IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received July 15, 2019, accepted July 23, 2019, date of publication August 12, 2019, date of current version August 28, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2933659

Radar Chart for Estimation Performance Evaluation

WEISHI PENG[1,2], YINGHONG LI[1], YANGWANG FANG[3], YUN WU[1], AND QING LI[2]

[1]Department of Aeronautics Engineering, Air Force Engineering University, Xi'an 710038, China
[2]Department of Equipment Management and Support, People Armed Police Engineering University, Xi'an 710086, China
[3]Unmanned System Research Institute, Northwestern Polytechnical University, Xi'an 710072, China

Corresponding author: Weishi Peng (peng_weishi@163.com)

This work was supported in part by the National Natural Science Foundation of China under Grant 71801222, in part by the National Science Foundation of Shaanxi Province of China under Grant 2018JQ6019, and in part by the National Postdoctoral Program for Innovative Talents under Grant BX201700104.

ABSTRACT Comprehensive measures for the estimation performance evaluation (EPE) has become increasingly prominent. This paper proposed a new radar chart evaluation method to measure the estimation performance. Firstly, the new radar chart index, which is composed of several popular incomprehensive measures, are presented, and the method of the weight of the each index is calculated based on vector ranking method. Secondly, the new comprehensive measures for the EPE is designed according to the fan area and the fan arc length. Finally, two cases study are provided to verify the effectiveness of this method.

INDEX TERMS Estimation algorithms, estimation performance evaluation, radar chart, decision support systems.

I. INTRODUCTION

Recently, an increasing number of estimation/filter algorithms were proposed in science and engineering, and many researchers claimed that their algorithms performed better than others. However, only a few of the convincing algorithms were applied to engineering practices. Therefore, it has extremely vital significance to explore and design more comprehensive measures for these evaluation algorithms. Clearly, performance evaluation of algorithms is as important as information fusion.

Estimation performance evaluation (EPE) has begun to gain more attention of researchers only in recent years though it was formally put forward by Li in 2006 ([1], [2]). Li found that the most commonly used root mean square error (RMSE) had two serious flaws, i.e., focusing on the greater errors excessively and no clear physical interpretation [3]. To solve the above problems, the average Euclidean error (AEE), harmonic average error (HAE), geometric average error (GAE), error mode (EM) and median error (ME) were provided by Li *et al.* [3]. Furthermore, Yin, et al, designed a new measure, i.e., the iterative mid-range error (IMRE), to solve the measure's robust problem [4]. As discussed above, we find that most existing measures on performance evaluation are some average of errors, which usually give 'big' or 'small' results to show the 'bad' or 'good' performance. Moreover, only one narrow aspect is considered by the above-listed measures in EPE, so, more comprehensive measures are desired.

Error spectrum (ES) is then proposed in [5], which is aggregating several commonly incomprehensive measures, such as the RMSE, AEE, GAE and so on. However, ES has several drawbacks, which attracts lots of researches to focus on the improvement of the ES. To solve the computation problem, Liu *et al.* presented the Mellin transform to compute the ES analytically [6], then, we proposed two algorithms to calculate the ES based on the power means error and the Gaussian mixture model, respectively [7], [12], [13]. Furthermore, to solve the dynamic evaluation problem, Mao *et al.* introduced a dynamic error spectrum (DES) to transform the ES into a single point at a time instant [8]. Obviously, the DES is an many-to-one mapping, which leads to the information loss problem. To tackle this problem, the range error spectrum induced area (RESA) and the DES induced area (DESA) were proposed in [9], [10]. Unfortunately, it's still hard to distinguish which estimator performs better by using the DES, RESA, and DESA. Consequently, Ma *et al.* proposed a volume error spectrum (VES) to further solve the dynamic systems evaluation problem [16]. However, the VES is still an many-to-one mapping.

The associate editor coordinating the review of this article and approving it for publication was Weimin Huang.

In many practical applications, radar chart, by which the status of to-be-evaluated object can be displayed intuitively [17], [18], is one of the most popular methods for comprehensive performance evaluation since its intuitive visualization. Radar Chart resembles a Plan Position Indicator-the typical two-dimensional layout of radar return from a radial trace, when used as a method of displaying data, it is much simpler than of doing statistical analyzes which can sometimes be complex because of the multi-dimensionality [19]. Furthermore, compared with traditional bar chart, radar chart has a stronger visual impact and can more intuitively display the characteristics of one object that are particularly prominent in an attribute [20]. Therefore, radar chart has been received a great amount of attention due to its increasing use in risk evaluation [20], multiple energy systems [21], e-waste management systems [22], process control [23], and medical field [24], etc. However, in traditional radar chart, the evaluation results are different due to the different order of indexes [25]. So, [26] proposed an improved radar chart to evaluate the performance of an object and the balanced degree of each index based on the fan area and the fan arc length, respectively; moreover, the included angles between adjacent index axes are equal, so it can not reflect the influence degree of indexes to evaluation objects.

In this paper, new radar chart evaluation method is proposed to EPE, which considers both incomprehensive measures and comprehensive measures for EPE. Firstly, the new radar chart index is composed of incomprehensive measures and comprehensive measures. Secondly, the vector ranking method is used to calculate the weight of the each index. Finally, the new comprehensive measures for the EPE is designed according to the fan area and the fan arc length of the radar chart.

This paper is organized as follows. The problem of the EPE is analyzed in III. Furthermore, the radar chart is proposed to estimation performance evaluation in Section III-D. In Section IV, numerical examples are used to show the validity and effectiveness of the new comprehensive measures. Finally, Section V concludes this paper.

II. PROBLEM FORMULATION

In EPE, the root mean square error (RMSE) is widely used to estimator comparison and estimation performance. However, the RMSE is seriously flawed. On one hand, the RMSE lacks clear physical interpretation. On the other hand, the RMSE is easily dominated by large error terms, for example, if all 100 terms of estimation error are around 1 except for one term of 500, then RMSE $\approx$ 55. Obviously, the evaluation result is determined by the biggest one term while ignored the other 99 terms. To see this, the most commonly measures for EPE is analyzed in the following part.

