Multi-attribute Modelling for Oil Condition Assessment Considering Uncertainties

Yan Pan(潘燕), Tonghai Wu(武通海), Yunteng Jing(敬云腾) and Pujian Wang(王普健)

***Abstract*—Lubricating oil carries the primary wear failure information of the critical tribological components, therefore, severs for the condition-based maintenance of equipment. However, the oil condition assessment presents low reliability due to the uncertainties originating from the variable working conditions and the redundant indicators. To address the uncertainties with multiple indicators, a knowledge-guided three-layer model is established for characterizing the multi-attribute oil state. Further, data dispersion is considered by assigning fuzzy probability among the attribute layer. The inconsistent decisions are solved by improved evidential reasoning embedding inference rules in the state layer. The effectiveness of the proposed approach is verified using the real-world lubricant oil monitoring data from vehicle engines.**

***Index Terms*—oil condition assessment, expert system, evidential reasoning, uncertainty.**

List of Symbols

OCA	Oil Condition Assessment
ER	Evidential Reasoning
D-S	Dempster-Shafer
ES	Expert Systems
ROC	Receiver Operating Characteristic
BPA	Basic Probability Assignment
Bel	Belief measure
Pl	Plausibility measure
FER	Fuzzy Evidential Reasoning
FERIES	Fuzzy Evidential Reasoning Integrating Expert Systems
TBN	Total Base Number
a_{ij}	The oil indicator with monitoring data
$\bar{a}_{ij}$	The normalized value for a_{ij}
μ	The mean of Gaussian function
σ	The variance of Gaussian function
N	The number of oil grades
H_c	The oil grade
$\mu_{H_c}(a_{ij})$	The state membership possibility of monitoring data a_{ij}
$m_i(H_c)$	The basic probability of the *i*-th attribute as a member of H_c
$M_i(H_c)$	The membership of the *i*-th attribute as a member of H_c
Θ	The complement of H
a_{ij0}	The initial value of indicator a_{ij}
a_{ijg}	The failure value of indicator a_{ij}
A	The oil attribute
w_{ij}	The weight of the indicator
W_i	The weight of the attribute
$\widetilde{w}_k$	The weight of the rule

I. Introduction

LUBRICATING oil is used to reduce friction and wear in mechanical parts. The information contained in the oil is a direct indicator of the machine's operation conditions [1]. Hence, oil condition assessment (OCA) provides the first-line defense for detecting early deterioration to prevent potential equipment failures [2]. However, variable working conditions and redundant indicators in off-line inspection often cause irregular data dispersions and inconsistent decisions. These unavoidable uncertainties in the monitoring often generate the root of low reliability for traditional models in OCA.

To comprehensively characterize the oil degradation, various indicators are inspected from different oil attributes, including oxidation, pollutant content, metal element content, and additive content et al. [2]. With continuous samples, the variations of these indicators can be modeled jointly or independently by principles of the data sequence. Such models have been adopted dominantly so far [3],[4] in preventive maintenance. However, some intrinsic limitations are accompanied throughout. First, the dispersions of data series are frequently encountered because machines often work under variable conditions. Hence, significant errors are often produced for data-driven models. On the other hand, inconsistent decisions cannot be avoided among dozens of indicators without the coupling knowledge [5]. Thereby, the redundant information is involved as the disturbances for the final decision [6].

Several methods have been proposed to handle uncertainties

Manuscript submitted October 16, 2021. This work was supported in part by the National Science Foundation of China under grant No. 51975455 and the Industry-University-Research Cooperation Project of Aero Engine Corporation of China (HFZL2020CXY016). *(Corresponding author: Tonghai Wu.)*

Yan Pan is with the Key Laboratory of Education Ministry for Modern Design and Rotor-Bearing System, Xi'an Jiaotong University, Xi'an 710049, China (e-mail:yanwuyu123@163.com).

Tonghai Wu is with the Key Laboratory of Education Ministry for Modern Design and Rotor-Bearing System, Xi'an Jiaotong University, Xi'an 710049, China (e-mail: tonghai.wu@mail.xjtu.edu.cn).

Yunteng Jing is with the Key Laboratory of Education Ministry for Modern Design and Rotor-Bearing System, Xi'an Jiaotong University, Xi'an 710049, China (e-mail: 1102794667@qq.com).

Pujian Wang is with the XCMG Research Institute, Xuzhou 221004, China (e-mail: wpj115@163.com).

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIM.2022.3161707, IEEE Transactions on Instrumentation and Measurement

and can be classified into two types: data-driven and model-based methods. To describe the data dispersions, Liu [7] introduced the random noise of Brownian motion to simulate the stochastic dispersions based on the equilibrium concentration of particles in the lubricating oil. Moreover, Ian [8] demonstrated that probabilistic reasoning methods could effectively characterize the inevitable uncertainties with sparsity in off-line oil data. Furthermore, amounts of probabilistic methods have been applied, such as Bayesian probability [9], belief network [10], and fuzzy logic [3]. However, the data-driven methods lack interpretability due to the complex nature and unknown mechanism. Some researchers attempted to trace the uncertainty using model-based methods. For example, Wang [9] characterized the effects of internal and external variables in the stochastic process, modelling the oil uncertain state. Valis [11] proposed a joint stochastic diffusion process and fuzzy approach instead of deterministic models to detect stochastic errors in OCA. Nevertheless, oil degradation is a gradual process with the interaction of different attributes. It is hardly practical to characterize the oil state accurately with a single or few attributes. Further, another uncertainty in multi-attribute OCA, inconsistent decisions due to redundancy, needs to be resolved.

The introduction of additional knowledge, such as additional data or experience [12], is a viable solution to mitigate the uncertainties in multi-attribute OCA. Evidential reasoning (ER), based on Dempster-Shafer (D-S) theory [13], is effective for handling uncertainty in multi-attribute making-decision. Evidence constructed with multi-attributes can provide a more comprehensive representation of oil nature. Moreover, it can accommodate a broader range of uncertainty by including the fuzzy belief structure based on plausibility [14]. However, inconsistent decisions of different attributes still exist. Therefore, more profound inspection should be involved, e.g., expert systems (ES) [15].

