methods of lithium-ion batteries.” Renewable and Sustainable Energy Reviews, 116, 109405.

Omri, N., Al Masry, Z., Mairot, N., Giampiccolo, S., Zerhouni, N. (2021). “Towards an adapted PHM approach: Data quality requirements methodology for fault detection applications.” Computers in Industry, 127, 103414.

Pecht, M., Jaai, R. (2010). “A prognostics and health management roadmap for information and electronics-rich systems.” Microelectronics Reliability, 50, 317–323.