Let $\tilde{x}_q$ be the estimation error, i.e., $\tilde{x}_q = x_q - \hat{x}_q$ and $\|\tilde{x}_q\|_2 = (\tilde{x}_q' \tilde{x}_q)$, where $\hat{x}_q$ is the estimator, x_q is the quantity to be estimated; M is the Monte Carlo runs. In EPE, the most popular measure is the RMSE, i.e.,

$$RMSE(\hat{x}) = \left(\frac{1}{M} \sum_{q=1}^{M} \|\tilde{x}_q\|^2 \right)^{\frac{1}{2}} \tag{1}$$

Since the RMSE is dominated by large error, it's usually used to show the performance of a system at its worst.

On the contrary, the HAE is proposed to show the performance of a system at its best because it focus on the small errors.

$$HAE(\hat{x}) = \left(\frac{1}{M} \sum_{q=1}^{M} \frac{1}{\|\tilde{x}_q\|} \right)^{-1} \tag{2}$$

Furthermore, the AEE is given by

$$AEE(\hat{x}) = \frac{1}{M} \sum_{q=1}^{M} \|\tilde{x}_q\| \tag{3}$$

Obviously, we can utilize the AEE to evaluate the average performance of the system.

In addition, the GAE is balanced since it is neither dominated by large error nor affected by small error, i.e.,

$$GAE(\hat{x}) = \left(\prod_{q=1}^{M} \|\tilde{x}_q\| \right)^{\frac{1}{M}} \tag{4}$$

Furthermore, the IMRE is proposed to measure the stability of system performance under normal conditions, i.e.,

$$IMRE(\hat{x}) = \frac{\min\{\|\tilde{x}_q\|\}_{q=1}^{M} + max\{\|\tilde{x}_q\|\}_{q=1}^{M}}{2} \tag{5}$$

For some cases, the EM is proposed to find the performance of a system under normal conditions, which is the location of the highest peak of the histogram for the given error.

Since the existed measures, such as the RMSE, AEE, GAE and HAE can reflect only one aspect in EPE, the error spectrum was presented in [5]. In EPE, ES can reveal more information because it aggregates several commonly incomprehensive measures, such as the RMSE, AEE, GAE and HAE. In the following, the ES is given as

$$\begin{aligned} S(r) &= (E[e^r])^{1/r} = \{\int e^r dF(e)\}^{1/r} \\ &= \begin{cases} \{\int e^r f(e) de\}^{1/r} & if\ e\ is\ continuous \\ (\sum p_i e_i^r)^{1/r} & if\ e\ is\ discrete \end{cases} \end{aligned} \tag{6}$$

where $e = \|\tilde{x}\|$ or $e = \|\tilde{x}\|/\|\hat{x}\|$ are the absolute or relative estimation error norm, respectively; $F(e)$, $f(e)$, and p_i are the cumulative distribution function (CDF), probability density function (PDF), and probability mass function (PMF), respectively.

As pointed out in [7]–[10], the ES is in fact the power mean of e. Therefore, the error spectrum for EPE implies that the RMSE, the AEE, the GAE and the HAE have equal weight. In other words, the evaluation results are clearly subjective. To solve this problem, the radar chart is applied to the EPE.

III. RADAR CHART FOR ESTIMATION PERFORMANCE EVALUATION

According to the principle of radar map, first the index of the radar cloud should be designed; and then, the weight of each index in radar cloud are calculated based on the ranking vector method; furthermore, the radar chart suitable for EPE is draw according to the above index and the weight of each index; finally, the new comprehensive measures for EPE is presented by the sector area and sector arc length of the radar chart.

A. INDEX CONSTRUCTION OF THE RADAR CHART

Firstly, we summarize the commonly used measures in Table. 1. Furthermore, the above measures are map onto the radar chart in order to evaluate the estimation performance.

TABLE 1. Incomprehensive measures.

Metric	Pros, Cons and Good for
$RMSE$	$Pros:$ $clear\ physical\ interpretation$ $Cons: large\ error\ prevalence$ $Good\ for:$ $evaluate\ the\ average$ $performance\ of\ the\ system$
HAE	$Pros:$ $robust, measuring\ and\ comparing$ $good\ perfomance$ $Cons: small\ error\ dominance\ unclear$ $physical\ interpretation$ $Good\ for:$ $show\ the\ performance\ of\ a$ $system\ at\ its\ best$
AEE	$Pros:$ $clear\ physical\ interpretation$ $Cons: large\ error\ prevalence$ $Goodfor:$ $evaluate\ the\ average$ $performance\ of\ the\ system$
GAE	$Pros:$ $robust, extreme, error$ $Cons:$ $not\ popular$ $Good\ for:$ $measure\ the\ stability$ $of\ overall\ system$ $performance$
$IMRE$	$Pros::$ $robust, extreme, error$ $Cons::$ $not\ popular$ $Good\ for:$ $find\ the\ stability\ of\ system$ $performance\ under$ $normal\ conditions$
EM	$Pros:$ $more\ popular$ $Cons:$ $scenario\ sensitive$ $Good\ for:$ $find\ the\ performance$ $of\ a\ system\ under$ $normal\ conditions$

After obtaining the index of radar cloud chart, the next step is to determine the weights of each index.

B. WEIGHT COMPUTATION OF EACH INDEX

In [2], Yin proposed a ranking vector method for multiple-attribute estimation ranking. Furthermore, we proposed a improved ranking vector method by using the ES in multiple-attribute estimation ranking [11]. Inspired by the above ranking vector methods, the weights of each index in radar chart are calculated as follows.

1) PITMAN'S CLOSENESS MEASURE

Firstly, we introduce the Pitman's closeness measure (PCM). In [14], Pitman proposed a criterion to obtain the "joint" information by comparing the relative closeness of estimator $\hat{x}$ with the estimate x. Let $m(1, 2; a_i)$ be the measure of the difference between two compared objects s_1 and s_2 with respect to the i-th attribute a_i, that is,

$$m(s_1, s_2; a_i) = \begin{cases} 1 & if\ s_1 \succ s_2 \\ 0.5 & if\ s_1 = s_2 \\ 0 & if\ s_2\ \prec s_1 \end{cases} \tag{7}$$

where $s_1 \succ s_2$ represents that s_1 is better than s_2, and a_i is the attribute of the objects s_1 or s_2.

So, the multiple-attribute competition measure (MCM) is given by

$$M_{MCM}(s_1, s_2; a) = \frac{1}{n}\sum_{i=1}^{n} m(1, 2; a_i) \tag{8}$$

where a is the vector of the attributes. For $M_{MCM}(s_1, s_2; a) > 0.5$, we argue that s_1 is MCM-better than s_2 for a.