Addressing the above uncertainties and viable solutions, a knowledge-guided model for OCA is proposed. Interaction mechanism is considered for building a three-layer model. To quantify the uncertainties, the distributed fuzzy membership is assigned between adjacent layers to form the basic probability assignment (BPA) of ER. By this, the quantified uncertainty is obtained from the multi-attribute decision with ER. Further, expert knowledge is used to guide the reasoning for treating inconsistent decisions. The major contributions are summarized.

1) A three-layer structure including indicator-attribute-state is proposed for OCA based on the oil degradation mechanism.
2) The uncertainty arising from data dispersion is quantified with the probability assignment in ER.
3) A modelling strategy that introduces expert knowledge for rule inference is proposed, where inference rules are combined by ER to solve inconsistent multi-attribute decisions.

The rest of this paper is organized as follows. In section 2, the related methods are illustrated. In section3, the knowledge guided model is introduced. Verifications of the proposed model are described in Section 4. Section 5 contains the conclusion.

II. Related Methods

In this section, related methods for uncertainty treatment have been illustrated, mainly involving the fuzzy membership assignment and D-S theory.

A. Fuzzy membership assignment

Fuzzy evaluation is used to characterize the data dispersion. The data series can be correlated with the oil state via membership functions.

Considering the long-duration and gradual degradation process, the state grade is defined as the quantified index H_c to categorize the full span. Here, c is the number of the current state grade with its maximum of N. Referring to the maintenance strategy of OCA of the target machine, N is set as 5, corresponding to the grades as $\{\text{excellent}, \text{good}, \text{average}, \text{poor}, \text{serious}\}$. Let $\{a_{i1}, \cdots, a_{ij}\}$ be the oil indicator set, and an oil state level H_c is treated as a fuzzy set $\mu_{H_c}(a_{ij})$, which denotes the possibility that a_{ij} belongs to H_c.Then the state is characterized by a membership function $\mu_{H_c}(a_{ij}) \in [0,1]$.

Normalization is adopted considering the different effects of the indicators on oil degradation. The indicators with positive effects, such as additive content, total base number (TBN), are called the value indicators. Other indicators are named cost indicators. $\bar{a}_{ij}$ denotes the normalized value for a_{ij}, that is,

$$\bar{a}_{ij} = \begin{cases} \dfrac{a_{ijg} - a_{ij}}{a_{ijN} - a_{ij0}}, & a_{ij} \in I_1 \\ \dfrac{a_{ij} - a_{ij0}}{a_{ijg} - a_{ij0}}, & a_{ij} \in I_2 \end{cases}, \tag{1}$$

where a_{ij0} and a_{ijg} are the initial and failure value of the full range dataset, I_1 represents the value-type indicator set, I_2 represents the cost-type indicator set.

Let $\mu_{H_c}(\bar{a}_{ij})$ be the membership of the normalized data $\bar{a}_{ij}$ that belongs to the state $H_c, c = 1, \cdots, N$. It can be computed from the fuzzy membership function as in Eq. (2),

$$\mu_{H_c}\left(\bar{a}_{ij}\right) = exp\left(-\left(\frac{\bar{a}_{ij} - \mu}{\sigma}\right)^2\right), \tag{2}$$

where μ and σ are respectively the mean and variance of Gaussian function, which are setting in parameters optimization. $\bar{a}_{ij}$ is the normalized indicator data of Eq. (1), and $\mu_{H_c}(\bar{a}_{ij})$ represents the state membership possibility of monitoring data a_{ij}.

B. D-S theory reasoning

To quantify the uncertainty, D-S theory is applied for multi-attributes decision-making in OCA [16]. Two parameters for measuring information, a belief measure (*Bel*) and a plausibility measure (*Pl*), are defined as:

$$Bel(H)=\sum_{A_k\subseteq H} m(A_i),\quad Pl(H)=\sum_{A_k\cap H\neq\varnothing} m(A_i), \tag{3}$$

where $m(A_i)$ is BPA of the evidence A_i, which indicates the degree of the belief in attribute A_k; $Pl(H)$ represents the plausibility that the oil state level members to H and can be described as,

$$Pl(H)=1-Bel(\Theta), \tag{4}$$

where Θ denotes the complement of H; $Bel(\Theta)$ represents the portion that uncertainty with evidence A_i, and the degree of uncertainty can be measured by it; The interval $[Bel(H), Pl(H)]$ is used to describe the decision confidence.

III. KNOWLEDGE GUIDED MODEL CONSIDERING UNCERTAINTIES

A. Three-layer structure modelling with multi-attribute

Essentially, lubricating oil degradation is the combined effects of the chemical and physical attributes. For OCA, indicators are often not traced to the attributes in monitoring. Accordingly, a three-layer structure of indicator-attribute-state is constructed for a comprehensive understanding of the degradation mechanism. Considering the uncertainty of data dispersion, a membership function is adopted. Similarly, attribute membership is determined with the associated indicators. The oil state can be reasoned by ER algorithm, and the membership of each attribute provides a BPA as the evidence.

The terminologies are defined as follows:

1) The state level is defined as grade set: $H=\{H_1,\cdots,H_c\}$, $c=2,\cdots,N$, where N is the number of oil grades.
2) The attribute set $A=\{A_1,\cdots,A_i\}$, $i=2,\cdots r$, where r is the number of attributes.
3) The indicator set is denoted as $\{a_{i1},\cdots,a_{ij}\}$, $j=2,\cdots g$, where g is the number of indicators of the i-th attribute.
4) The weight of the indicator set $\{w_{i1},\cdots w_{ij}\}$.
5) The weight of the attribute set $\{W_1,\cdots,W_i\}$.
6) The weight of rule set $\{\widetilde{w}_1,\cdots,\widetilde{w}_k\}$, $k=2,\cdots n$, where n is the number of rules.

In the indicator layer, each indicator is used to match the pre-set state level with fuzzy evaluation. In the attribute layer, the membership of an attribute is a combination of the corresponding indicators. In the state layer, the above attribute evidence determines the oil state by ER. This process is illustrated in Fig. 1.