2) MULTIPLE-ATTRIBUTE COMPETITION MEASURE MATRIX

According to Eq. (8), we obtain the multiple-attribute competition measure matrix, i.e.,

$$X_{MCM} = \begin{bmatrix} M(s_1, s_1; a) & \cdots & M(s_1, s_m; a) \\ \vdots & \ddots & \vdots \\ M(s_m, s_1; a) & \cdots & M(s_m, s_m; a) \end{bmatrix} \tag{9}$$

where m is the number of the object.

Particularly, if $M(s_1, s_2; a) = 0$, let $M(s_1, s_2; a)$ equal 0.0001. So, there exists an eigenvector $\lambda > 0$ for the MCM matrix according to the Perron-Forbenius theorem [15].

Furthermore, we have

$$X_{MCM} \cdot \omega = \lambda \cdot \omega \tag{10}$$

where ω is the only eigenvalue in the spectral circle of X_{MCM}

3) WEIGHT ATTRIBUTE MATRIX

Assume that n is the number of index, so, the weight attribute matrix is defined as

$$R = \begin{array}{c c} & \begin{array}{cccc} a_1 & a_2 & \cdots & a_n \end{array} \\ \begin{array}{c} s_1 \\ s_2 \\ \vdots \\ s_m \end{array} & \begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1n} \\ r_{21} & r_{22} & \cdots & r_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ r_{m1} & r_{m2} & \cdots & r_{mn} \end{bmatrix} \end{array} \tag{11}$$

where r_ij is the values which can be obtained by means of expert scoring or questionnaire survey.

According to section A, Eq.(11) is rewritten as

$$R = \begin{array}{c c} & \begin{array}{ccc} RMSE & HAE & AEE \end{array} \\ \begin{array}{c} s_1 \\ \vdots \\ s_m \end{array} & \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ \vdots & \vdots & \vdots \\ r_{m1} & r_{m2} & r_{m3} \end{bmatrix} \end{array}$$

$$\begin{array}{ccc} GAE & IMRE & EM \\ \left. \begin{array}{c} r_{14} \\ \vdots \\ r_{m4} \end{array} \right. & \begin{array}{c} r_{15} \\ \vdots \\ r_{m5} \end{array} & \left. \begin{array}{c} r_{16} \\ \vdots \\ r_{m6} \end{array} \right] \end{array} \quad (12)$$

According to Eq. (9) and Eq. (10), we can calculate the eigenvalue ω of the weight attribute matrix. Furthermore, the ground elements in the eigenvectors are normalized

$$w_i^* = \frac{w_i}{\sum_{i=1}^{n} w_i} \quad (13)$$

In this paper, w_i^* is the weight of each measurements.

C. RADAR CHART DRAWING OF THE ESTIMATOR

According to the index in Table. 1, the radar chart for the estimation performance evaluation is designed as follows.

Step 1: Index normalization. For convenient, we apply the following equation to normalize the index in Table. 1. Let $r_{ji} \in \{r_{RMSE}, r_{HAE}, r_{AEE}, r_{GAE}, r_{IMRE}, r_{EM}\}$ for the j-th evaluation object ($j = 1, 2, \cdots, m$), the normalized index is given as

$$r_{ji}^* = r_{ji} / \sum_{i=1}^{n=6} r_{ji}$$

where n is the number of the index.

Step 2: Draw concentric circles. Let the number of concentric circles equal to n, and the radius $\{R_k^{cc}\}_{k=1}^{n}$ of the concentric circle is the sorted value of the index, i.e.,

$$R_1^{cc} = \min\{sort[\{r_{ji}^*\}_{i=1}^{n}]\} \leq \cdots \leq R_n^{cc} = \max\{sort[\{r_{ji}^*\}_{i=1}^{n}]\}$$

where $sort[\,\cdot\,]$ represents the ascending sort function.

Step 3: Design index axis. According to the number of index, the above concentric circles are separated by a number axis, and the included angle between the number axes is

$$\alpha_i = 2\pi w_i^*$$

where w_i^* is calculated by Eq. (13).

Furthermore, the RMSE, AEE, GAE, HAE, IMRE and EM are marked in the circle of the radar chart.

Step 4: Draw sector. Let the sector center angle and sector radius equal α_i and r_{ji}^*, respectively. Furthermore, we can obtain the sector of each index, as shown in Fig. 1. Furthermore, we can see that the smaller the area of the radar chart is the better performance of the system will be since the index of the radar chart are the error metrics.

D. NEW COMPREHENSIVE MEASURES FOR EPE

1) FEATURE DETECTION OF THE RADAR CHART

From Fig. 1, we have

$$\begin{cases} S_{ji}^s = \pi w_i^* v_{ij}^2 \\ L_{ji}^s = 2\pi w_i^* v_{ij} \end{cases} \quad (14)$$

where S_{ji}^s and L_{ji}^s, $i = 1, 2, \cdots, n$ and $j = 1, 2, \cdots, m$ are the sector area and sector arc length of the i-th index for the j-th evaluation object, respectively.

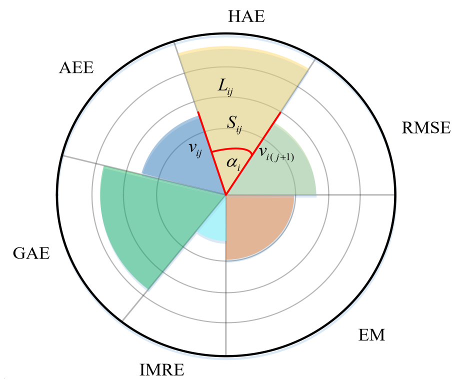

FIGURE 1. The radar chart of the estimated.

then, we obtain

$$\begin{cases} S_j^s = \dfrac{S_j}{\max\{S_j\}} = \dfrac{\sum_{i=1}^{n} \pi w_i^* v_{ij}^2}{\max\{\sum_{i=1}^{n} \pi w_i^* v_{ij}^2\}_{i=1}^{m}} \\ L_j^s = \dfrac{L_j}{2\pi\sqrt{S_j/\pi}} = \dfrac{\sum_{i=1}^{n} 2\pi w_i^* v_{ij}}{2\pi\sqrt{\sum_{i=1}^{n} w_i^* v_{ij}^2}} \end{cases} \quad (15)$$

where S_j^s is used for measuring the system efficiency of the evaluation object, and L_j^s is applied to evaluate the equilibrium degree of each individual index for the evaluation object.