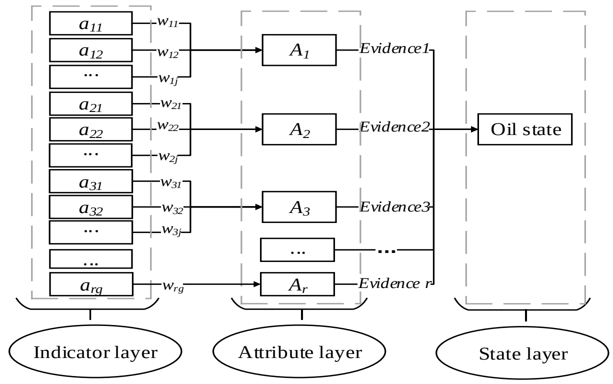


Fig. 1 Three-layer modelling structure in OCA.

The essence of the modelling is ER. Based on the D-S theory, the oil state is deduced by ER from the indicator-to-attribute. Two connections are involved: the fuzzy evaluation of the associated indicators provides the assessment for the corresponding attribute. Besides, the decision of the oil state is reasoned from attributes as distinctive pieces of evidence.

ER is used to handle uncertainty with multi-attribute decision-making, for instance, two attributes A_1 and A_2 belonging to a level H can be calculated as follows,

$$m_{(2)}(H)=(A_1\oplus A_2)(H)=\frac{\sum_{H(A_1)\cap H(A_2)=H}A_1A_2}{1-\sum_{H(A_1)\cap H(A_2)=\varnothing}A_1A_2}, \tag{5}$$

where $H(A_1)$ and $H(A_2)$ represent the focal elements in state, $\sum_{H(A_1)\cap H(A_2)=\emptyset}A_1A_2$ denotes the uncertainty degree in combination, $\oplus$ means the combination operator.

The quantity of belief assignment for the uncertainty $m_{(2)}(\Theta)$is:

$$m_{(2)}(\Theta)=(A_1\oplus A_2)(\varnothing)=\frac{\sum_{H(A_1)\cap H(A_2)=\varnothing}A_1A_2}{1-\sum_{H(A_1)\cap H(A_2)=\varnothing}A_1A_2}. \tag{6}$$

Besides, the comprehensive oil state can be obtained by aggregating multiple attributes as evidence and uses the orthogonal sum to combine multiple attributes as $A_1\oplus A_2\oplus\cdots\oplus A_r$. The recursive ER can be expressed with multi-attributes of multiple states [19]. Suppose that any state level H_c of two adjacent attributes performs evidence combination,

$$\begin{aligned} m_{i,i+1}(H_c)&=\left[m_i(H_c)+m_i(\Theta)\right]\left[m_{i+1}(H_c)+m_{i+1}(\Theta)\right]\\ &-m_i(\Theta)m_{i+1}(\Theta)\\ &=\prod_{i=i}^{i+1}\left[m_i(H_c)+m_i(\Theta)\right]-\prod_{i=i}^{i+1}m_i(\Theta). \end{aligned} \tag{7}$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIM.2022.3161707, IEEE Transactions on Instrumentation and Measurement

Evidence combination on r attributes yield the following recursive formula:

$$m_{(1,r)}\left(H_c\right)=\prod_{i=1}^{r}\left[m_i\left(H_c\right)+m_i\left(\Theta\right)\right]-\prod_{i=i}^{i+1}m_i\left(\Theta\right). \tag{8}$$

Similarly, the two parts of $m_{(1,r)}(\Theta)$ after evidence combination: the assignment probability $\tilde{m}_{(1,r)}(\Theta)$ due to incomplete evidence and the assignment probability $\bar{m}_{(1,r)}(\Theta)$ due to incomplete weight assignment, can be obtained.

$$\begin{aligned}\tilde{m}_{(1,r)}\left(\Theta\right)&=\prod_{i=1}^{r}m_i\left(\Theta\right)-\prod_{i=1}^{r}\bar{m}_i\left(\Theta\right),\\ \bar{m}_{(1,r)}\left(\Theta\right)&=\prod_{i=1}^{r}\bar{m}_i\left(\Theta\right).\end{aligned} \tag{9}$$

Since the distribution of information in the evidence synthesis is complete, the following equivalence is constructed,

$$\sum_{c=1}^{N}K\left[m_{(1,r)}\left(H_c\right)+\tilde{m}_{(1,r)}\left(\Theta\right)+\bar{m}_{(1,r)}\left(\Theta\right)\right]=1. \tag{10}$$

The value of K is obtained by solving the equation:

$$\begin{aligned}K^{-1}=&\prod_{i=1}^{r}\left(W_i m_i\left(H_c\right)+1-W_i\sum_{c=1}^{N}m_i\left(H_c\right)\right)\\ &-\left(N-1\right)\prod_{i=1}^{r}\left(1-W_i\sum_{c=1}^{N}m_i\left(H_c\right)\right),\end{aligned} \tag{11}$$

The following assignments of probabilities after composition can be obtained.

$$\begin{aligned}m_{(r)}\left(H_c\right)&=\frac{K\times m_{(1,r)}\left(H_c\right)}{1-K\times\bar{m}_{(1,r)}\left(\Theta\right)},\\ m_{(r)}\left(\Theta\right)&=\frac{K\times\tilde{m}_{(1,r)}\left(H_c\right)}{1-K\times\bar{m}_{(1,r)}\left(\Theta\right)}.\end{aligned} \tag{12}$$

In the recursive ER, the unknown parameters $m_i(H_c)$, which is defined as BPA in ER, needs to be computed. To construct the BPA, namely $m_i(H_c)$, the membership of corresponding indicator is obtained by Eq.(2). The membership of each attribute to the states can be defined with the joint possibility of the associated indicators. With the assigned weights for the indicators, the joint membership for the i-th attribute can be defined as:

$$M_i\left(H_c\right)=w_{ij}\mu_{H_c}\left(\bar{a}_{ij}\right), \tag{13}$$

where w_{ij} is the weights of the indicators in the i-th attribute needed to be optimized.

To assign the uncertainty in ER process, a correction factor α_R indicates that each evidence has the unknown portion, whose value needs to satisfy Eq.(14),

$$\prod_{i=1}^{r}\left(1-\alpha_R\times\frac{W_i}{\max\left(W_i\right)}\right)\le\delta, \tag{14}$$

where w_i is the weights of the attributes, δ is a sufficiently small non-negative real number denoting the decision uncertainty [17]. To satisfy the above constraint α_R is assigned by the decision-maker as 0.9 [18].