2) DEFINITION OF THE NEW COMPREHENSIVE MEASURES

As pointed out in [10], the more flatness of an estimator means that the PDF of the estimation error is more concentrative than the desired one, which further illustrates the estimator is better. Since the sector area and sector arc length of the radar chart are essentially calculated by the estimation errors, we certainly hope both of them to be as small as possible. Therefore, the EPE problem using both S_j^s and L_j^s can be naturally changed into a bi-objective optimization problem [10]. Obviously, the most critical things in EPE is how to transform the two objective functions (i.e., S_j^s and L_j^s) into a single objective function. Here, the bi-objective optimization problem is defined as

$$NCM = min\, f(S_j^s, L_j^s) \quad (16)$$

where $f(\cdot, \cdot)$ represents the utility function. Next, we consider two popular forms of the utility function.

If prior preference about the weights is available, the new comprehensive measure (NCM) is defined as

$$NCM = f(M_{j1}, M_{j2}) = \sum_{l=1}^{2} \beta_l M_{jl} \quad (17)$$

where M_{j1} and M_{j2} are defined as follows.

$$\begin{cases} M_{j1} = \dfrac{S_j}{\min\{S_j\}_{j=1}^{m}} \\ M_{j2} = \dfrac{L_j}{\min\{L_j\}_{j=1}^{m}} \end{cases} \quad (18)$$

and β_1 and β_2 are the weights associated with the M_{j1} and M_{j2}, respectively, which holds for $\sum_{l=1}^{2} \beta_l = 1$.

In practical applications, the weights are determined by the users. It can be easily seen from Eq. (17) that if we focus the flatness of the estimator in EPE, we can let the weights $\beta_1 > \beta_2$. And if the weights satisfy $\beta_1 < \beta_2$, it means that the $f(M_{j1}, M_{j2})$ focuses more on the estimation accuracy of the estimator.

Certainly, if there is no information about the weights (β_1 and β_2), Eq. (17) is redefined as

$$NCM = \frac{M_{j1} + M_{j2}}{2} \tag{19}$$

In fact, Eq. (19) is the arithmetic mean which is dominated by large terms. To solve this problem, we use the geometric mean since it is neither dominated by large terms nor by small ones.

Thus, the NCM is rewritten as

$$NCM = \left(M_{j1} \times M_{j2}\right)^{1/2} \tag{20}$$

Clearly, Eq. (25) indicates that M_{j1} and M_{j2} are equally important in EPE.

From Eq. (17), Eq. (19) and Eq. (20), we obtain

$$NCM = \begin{cases} \sum_{k=1}^{2} \beta_k M_{ik} & \beta_k \ \ available \\ \begin{cases} (M_{j1} + M_{j2})/2 \\ (M_{j1} \times M_{j2})^{1/2} \end{cases} & \beta_k \ unavailable \end{cases} \tag{21}$$

Obviously, the estimator with a smaller *NCM* is better.

Theorem 1 (Properties of the NCM):

(a) $f(M_{j1}, M_{j2})$ satisfies the Regularity, i.e.,

$$f(0, 0) = 0 \tag{22}$$

(b) $f(M_{j1}, M_{j2})$ is monotonic, i.e., $\forall v, u, \quad v, u = 1, 2, \cdots,$ we have

$$f(M_{v1}, M_{v2}) \leq f(M_{u1}, M_{u2}) \tag{23}$$

when

$$\begin{cases} M_{v1} \leq M_{u1} \\ M_{v2} \leq M_{u2} \end{cases} \tag{24}$$

(c) $f(M_{j1}, M_{j2})$ is a continuous function, that is,

$$\lim_{\substack{M_{i1} \to M_{u1} \\ M_{v2} \to M_{u2}}} f(M_{v1}, M_{u2}) = f(M_{u1}, M_{u2}) \tag{25}$$

IV. COMPUTATIONAL PSEUDOCODE OF THE RADAR CHART METHOD

With the above preparations, we begin to present the radar chart method.

V. NUMERICAL EXAMPLES

Hereafter, the parament estimation case and state evaluation are proposed to illustrate the superiority of the radar chart method in EPE.

TABLE 2. Computational pseudocode of the radar chart method.

Input: estimation error $\{\tilde{x}_q\}_{q=1}^{M}$, the number of the index N.
Output: evaluation results

Step 1: Substituting the estimation error into RMSE, HAE, AEE, GAE, IMRE and EM, the index in the radar chart is given.

Step 2: Using the ranking vector method to calculate the weight of each index in the radar chart.

$$w_i^* = \frac{w_i}{\sum_{i=1}^{n} w_i}$$

Step 3: According to the section III-C, the radar chart for EPE is obtained.

Step 4: The new comprehensive measures are calculated

$$NCM = \begin{cases} \sum_{k=1}^{2} \beta_k M_{ik} & \beta_k \ \ available \\ \begin{cases} (M_{j1} + M_{j2})/2 \\ (M_{j1} \times M_{j2})^{1/2} \end{cases} & \beta_k \ unavailable \end{cases}$$

Finally, output the evaluation results.

A. PARAMETER ESTIMATION PERFORMANCE EVALUATION

1) PARAMETER ESTIMATION MODEL

Utilizing the following single noisy measurement [1], [5]:

$$z = x + u \tag{26}$$

where u follows a Gaussian distribution with zero mean and one variance, i.e., $u \sim \mathcal{N}(0, 1)$. x is generated from the following exponential prior:

$$f(x) = \begin{cases} \lambda \exp(-\lambda x) & x > 0 \\ 0 & x \leq 0 \end{cases} \tag{27}$$

Hereafter, we apply the maximum a posteriori (MAP) estimator and the minimum mean square error (MMSE) estimator to estimate the true x, respectively, where the former estimator is given by [5]

$$\hat{x}^{MAP}(\lambda) = max(z - \lambda, 0) \tag{28}$$

and the latter is calculated as

$$\hat{x}^{MMSE}(\lambda) = \frac{\exp(\frac{-(z-\lambda)^2}{2} + z - \lambda)}{\sqrt{2\pi}(1 - \Phi(\lambda - z))} \tag{29}$$

where $\Phi(\cdot)$ is the cumulative distribution function (CDF) of the standard normal distribution.