To define the uncertainty of the attribute, the relative corrected BPA are modified according to Eq.(15),

$$\begin{aligned}m_i\left(H_c\right)&=\alpha_R\times\frac{W_i}{\max\left(w_i\right)}\times M_i\left(H_c\right),\\ m_i\left(\Theta\right)&=1-\alpha_R\times\frac{W_i}{\max\left(W_i\right)}\times\sum_{c=1}^{N}M_i\left(H_c\right).\end{aligned} \tag{15}$$

Finally, the comprehensive assessment of oil state can be obtained by Eq. (12). In that case, two quantitative indicators are computed, namely belief degree $m_{(r)}(H_c)$ and uncertainty degree $m_{(r)}(\Theta)$. Where $m_{(r)}(H_c)$ denotes the aggregated probability assignment that the oil state is characterized as H_c, and $m_{(r)}(\Theta)$ measures the degree of uncertainty by r pieces of evidence.

B. Knowledge embedding in the three-layer structure

Based on the fuzzy membership assignment of the oil attribute in FER, each attribute as evidence can be linked to a probability assessment. However, insufficient evidence is unavailable to ensure reliable assessment, mainly indicated as uncertainty is too large for decision-making. To reduce uncertainty, a rule base with expert knowledge is applied to handle inconsistent attribute assessment. In addition, multiple rules in modeling instead of inconsistent attribute evidence further eliminate decision uncertainty. Based on this strategy, an improved model (abbreviation as FERIES) is illustrated in Fig. 2.

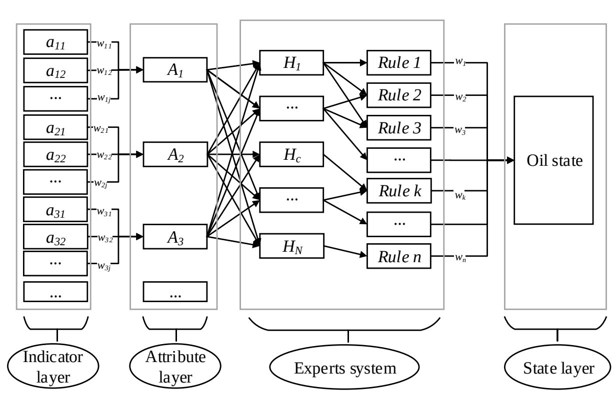


Fig. 2 The flowchart of FERIES.

In FERIES, limited pieces of evidence are orthogonally combined into multiple rules. Then the inference is made on each rule based on the prior expert knowledge, which is originated from two parts: the expertise and standard for the oil state judgment and the training samples in parameters optimization. Therefore, the combination of multiple rules compensates for the inconsistent assessment in ER.

Practically, the *k*-th inference rule is formed as follow.

IF: A_1^k is H_1 and $\cdots$ and A_r^k is H_N,
THEN:$\left\{\left(H_1, m_1^k(H_1)\right), \cdots, \left(H_N, m_r^k(H_N)\right), \left(\Theta, m^k(\Theta)\right)\right\}$,

where $A_i^k(i = 1,2,\cdots,r)$ is the attributes in the *k*-th rule, Θ represents the uncertain information without evidence, m^k represents the probability assignment in the *k*-th rule.

An orthogonal rule is developed based on expertise and standard for OCA. Each of rules contains the assessed level H and uncertainty Θ. A non-zero probability is assigned to any state level to prevent the zero-belief paradox in the evidence combination.

The activation weight $\widetilde{w}_k$ should be incorporated with the membership assignment, in which $\widetilde{w}_k$ of the *k*-th rule is,

$$w_k = \frac{\prod_{c=1}^{N}\prod_{i=1}^{r} M_i^k\left(H_c\right)}{\sum_{k=1}^{n}\prod_{c=1}^{N}\prod_{i=1}^{r} M_i^k\left(H_c\right)}, \tag{16}$$

where *N* is the number of states, *r* is the number of oil attributes, *n* is the number of inference rules, $M_i^k(H_c)$ is the membership of the *i*-th attribute as a member of H_c in the *k*-th rule.

It should be noted that the multi-attribute of oil is combined in FER, while FERIES is aimed at the inference rules instead of the limited attributes. Two factors ensure the reduction of uncertainty: the more knowledge with rule base and the more evidence from rules.

The combination of the inference rules is taken by Eqs. (7)-(12) to obtain the comprehensive assessment.

The definition of the BPA for inferential results should be considered based on the activation weight of the rule. Similar to FER, the BPA of each rule is assigned as,

$$\begin{aligned} m_i^k\left(H_c\right) &= \alpha_R \times \frac{w_k}{\max\left(w_k\right)} \times M_i^k\left(H_c\right), \\ m^k\left(\Theta\right) &= 1 - \alpha_R \times \frac{w_k}{\max\left(w_k\right)} \times \sum_{c=1}^{N} M_i^k\left(H_c\right), \end{aligned} \tag{17}$$

where $\max(\widetilde{w}_k)$ denotes the maximum weight in the activation weights, $M_i^k(H_c)$ represents the membership in the *k*-th rule.

C. Parameter and decision optimization

Parameter optimization includes two parts: the determination of state grades and the setting of parameters. The optimization strategies are construed with training samples.

To determine the state grades $\{H_1, \cdots, H_N\}$ with the oil indicators $\{a_{i1}, \cdots, a_{ij}\}$, several intervals can be categorized as$\{[a_0, a_1], (a_1, a_2], \cdots, (a_{c-1}, a_c], \cdots, (a_{N-1}, a_N]\}$, where a_c is the boundary of the corresponding state grade. Generally, the above categorization may vary for different oil indicators considering the different ranges. Therefore, the a_c should be determined independently for each indicator by optimization. Currently, the Receiver Operating Characteristic (ROC) curve is adopted as the optimization accordance [20]. The optimal boundary is selected from all possible interval boundaries based on the maximum Youden index point [21] located at the upper-left of the ROC. Thus, the primary classification can be accomplished by labelled data training.