Furthermore, the estimation error is defined to calculate the existing measures and the proposed measures, i.e.,

$$\tilde{x} = x - \hat{x} \tag{30}$$

So, the estimation error of the MAP estimator is

$$\hat{x}^{MAP}(\lambda) = x - \hat{x}^{MAP} = x - max(x + u - \lambda, 0) \tag{31}$$

and one of the MMSE estimator is [29]

$$\hat{x}^{MMSE}(\lambda) = x - \hat{x}^{MMSE}(\lambda) \approx \lambda - u - \Delta \tag{32}$$

where $\Delta = (0.661|\lambda - z| + 0.3999\sqrt{(\lambda - z)^2 + 5.51})$.

2) SIMULATION RESULT ANALYSIS

Assume that the true value x is generated from Eq. (38) with the parameter $\lambda = 1$. Then, the parameter λ of the MAP and MMSE estimators are equal to 1.8, respectively, i.e., $\tilde{x}^{MAP}(\lambda_{MAP} = 1.8)$ and $\tilde{x}^{MMSE}(\lambda_{MMSE} = 1.8)$. Over 100,000 Monte Carlo runs, the metrics discussed in Table. 1 are listed in Table. 3.

TABLE 3. Performance evaluation measures value.

Measure	RMSE	HAE	AEE	GAE	IMRE	EM
MAP	1.02	0.06	0.77	0.46	0.72	0.23
MMSE	0.74	0.07	0.54	0.34	0.41	0.23

Furthermore, in practical applications, the values in the weight attribute matrix (R) can be obtained by means of expert scoring or questionnaire survey. Here, to verify the correctness of the proposed radar chart method, the values in the weight attribute matrix are generated randomly by the MATLAB software, i.e.,

$$R = \begin{array}{c} \\ MAP \\ MMSE \\ \\ MAP \\ MMSE \end{array} \begin{array}{ccc} RMSE & HAE & AEE \\ 0.45 & 0.40 & 0.18 \\ 0.11 & 0.42 & 0.49 \\ GAE & IMRE & EM \\ 0.39 & 0.84 & 0.32 \\ 0.37 & 0.98 & 0.49 \end{array} \tag{33}$$

Therefore, the eigenvalues of the above weight attribute matrix is

$$w^* = \begin{array}{cccccc} RMSE & HAE & AEE & GAE & IMRE & EM \\ [0.14 & 0.16 & 0.12 & 0.10 & 0.34 & 0.14] \end{array} \tag{34}$$

According to Table. 3, the radar chart and the ES of the MAP and MMSE estimators are as shown in Figs. 2 and 3, respectively.

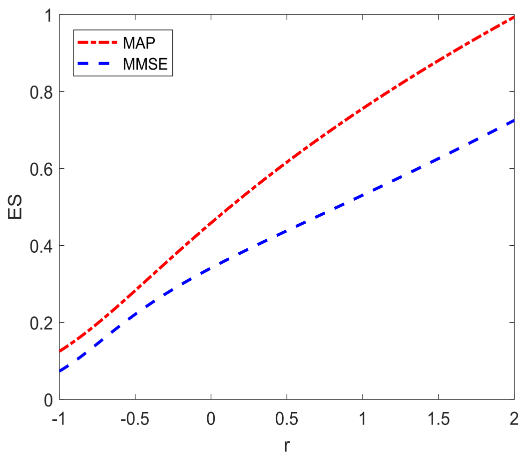

FIGURE 2. The ES curves of the MAP and MMSE estimators.

From Fig. 2, the ES curve of the MMSE estimation estimator is lower than that of the MAP estimator. Therefore, the MMSE estimator is better than the MAP estimator, i.e.,

$$MMSE \succ MAP \tag{35}$$

where $A \succ B$ means that A outperforms B.

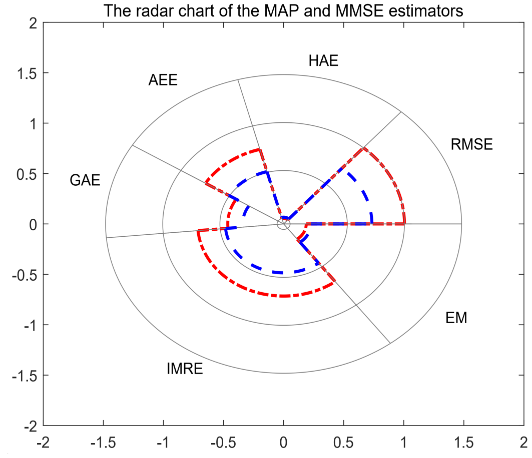

FIGURE 3. The radar chart of the MAP and MMSE estimators.

Furthermore, we can see from Fig. 3 that the radar chart of the MAP estimator is larger than the MMSE estimator. Hence, the MMSE estimator is superior to the MAP estimator in the case of $\lambda_{MAP} = \lambda_{MMSE} = 1.8$. Obviously, the radar chart can better reflect the performance of the estimator.

To verify this, we calculate the NCM according to Eq. (21), i.e.,

$$NCM_{MAP}{=}1.9930 > NCM_{MMSE}{=}1.2092 \tag{36}$$

Clearly, Eq. (36) shows that the MMSE estimator outperforms the MAP estimator, which is consistent with the Eq.(35). Furthermore, this result is still consistent with the results of literature [5], which further illustrates the correctness of the proposed method.

Next, the radar method is presented to evaluate the state estimation performance.