To set the parameters in modelling, the optimal parameters are searched by the optimization algorithm. The optimization includes two types of parameters, 1) the weights including $W_i, w_{ij}, \widetilde{w}_k$, and 2) the mean μ and variance σ^2 in Eq. (2). The constraints of the parameters are constructed:

$$\begin{aligned} &\text{s.t.} 0 \le W_i \le 1, 0 \le w_{ij} \le 1, 0 \le w_k \le 1, 0 \le \mu \le 1, \sigma > 0 \\ &\sum_{i=1}^{r} W_i = 1, \sum_{i=1}^{r}\sum_{j=1}^{g} w_{ij} = 1, \sum_{k=1}^{n} w_k = 1. \end{aligned} \tag{18}$$

The labelled samples with known states are used as the training set, and the gradient descent or PSO algorithm [22] is applied to optimize the parameters.

Finally, with the optimal parameters, a decision-making criterion for ER is defined as,

$$\begin{cases} m_{(r)}\left(H_{N1}\right) - m_{(r)}\left(H_{N2}\right) > \varepsilon_0 \\ m_{(r)}\left(\Theta\right) < \varepsilon_1 \end{cases}, \tag{19}$$

where $m_{(r)}(H_{N1})$ is the maximum combined probability assignment for state H_{N1}, $m_{(r)}(H_{N2})$ is the second most combined probability corresponding to the state H_{N2}, $m_{(r)}(\Theta)$ is the measured probability for uncertainty. The threshold value ε_0 and ε_1 are set based on experiment and expertise. For example, ε_0=0.01 and ε_1=0.04 are feasible settings [18]. Consequently, H_{N1} that satisfies the constraints is the oil state. The properties can be summarized as follows:

1) The assessed oil state should have the maximum combined probability assignment and should be greater than that of other states by a certain threshold ε_0.
2) The measured probability for uncertainty should be less than a threshold ε_1.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIM.2022.3161707, IEEE Transactions on Instrumentation and Measurement

IV. CASE STUDY

To identify the uncertainty in oil states, the selected case includes a group of oil monitoring data with multiple attributes. The selected oil samples are periodically collected from real-world operating vehicles.

A. Procedures

The procedure of OCA includes:

Step1: Match the measurements with indicators to their corresponding states. Set the interval and parameters with the training samples optimization.

Step2: Obtain the indicator membership with the fuzzy approach. Weight indicators with the same attribute to obtain the membership of the oil attribute.

Step3: Formulate the rule base for oil state inference based on expert knowledge, then construct multiple rules with attributes as antecedents of the rule.

Step4: Combine the activation rules with FER to obtain a comprehensive assessment. Obtain the maximum probability assignment $m_{(r)}(H_c)$ and the suitable uncertainty of the oil state as the output.

The implementation flowchart is shown in Fig. 3.

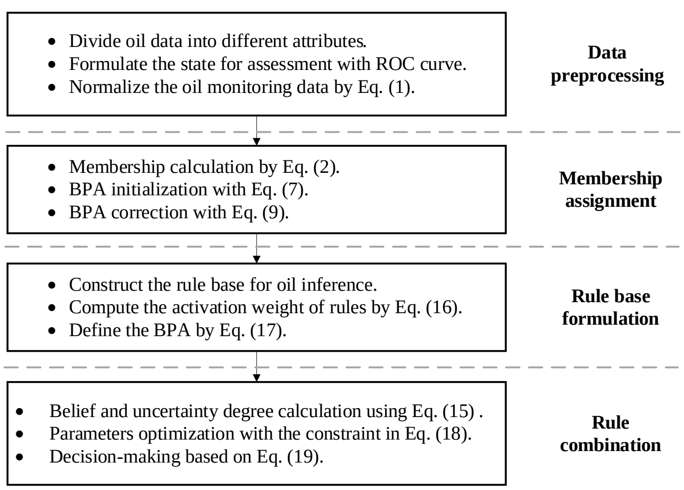


Fig. 3 Flowchart of state assessment by FERIES

B. OCA for loaders

As the most convincing evidence, real data experiments are necessary for our method in the paper. The real oil data is from the OCA project of a construction machinery company. The case is derived from the oil condition monitoring of the loader's hydraulic system. For two years, hydraulic oil was sampled regularly on the six loaders under different working conditions and operating environments. This project aims to establish intelligent OCA for preventive maintenance of loader machinery under various operating conditions. Ten indicators with failure thresholds are adopted according to the test specification for hydraulic oil. From the examination for each oil sample, five indicators with the highest weights are selected for further evaluation.

The indicator set $\{a_{11}, a_{12}, a_{21}, a_{22}, a_{31}\}$ includes viscosity, TAN, particle number, Fe content, and zinc content. The attribute set $\{A_1, A_2, A_3\}$ is formed, where A_1 represents the oil physicochemical attribute, A_2 represents the pollution attribute, A_3 represents the oil additive attribute. The setting of a_{ijg} refers to the criterion [23] and a_{ij0} is set as the value of new oil. The state set is defined as $H = \{H_1, H_2, H_3, H_4, H_5\}$ corresponding to the state set {excellent, good, average, poor, serious}. 36 oil samples were collected as shown in Fig. 4.

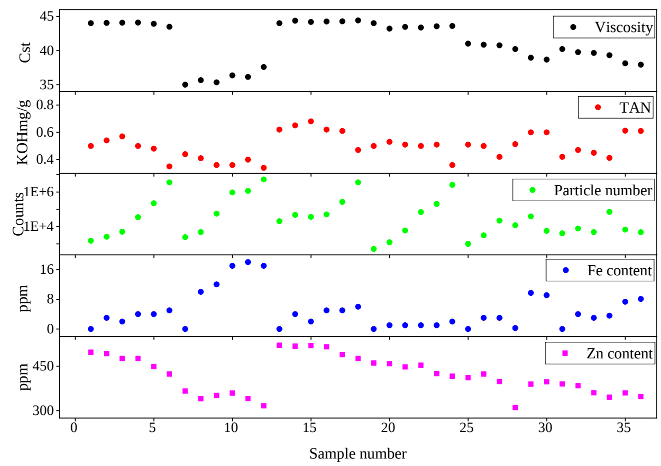


Fig. 4 The original oil data with five indicators

With the training samples shown in the literature [22], the weight sets of indicators and attributes are set as $\{0.42, 0.58; 0.59, 0.41; 1\}$ and $\{0.41, 0.30, 0.28\}$. Moreover, the ROC curves are used to determine the boundaries of the indicators, as shown in Fig. 5. With the well training model, the proposed methods are adopted for verification with 36 oil samples.