B. STATE ESTIMATION PERFORMANCE EVALUATION

1) THE STATE ESTIMATION MODELS

Suppose a nonlinear non-Gaussion model is given by

$$\begin{cases} \boldsymbol{x}_{k+1} = f(\boldsymbol{x}_k) + \boldsymbol{w}_k \\ \boldsymbol{y}_{k+1} = h(\boldsymbol{x}_{k+1}) + \boldsymbol{v}_{k+1} \end{cases} \tag{37}$$

where $f(\cdot)$ and $h(\cdot)$ are nonlinear functions as follows

$$\begin{aligned} f(\boldsymbol{x}_{k+1}) &= 0.5\boldsymbol{x}_k + \sin(0.04 \times \pi \times t) + 1 \\ h(\boldsymbol{x}_{k+1}) &= \begin{cases} {\boldsymbol{x}_{k+1}}^2/5 + \sin(\boldsymbol{x}_{k+1}) & k < 30 \\ \boldsymbol{x}_{k+1}/2 - 2 & Otherwise \end{cases} \end{aligned} \tag{38}$$

and the state noise $\boldsymbol{w}_k$ and the measurement noise $\boldsymbol{v}_k$ are

$$\begin{cases} w_k \sim Gamma(0, 0.01) \\ v_{k+1} \sim N(0, 0.01) \end{cases} \tag{39}$$

where $Gamma(\alpha, \beta)$ is the Gamma distribution with the shape $\alpha = 0$ and date parameters $\beta = 0.01$, and $N(\mu, \Sigma)$ is the Gaussian distribution with the mean $\mu = 0$ and variance $\Sigma = 0.01$.

Furthermore, we used four types of nonlinear estimation algorithm (NEA), i.e., extended Kalman filter (EKF) [30], unscented Kalman filter (UKF) [31], Particle Filter (PF) [32], and Gaussian Sum Filter (GSF) [33] to estimate the above nonlinear Gaussian models.

2) PARAMETER INITIALIZATION

The initialization parameters of the EKF, UKF, PF and GPF were summarized in Table 4. Furthermore, over the Monte-Carlo runs, the tracking results and the corresponding estimation error of the above four NEAs are as shown in Figs. 4 and 5, respectively.

TABLE 4. Initialization parameters.

Algorithm	Initialization parameters	Time
EKF	$\hat{x}_0 = 1, P_0 = 3/4$ $Q_0 = 3/4, R_0 = 0.01$	50s
UKF	$\hat{x}_0 = 1,\ P_0 = 3/4$ $Q_0 = 3/4, R_0 = 0.01$ $\alpha = 1,\ \beta = 0,\ \lambda = 0$	50s
PF	$N = 1000,\ x_0^{(i)} = 1,$ $w_0^{(i)} = 1,\ R_0 = 0.01$	50s
GPF	$Q_0^{(1)} = \frac{2}{4},\ Q_0^{(2)} = \frac{3}{4}$ $w_w^{(1)} = 0.6,\ R_0^{(1)} = 0.1$ $R_0^{(2)} = 0.2,\ w_v^{(1)} = 0.8$ $w_v^{(2)} = 0.2$	50s

For the radar chart, the weight attribute is similarly given by the MATLAB software, i.e.,

$$
R = \begin{array}{c} \\ EKF \\ UKF \\ PF \\ GSF \end{array}
\begin{array}{c}
\begin{array}{ccc} RMSE & HAE & AEE \end{array} \\
\left[\begin{array}{ccc} 0.66 & 0.47 & 0.18 \\ 0.11 & 0.42 & 0.49 \\ 0.36 & 0.37 & 0.18 \\ 0.46 & 0.27 & 0.68 \end{array}\right.
\end{array}
\qquad (40)
$$

$$
\begin{array}{c} \\ EKF \\ UKF \\ PF \\ GSF \end{array}
\begin{array}{c}
\begin{array}{ccc} GAE & IMRE & EM \end{array} \\
\left.\begin{array}{ccc} 0.39 & 0.54 & 0.32 \\ 0.37 & 0.38 & 0.49 \\ 0.89 & 0.44 & 0.72 \\ 0.27 & 0.62 & 0.39 \end{array}\right]
\end{array}
$$

So, we have

$$
w^* = \begin{array}{cccccc} RMSE & HAE & AEE & GAE & IMRE & EM \\ [0.15 & 0.14 & 0.16 & 0.15 & 0.21 & 0.19] \end{array} \qquad (41)
$$

Finally, the radar chart of the above four NEAs are as shown in Fig. 6.

3) ANALYSIS OF THE SIMULATION RESULTS

Clearly, we can see from Fig. 6 that the *RMSE* in the radar chart holds the following inequality

$$
RMSE_{EKF} > RMSE_{GSF} > RMSE_{UKF} > RMSE_{PF} \qquad (42)
$$

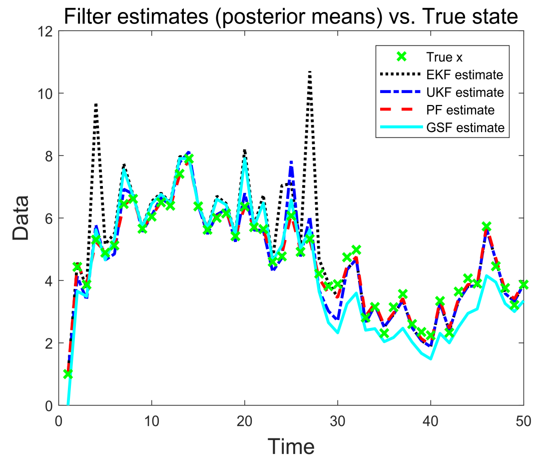

FIGURE 4. The tracking results of the above four NEAs.

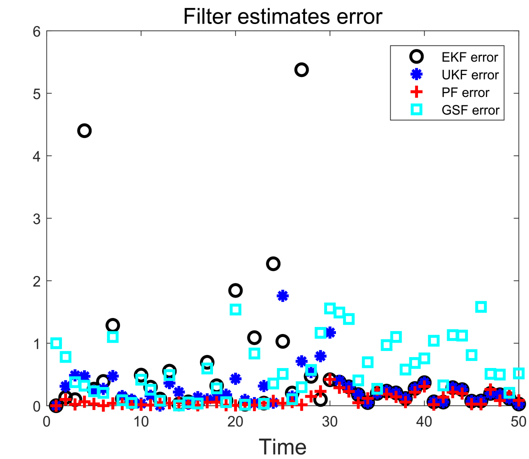

FIGURE 5. The estimation error of the above four NEAs.