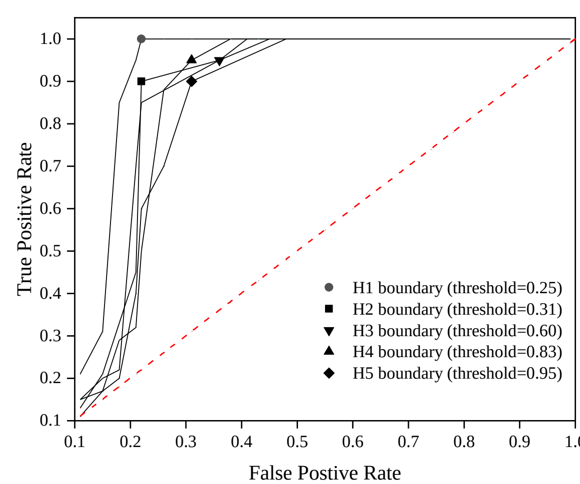


Fig. 5 The ROC determination for boundaries of Fe indicator

Furter, the 36 oil samples collected from the hydraulic system are processed by FER and FERIES. The assessments of three of these samples are shown in TABLE I. The accuracy of the FER method presents poor performance without considering evidence conflict. The $m_r(\Theta)$ in samples 1 and 3 exceeds the threshold in FER, which means the unachievable assessment.

TABLE I
THE RESULTS OF OCA WITH FER AND FERIES

Method	No.	$m_r(\Theta)$	$m(H_1)$	$m(H_2)$	$m(H_3)$	$m(H_4)$	$m(H_5)$
FERIES	1	0.01	0.01	0.01	**0.89**	0.07	0.01
FER	1	**0.05**	0.11	0.22	0.10	0.00	**0.53**
FERIES	2	0.00	0.00	0.01	**0.95**	0.03	0.00
FER	2	0.03	0.07	**0.53**	0.00	0.00	0.36
FERIES	3	0.00	0.00	0.00	0.02	**0.93**	0.04
FER	3	**0.05**	0.00	0.13	0.13	0.18	**0.51**

Uncertainties in OCA are effectively detected, as shown in Fig. 6. In FER, the uncertainty, arising from inconsistent attribute assessment, has exceeded the threshold specified in Eq. (19). In FERIES, the evidence is enhanced by the introduction of expert knowledge, and smaller uncertainties are obtained based on FER, indicating that a reduction in uncertainty is effectively achieved using FERIES method.

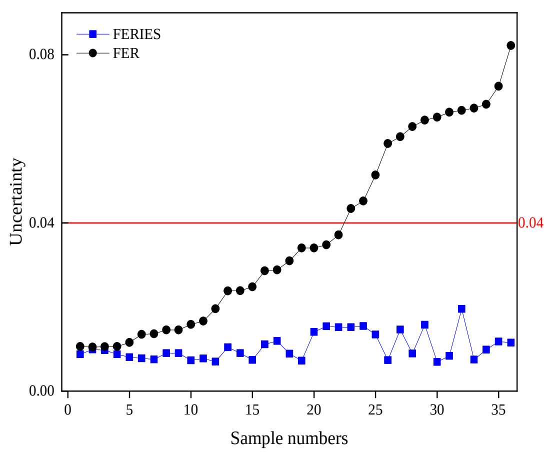


Fig. 6 The uncertainty detection with FER and FERIES methods

The assessments are shown in Fig. 7 with two methods. It can be seen that FERIES presents higher accuracy and consistency with real value, which has been assessed based on expert knowledge. The discrepancy between the FERIES and the real values is mainly manifested in the inconsistency of adjacent intermediate states (H2, H3, H4), which has less impact on the failure determination. On the other hand, the error in the FER decision is mainly expressed in the incorrect detection of the invalid state H5, which may lead to Type I errors [24], namely, incorrectly determining the failure state. The essence is that a single indicator failure is considered in the formulation of the rule. That is, the rule consequent is determined as a failure if anyone indicator exceeds the threshold. Thus, the risk of local failure for maintenance is avoided.

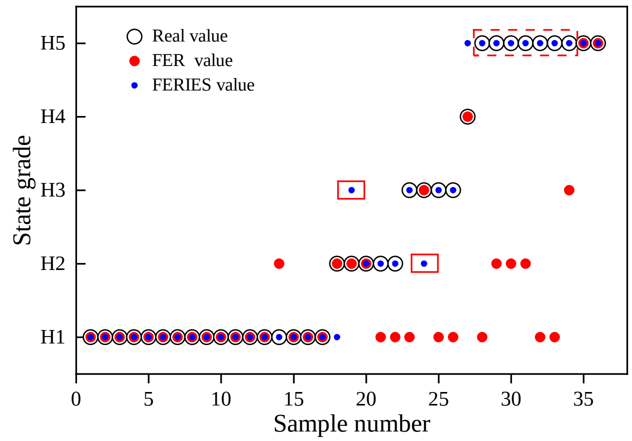


Fig. 7 The assessed comparison with FER and FERIES

C. Performance evaluation

Accuracy is a common and critical criterion to assess performance. However, accuracy itself is insufficient since high accuracy may give low consistency. Therefore, the accuracy rate (A) and the consistency rate (C) should be considered in OCA. The rates A and C are computed by Eqs. (20) and (21) [25],

$$A = \frac{T_{sp}}{T_{tc}}, \tag{20}$$

$$C = \sum_{i=1}^{N} \frac{P_i}{NS_{(i)}}, \tag{21}$$

where T_{sp} is the total number of samples with successful assessment and T_{tc} is the total number of samples, P_i is the number of the i-th state with successful assessment, $S_{(i)}$ is the total number of the ith states, N is the total number of states.

As illustrated in Table II, the attribute with higher weights is selected for ER. The oil physicochemical attribute, and the pollution attribute are selected in FERIES with two attributes. In FERIES with three attributes, all three attributes are used to compare. It can be found that both models present high performance, but FERIES with three attributes can obtain better accuracy and consistency.

TABLE II
THE COMPARISON OF DIFFERENT MODELS

Method	The number of attributes	Accuracy (A)%	Consistency (C)%
FE	3	69.4%	59.4%
FER	2	47.2%	23.9%
FER	3	50.0%	26.1%
FERIES	2	61.1%	38.2%
FERIES	3	94.4%	76.0%

To verify the performance of the models, the oil states are assessed by fuzzy evaluation (FE) [26], FER [18], and FERIES. Considering the impact of different numbers of indicators on the assessment, three and two attribute evidence are constructed respectively. The evaluation results are shown in Table II. It is obvious that the more attributes prove the better assessment. Compared with fuzzy evaluation, FER presents a worse assessment due to the uncertainty arising from inconsistent evidence, further validating that the inconsistency of multiple attributes probably brings additional uncertainty in decision-making. FERIES with the same attributes presents the best performance, which demonstrates the effectiveness to eliminate the uncertainty.

As the improved method, multi-attribute decision making is determined with FER, but it cannot remove uncertainty due to inconsistent attribute assessment. For the proposed model that combines ES and FER, there are two critical factors determining better performance. One is the enhancement of evidence with the rule base combination. The other is the introduction of expert knowledge, and the inference is much consistent with the real-world data, which is determined by expert knowledge. The assessment errors of the different methods are shown in Fig. 8, and it is confirmed that FERIES with three attributes shows excellent robustness and consistency. Therefore, the epistemic uncertainty can be eliminated by the introduction of knowledge, resulting in accurate oil condition assessment and identification.

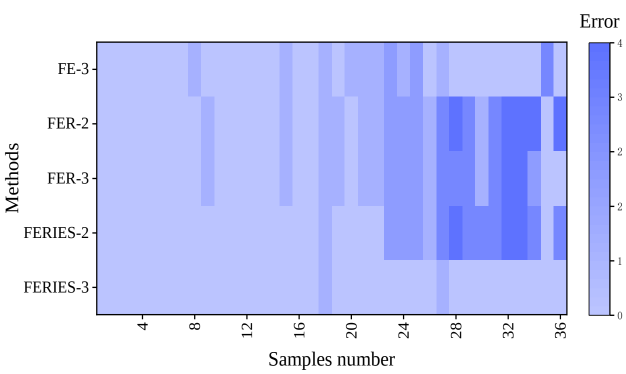


Fig. 8 The comparison of the errors of different methods

V. Conclusion

A knowledge-guided approach has been proposed shooting at the uncertainty problem in OCA. To deal with the uncertainty caused by data dispersions, a mechanism-driven three-layer model is used to integrate oil attributes based on the FER method. To reduce the uncertainty caused by inconsistent decisions, the improved model is proposed based on the introduction of inference rules. The main conclusions are as follows:

1) To quantify the uncertainty arising from data dispersions, an indicator-attribute-state three-layer model is used to assess the oil state combining the probability assignment and ER.
2) Modelling with the rule base in ER provides an effective solution for the uncertainty due to inconsistent decisions with multiple attributes.
3) The performance of the proposed model is verified with the multi-attribute data in real-world OCA.

There are still some limitations in the proposed method, and the accuracy of the model directly relies on the quantity of irregular oil data. In practice, limited by the complex condition and measured uncertainties, the oil data is dominated by small samples. We have taken the expertise and multiple attributes to deal with the decision uncertainty. To solve the uncertainty problem with insufficient data, the viable solution is to explore the essence of data law.

References

[1] Y. Du, C. Duan, and T. Wu, “Replacement Scheme for Lubricating Oil Based on Bayesian Control Chart,” *IEEE Trans. Instrum. Meas.*, vol. 70, pp. 1–10, 2021, doi: 10.1109/TIM.2020.3042231.
[2] J. M. Wakiru, L. Pintelon, P. N. Muchiri, and P. K. Chemweno, “A review on lubricant condition monitoring information analysis for maintenance decision support,” *Mechanical Systems and Signal Processing*, vol. 118, pp. 108–132, Mar. 2019, doi: 10.1016/j.ymssp.2018.08.039.
[3] D. Vališ, L. Žák, O. Pokora, and P. Lánský, “Perspective analysis outcomes of selected tribodiagnostic data used as input for condition based maintenance,” *Reliability Engineering & System Safety*, vol. 145, pp. 231–242, Jan. 2016, doi: 10.1016/j.ress.2015.07.026.
[4] L. Wei *et al.*, “Motor oil condition evaluation based on on-board diagnostic system,” *Friction*, vol. 8, no. 1, pp. 95–106, Feb. 2020, doi: 10.1007/s40544-018-0248-0.
[5] M. Maharana, N. Baruah, S. K. Nayak, N. Meher, and P. K. Iyer, “Condition Assessment of Aged Ester-Based Nanofluid Through Physicochemical and Spectroscopic Measurement,” *IEEE Trans. Instrum. Meas.*, vol. 68, no. 12, pp. 4853–4863, Dec. 2019, doi: 10.1109/TIM.2019.2900883.
[6] I. Morgan and H. Liu, “Computational Analysis of Sparse Datasets for Fault Diagnosis in Large Tribological Mechanisms,” *IEEE Trans. Syst., Man, Cybern. C*, vol. 41, no. 5, pp. 617–629, Sep. 2011, doi: 10.1109/TSMCC.2010.2073703.
[7] L. Yan, X. Youbai, Z. Fang, and Y. Zhigang, “Revision to the concept of equilibrium concentration of particles in lubrication system of machines,” *Wear*, vol. 215, no. 1–2, pp. 205–210, Mar. 1998, doi: 10.1016/S0043-1648(97)00269-X.
[8] J. Gajewski and D. Vališ, “The determination of combustion engine condition and reliability using oil analysis by MLP and RBF neural networks,” *Tribology International*, vol. 115, pp. 557–572, Nov. 2017, doi: 10.1016/j.triboint.2017.06.032.
[9] W. Wang and B. Hussin, “Plant residual time modelling based on observed variables in oil samples,” *Journal of the Operational Research Society*, vol. 60, no. 6, pp. 789–796, Jun. 2009, doi: 10.1057/palgrave.jors.2602621.
[10] X. Xu, X. Yan, C. Sheng, C. Yuan, D. Xu, and J. Yang, “A Belief Rule-Based Expert System for Fault Diagnosis of Marine Diesel Engines,” *IEEE Trans. Syst. Man Cybern, Syst.*, pp. 1–17, 2017, doi: 10.1109/TSMC.2017.2759026.
[11] D. Vališ, “Perspective approach in using anti-oxidation and anti-wear particles from oil to estimate residual technical life of a system,” *Tribology International*, p. 14, 2018.
[12] Y. Pan, Y. Jing, T. Wu, and X. Kong, “Knowledge-based data augmentation of small samples for oil condition prediction,” *Reliability Engineering & System Safety*, vol. 217, p. 108114, Jan. 2022, doi: 10.1016/j.ress.2021.108114.
[13] A. Ferrero and S. Salicone, “Modeling and Processing Measurement Uncertainty Within the Theory of Evidence: Mathematics of Random–Fuzzy Variables,” *IEEE Trans. Instrum. Meas.*, vol. 56, no. 3, pp. 704–716, Jun. 2007, doi: 10.1109/TIM.2007.894907.