Furthermore, compared with the area of the radar chart, the PF has the smallest radar chart; the second is the UKF; then the GSF and the EKF. To see this, the NCM is computed by the Eq.(21)

$$
\begin{cases} NCM_{EKF} = 4.1605 \\ NCM_{UKF} = 0.8157 \\ NCM_{PF} = 0.6667 \\ NCM_{GSF} = 1.8140 \end{cases} \qquad (43)
$$

That is,

$$
NCM_{EKF} > NCM_{GSF} > NCM_{UKF} > NCM_{PF} \qquad (44)
$$

Clearly, all the Eq. (42), Eq. (45) and Eq. (44) shown that if considering only the estimation accuracy, the PF is the best NEA; the next best is the UKF; then the GSF; the estimation accuracy of the EKF is the poorest among the four algorithms.i.e.,

$$
PF \succ UKF \succ GSF \succ EKF \qquad (45)
$$

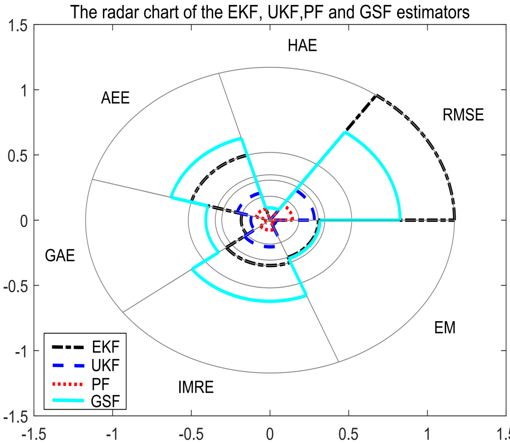

FIGURE 6. The radar chart of the above four NEAs.

Therefore, we can conclude that the proposed method can be used in EPE, and can reflect the performance of estimator objectively and comprehensively.

VI. CONCLUSION

The main contribution of this paper is threefold. First, the new radar chart has been proposed to estimation performance evaluation. Second, the fan area and the fan arc length, have been presented, where the former measures the estimation accuracy of an estimator and the latter quantifies the flatness of an estimator. Third, new comprehensive measures have been proposed to EPE, which includes the above two new measures. Finally, simulations show that the radar chart can be used to a variety of EPE directly, due to the consideration of more information, and the proposed measures can give more impartial evaluation results. Moreover, the other combinating form of the new comprehensive measures will be studied in the future work.

REFERENCES

[1] X. R. Li and Z. Zhao, "Evaluation of estimation algorithms part I: Incomprehensive measures of performance," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 42, no. 4, pp. 1340–1358, Oct. 2006.

[2] H. Yin, X. R. Li, and J. Lan, "Pairwise comparison based ranking vector approach to estimation performance ranking," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 48, no. 6, pp. 942–953, Jun. 2018.

[3] X. R. Li and Z. L. Zhao, "Measures of performance for evaluation of estimators and filters," *Proc. SPIE*, vol. 4473, Nov. 2001, pp. 530–541.

[4] H. Yin, X. R. Li, and J. Lan, "Iterative mid-range with application to estimation performance evaluation," *IEEE Signal Process. Lett.*, vol. 22, no. 11, pp. 2044–2048, Nov. 2015.

[5] X. R. Li, Z. L. Zhao, and Z. S. Duan, "Error spectrum and desirability level for estimation performance evaluation," in *Proc. Workshop Estimation Tracking Fusion A, Trbute Fred Daum*, Monterey, CA, USA, May 2007, pp. 1–7.

[6] Y. Liu and X. R. Li, "Computation of error spectrum for estimation performance evaluation," in *Proc 16th Int. Conf. Inf. Fusion*, Istanbul, Turkey, Jul. 2013, pp. 477–483.

[7] W.-S. Peng, Y.-W. Fang, R.-J. Zhan, and Y.-L. Wu, "Two approximation algorithms of error spectrum for estimation performance evaluation," *Optik*, vol. 127, no. 5, pp. 2811–2821, 2016.

[8] Y. Mao, C. Han, and Z. Duan, "Dynamic error spectrum for estimation performance evaluation: A case study on interacting multiple model algorithm," *IET Signal Process.*, vol. 8, no. 2, pp. 202–210, Apr. 2014.

[9] W.-S. Peng, Y.-W. Fang, and D. Chai, "Enhanced dynamic error spectrum for estimation performance evaluation in target tracking," *Optik*, vol. 127, no. 8, pp. 3943–3949, 2016.

[10] W. Peng, Y. Fang, Z. Duan, and B. Wang, "Enhanced error spectrum for estimation performance evaluation," *Optik*, vol. 127, no. 12, pp. 5084–5091, 2016.

[11] W. Peng, Y. Fang, X. Yong, and F. Wang, "Measures for multiple-attribute estimation ranking," *Optik*, vol. 127, no. 20, pp. 9479–9487, 2016.

[12] W. Peng, Y. Fang, R. Zhan, and W. Wang, "Weapon systems accuracy evaluation using the error spectrum," *Aerosp. Sci. Technol.*, vol. 58, no. 1, pp. 369–379, Nov. 2016.

[13] W. Peng, Y. Fang, and R. Zhan, "A variable step learning algorithm for Gaussian mixture models based on the Bhattacharyya coefficient and correlation coefficient criterion," *Neurocomputing*, vol. 239, no. 1, pp. 28–38, 2017.

[14] E. J. G. Pitman, "The 'closest' estimates of statistical parameters," *Math. Proc. Cambridge Philos. Soc.*, vol. 33, no. 2, pp. 212–222, 1937.

[15] R. B. Bapat and T. E. S. Rnghavan, *Nonnegative Matrices and Applications*. Cambridge, U.K.: Cambridge Univ. Press, 1997, pp. 4–7.

[16] Y. Ma, W. Peng, and H. Liu, "Volume error spectrum for estimation performance evaluation in dynamic systems," *Optik*, vol. 144, pp. 655–663, Sep. 2017.

[17] S. Wang, L. Hou, J. Lee, and X. Bu, "Evaluating wheel loader operating conditions based on radar chart," *Automat. Construct.*, vol. 84, pp. 42–49, Dec. 2017.

[18] Z. Wang and S. Qin, "Optimization of matching on torque converter with engine based on improved radar chart method," in *Proc. Int. Conf. Comput. Netw. Electron. Automat. (ICCNEA)*, Sep. 2017, pp. 370–373.

[19] M. J. Saary, "Radar plots: A useful way for presenting multivariate health care data," *J. Clin. Epidemiol.*, vol. 61, pp. 311–317, Apr. 2008.

[20] Y. Yanfei, "Comparison of radar chart creation methods for risk evaluation data," in *Proc. Int. Conf. Big Data Eng. Technol.*, 2018, pp. 51–55.