[14] V. Kaftandjian, Y. M. Zhu, O. Dupuis, and D. Babot, "The combined use of the evidence theory and fuzzy logic for improving multimodal nondestructive testing systems," *IEEE Trans. Instrum. Meas.*, vol. 54, no. 5, pp. 1968–1977, Oct. 2005, doi: 10.1109/TIM.2005.854255.

[15] A. Haji and M. Assadi, "Fuzzy expert systems and challenge of new product pricing," *Computers & Industrial Engineering*, vol. 56, no. 2, pp. 616–630, Mar. 2009, doi: 10.1016/j.cie.2007.03.005.

[16] A. Ferrero and S. Salicone, "Fully comprehensive mathematical approach to the expression of uncertainty in measurement," *IEEE Trans. Instrum. Meas.*, vol. 55, no. 3, pp. 706–712, Jun. 2006, doi: 10.1109/TIM.2006.873799.

[17] Jian-Bo Yang and Pratyush Sen, "A general multi-level evaluation process for hybrid MADM with uncertainty," *IEEE Trans. Syst., Man, Cybern.*, vol. 24, no. 10, pp. 1458–1473, Oct. 1994, doi: 10.1109/21.310529.

[18] R. Liao, H. Zheng, S. Grzybowski, L. Yang, Y. Zhang, and Y. Liao, "An Integrated Decision-Making Model for Condition Assessment of Power Transformers Using Fuzzy Approach and Evidential Reasoning," *IEEE Trans. Power Delivery*, vol. 26, no. 2, pp. 1111–1118, Apr. 2011, doi: 10.1109/TPWRD.2010.2096482.

[19] Y.-M. Wang, J.-B. Yang, and D.-L. Xu, "Environmental impact assessment using the evidential reasoning approach," *European Journal of Operational Research*, vol. 174, no. 3, pp. 1885–1913, Nov. 2006, doi: 10.1016/j.ejor.2004.09.059.

[20] T. Fawcett, "An introduction to ROC analysis," *Pattern Recognition Letters*, vol. 27, no. 8, pp. 861–874, Jun. 2006, doi: 10.1016/j.patrec.2005.10.010.

[21] R. Fluss, D. Faraggi, and B. Reiser, "Estimation of the Youden Index and its Associated Cutoff Point," *Biom. J.*, vol. 47, no. 4, pp. 458–472, Aug. 2005, doi: 10.1002/bimj.200410135.

[22] Y. Pan, Y. Jing, T. Wu, and X. Kong, "An Integrated Data and Knowledge Model Addressing Aleatory and Epistemic Uncertainty for Oil Condition Monitoring," *Reliability Engineering & System Safety*, vol. 210, p. 107546, Jun. 2021, doi: 10.1016/j.ress.2021.107546.

[23] "ASTM D6224 - 16 Standard Practice for In-Service Monitoring of Lubricating Oil for Auxiliary Power Plant Equipment," 2016. https://www.astm.org/Standards/D6224.htm (accessed May 14, 2021).

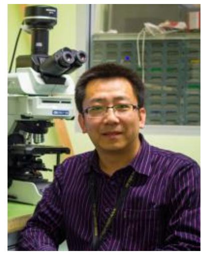
[24] J. A. de Diego, "TESTING TESTS ON ACTIVE GALACTIC NUCLEI MICROVARIABILITY," *AJ*, vol. 139, no. 3, pp. 1269–1282, Feb. 2010, doi: 10.1088/0004-6256/139/3/1269.

[25] T. Kari *et al.*, "An integrated method of ANFIS and Dempster-Shafer theory for fault diagnosis of power transformer," *IEEE Trans. Dielect. Electr. Insul.*, vol. 25, no. 1, pp. 360–371, Feb. 2018, doi: 10.1109/TDEI.2018.006746.

[26] C.-C. Sun, "A performance evaluation model by integrating fuzzy AHP and fuzzy TOPSIS methods," *Expert Systems with Applications*, vol. 37, no. 12, pp. 7745–7754, Dec. 2010, doi: 10.1016/j.eswa.2010.04.066.

Yan Pan received the M.S. degree in Applied Chemistry from Heifei University of Technology, Hefei, China, in 2011. He is currently pursuing the Ph.D. degree in mechanical engineering with the Key Laboratory of Education Ministry for Modern Design and Rotor-Bearing System, Xi'an Jiaotong University, Xi'an, China.

His current research interests include machinery condition monitoring, intelligent machinery diagnosis and prognosis, and reliability maintenance.



Tonghai Wu received the Ph.D. degree in Mechanical Engineering from Xi'an Jiaotong University, Xi'an, China, in 2006. He is currently a Professor with the School of Mechanical Engineering, Xi'an Jiaotong University.

From 2013 to 2014, he was a Visiting Scholar with University of New South Wales, Australia. His current research interests include machinery diagnosis and prognosis, machine vision on wear particle and surface analysis.

Yunteng Jing received the B.S. degree in Mechanical Engineering from Heifei University of Technology, Hefei, China, in 2020. He is currently pursuing the M.S. degree in mechanical engineering with the Key Laboratory of Education Ministry for Modern Design and Rotor-Bearing System, Xi'an Jiaotong University, Xi'an, China.

His current research interests include machinery condition monitoring, intelligent machinery diagnosis.

Pujian Wang received the M.S. degree in Chemistry from China University of Geosciences, Wuhan, China, in 2013. He is currently the senior engineer in the XCMG Research Institute, Xuzhou, China.

His current research interests include lubricant, oil condition monitoring.