[21] S. Liu, Y. Yao, L. Ren, T. Zhu, J. Ren, J. Xu, Y. Jiang, C. Han, K. Xiao, L. Yang, and Z. Lin, "Combination weight and radar chart based evaluation method for utility tunnel planning in multiple energy systems," in *Proc. 2nd IEEE Conf. Energy Internet Energy Syst. Integr. (EI2)*, Beijing, China, Oct. 2018, pp. 1–6.

[22] K. Chaudhary and P. Vrat, "Case study analysis of e-waste management systems in Germany, Switzerland, Japan and India: A radar chart approach," *Benchmarking, Int. J.*, vol. 25, no. 9, pp. 3519–3540, 2018.

[23] S. Pongswatd and K. Smerpitak, "Applying radar chart for process control behavior," in *Proc. 3rd Int. Conf. Control Robot. Eng. (ICCRE)*, Nagoya, Japan, Apr. 2018, pp. 90–93.

[24] Y. A. A. Saito and Y. Aoki, "Nursing care using KT (Kuchi-kara Taberu) index radar chart enabling elderly patients with dysphagia to live like human beings after initiating gastrostomy feeding," *Proc. Singapore Healthcare*, vol. 2, no. 27, pp. 136–138, 2018.

[25] M. Friendly and E. Kwan, "Effect ordering for data displays," *Comput. Statist. Data Anal.*, vol. 43, pp. 509–539, Aug. 2003.

[26] L. Hongliang, L. Anxin, Z. Bin, Z. Tiefu, and Z. Xin, "A fuzzy comprehensive evaluation method of maintenance quality based on improved radar chart," in *Proc. Int. Colloq. Comput. Commun. Control Manage. (ISECS)*, Aug. 2008, pp. 638–642.

[27] Z. Zhao, X. R. Li, and Y. Zhu, "Interaction between estimator and estimation criteria," in *Proc. 7th Int. Conf. Inf. Fusion*, Philadelphia, PA, USA, Jul. 2005, p. 6.

[28] Z. Zhao and X. R. Li, "Two classes of relative measures of estimation performance," in *Proc. 10th Int. Conf. Inf. Fusion*, Montreal, QC, Canada, Jul. 2007, pp. 1–7.

[29] X. R. Li, *Probability, Statistics, and Random Signals*. Boca Raton, FL, USA: CRC Press, 1999.

[30] A. H. Jazwinski, *Stochastic Processes and Filtering Theory*. San Diego, CA, USA: Academic, 1970.

[31] S. J. Julier and J. K. Uhlmann, "Unscented filtering and nonlinear estimation," *Proc. IEEE*, vol. 92, no. 3, pp. 401–422, Mar. 2004.

[32] M. S. Arulampalam, S. Maskell, N. Gordon, and T. Clapp, "A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking," *IEEE Trans. Signal Process.*, vol. 50, no. 2, pp. 174–188, Feb. 2002.

[33] D. L. Alspach and H. W. Sorenson, "Nonlinear Bayesian estimation using Gaussian sum approximations," *IEEE Trans. Autom. Control*, vol. AC-17, no. 4, pp. 439–448, Aug. 1972.

WEISHI PENG received the B.S. and M.S. degrees from the People Armed Police Engineering University, Xi'an, Shaanxi, China, in 2010 and 2013, respectively, and the Ph.D. degree from the Department of Aeronautics and Astronautics Engineering, Air Force Engineering University, Xi'an, Shaanxi, in 2016. He holds a postdoctoral position with the School of Aeronautical Engineering, Air Force Engineering University, and is currently a Lecturer with the Department of Equipment Engineering, People Armed Police Engineering University. He has published more than 40 papers in journal, book, and conference proceedings, including journals such as the *Aerospace Science and Technology*, the *Neurocomputing*, the *IET Radar, Sonar & Navigation*, and the *International Journal for Light and Electron Optics*. He has published one book. His main research interests include the accuracy analysis and evaluation of the missile, machine learning algorithm, and estimation performance evaluation. He chaired the National Natural Science Foundation of China through Grant 71801222, the National Science Foundation of Shaanxi Province of China through Grant 2018JQ6019, and the National Postdoctoral Program for Innovative Talents through Grant BX201700104.

YINGHONG LI received the B.S. degree from the Air Force Engineering College, Xi'an, Shaanxi, China, in 1983, and the M.S. degree from the Nanjing University of Science and Technology, Nanjing, Jiangsu, China, in 1989. He is currently a Professor with the School of Aeronautical Engineering, Air Force Engineering University. He has published more than 170 papers in journal, book, and conference proceedings. His main research interests include the aviation propulsion theory and engineering research.

YANGWANG FANG received the B.S., M.S., and Ph.D. degrees from Huaibei Normal University, Huaibei, Anhui, Shaanxi Normal University, Xi'an Jiaotong University, Xi'an, Shaanxi, China, in 1987, 1990, and 1998, respectively. He is currently a Professor with the Department of Aeronautics and Astronautics Engineering, Air Force Engineering University. He has published more than 200 papers in journal, book, and conference proceedings, including journals such as the IEEE Transactions on Aerospace and Electronic Systems, the *Aerospace Science and Technology*, the *Neurocomputing*, the *IET Radar, Sonar & Navigation*, and the *International Journal for Light and Electron Optics*. He has published 10 books. His main research interests include the guidance and control of the missile, stochastic optimal control theory and application, and nonlinear control.

YUN WU received the B.S. and Ph.D. degrees from the Department of Aeronautics and Astronautics Engineering, Air Force Engineering University, Xi'an, Shaanxi, China, in 2003 and 2008, respectively. He is currently a Professor with the School of Aeronautical Engineering, Air Force Engineering University, and with the School of Mechanical Engineering, Xi'an Jiaotong University. He has published more than 100 papers in journal, book, and conference proceedings. His main research interests include aviation propulsion, plasma-assisted flow control, ignition, and combustion.

QING LI received the B.S. degree from the People Armed Police Engineering University, Xi'an, Shaanxi, China, in 2012, where she is currently a Teaching Assistant Lecturer with the Department of Equipment Engineering. Her main research interests include the accuracy analysis and evaluation of the missile, the development and design of non-lethal weapons, and police equipment and estimation performance evaluation.