Article

Performance Evaluation Metrics and Approaches for Target Tracking: A Survey

Yan Song [1] [ID], **Zheng Hu** [1] [ID], **Tiancheng Li** [1,*] [ID] and **Hongqi Fan** [2] [ID]

[1] Key Laboratory of Information Fusion Technology, Northwestern Polytechnical University, Xi'an 710072, China; syzx@mail.nwpu.edu.cn (Y.S.); zheng.hoo@mail.nwpu.edu.cn (Z.H.)
[2] Key Laboratory of Science and Technology on ATR, National University of Defense Technology, Changsha 410073, China; fanhongqi@nudt.edu.cn
[*] Correspondence: t.c.li@nwpu.edu.cn; Tel.: +86-029-88431306

Abstract: Performance evaluation (PE) plays a key role in the design and validation of any target-tracking algorithms. In fact, it is often closely related to the definition and derivation of the optimality/suboptimality of an algorithm such as that all minimum mean-squared error estimators are based on the minimization of the mean-squared error of the estimation. In this paper, we review both classic and emerging novel PE metrics and approaches in the context of estimation and target tracking. First, we briefly review the evaluation metrics commonly used for target tracking, which are classified into three groups corresponding to the most important three factors of the tracking algorithm, namely correctness, timeliness, and accuracy. Then, comprehensive evaluation (CE) approaches such as cloud barycenter evaluation, fuzzy CE, and grey clustering are reviewed. Finally, we demonstrate the use of these PE metrics and CE approaches in representative target tracking scenarios.

Keywords: performance evaluation; cloud barycenter evaluation; fuzzy CE; grey clustering

1. Introduction

Target tracking is widely involved in many problems of significance, such as military defense, automation/driverless transportation, intelligent robots, and so on. There are many outstanding works to provide guidance on the implementation of target-tracking algorithms. For the issue of multitarget tracking, the International Society of Information Fusion (ISIF) took track estimation, data association, and performance evaluation into account many years ago. The textbook [1], which is a tutorial of the known target-tracking algorithms and the class material [2], can assist researchers in re-implementing these methods and developing advanced methods. Some common toolboxes include the Recursive Bayesian Estimation Library (ReBEL) [3–5], the nonlinear estimation framework [6–8], and the Tracker Component Library [9]. The Open Source Tracking and Estimation Working Group (OSTEWG), which is the working group of the ISIF, revisited the current widely used methods to make up an open-source framework in [10,11], which is named Stone Soup and is available from the website https://github.com/dstl/Stone-Soup/ (accessed on 10 November 2021). Another ISIF working group, the Evaluation of Techniques of Uncertainty Reasoning Working Group (ETURWG), develops the uncertainty description in the domain of target tracking and introduced the Uncertainty Representation and Reasoning Evaluation Framework (URREF) [12] to Stone Soup.

This paper focuses on the performance evaluation (PE) of target-tracking algorithms, which plays a key role in the comparison of existing algorithms and in putting forward a new algorithm. PE is referred to the assessment and the evaluation of various performance metrics of a system [13,14], whose significance lies in providing evaluation results of the system performance [15], as well as the reference basis for the optimization of the system performance.

The basic process required for tracking evaluation is that when both truth targets and tracks are available, the first step is to find an association between the true target and a track so that performance measures can be computed. However, it was assumed that there is a unique association in this paper. The association algorithms can be found in [16–19]. After the tracks were assigned to targets, we computed the various performance measures to analyze the target-tracking algorithms and optimize these algorithms.

For the practitioner, the PE problems can be divided into two stages. The first is to choose the relevant effective metrics, and the second is to evaluate a single score through these metrics [20]. In this paper, we reviewed the measures to evaluate the performance of the target tracking system. These metrics were grouped into the categories of the correctness, timeliness, and accuracy. The assessed results of each metric were weighted and combined to give an overall performance measure. Further, we designed simulations that employed several PE approaches based on these metrics to illustrate their use in the target-tracking problem.

The rest of this paper is organized as follows. Section 2 introduces general evaluation metrics and sorts categories by the characteristics of each metric. Section 3 reviews classic PE approaches. The simulation results of our approach in the context of target tracking are given in Section 4. Section 5 concludes our work and faces remaining challenges.

2. A Classification of the Comprehensive Evaluation Metrics

In the context of target tracking, a variety of evaluation metrics with physical significance have been proposed, which can evaluate the practicability of the tracking algorithm and the consistency of the expected and assessed results. These metrics can be divided into three categories: effectiveness, timeliness, and accuracy, which can be seen in [21–24]. This paper also followed this division criterion for convenience. What the correctness [18,25–27] usually represents is the number of missed/false targets, etc. The timeliness assesses the time performance of the estimated track [23,28], which is a crucial measure for online target tracking. The accuracy metrics can be defined in different ways according to the different scenario requirements, in which the (root-) mean-squared error ((R)MSE) is commonly used in trajectory error (TE), tracking position error (TPE), and tracking velocity error (TVE) [29], and the other accuracy metrics refer to [28]. References [30,31] combined the multiple-object-tracking precision (MOPT) and the multiple-object-tracking accuracy (MOTA) to describe the effectiveness and the timeliness of the multiple-object-tracking systems, but they disregarded the effect of the error. In addition, the measures, such as the cross-platform commonality, the track purity, the processor loading, etc., were considered in [7,26,32–36]. The comprehensive evaluation (CE) metric system is shown in Figure 1.
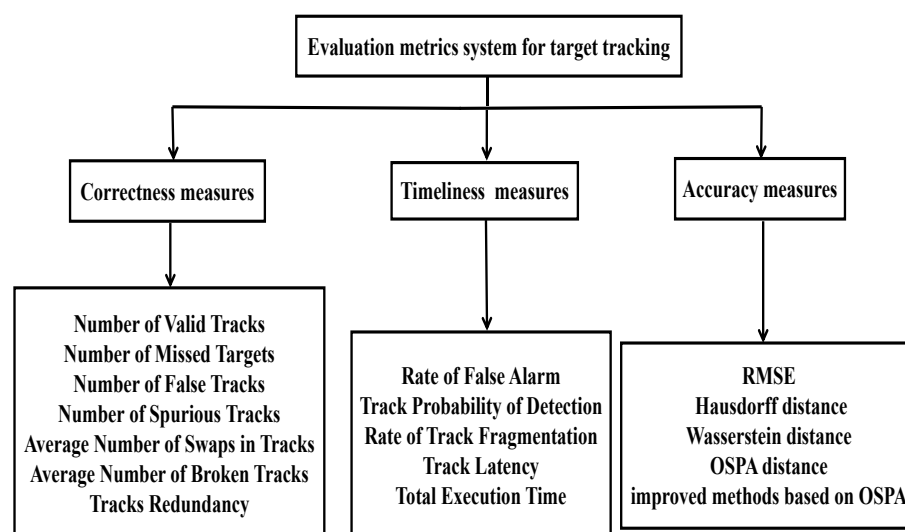

Figure 1. Classification of representative CE metrics.

2.1. Correctness Measures

The correctness measures the numerical characteristics of the acquired data and calculates how many mistakes the tracker made in terms of misses, false tracks, and so forth, which can be briefly described by Figure 2, where the small and large circles represent the truth target and the estimated track, respectively. The solid and dashed lines denote the trajectory of the target and the curve of the track, respectively. Given a time interval $t \in [t_1, t_2]$, the correspondence between the target and the track is established in Figure 2. These measures are explained in detail below.
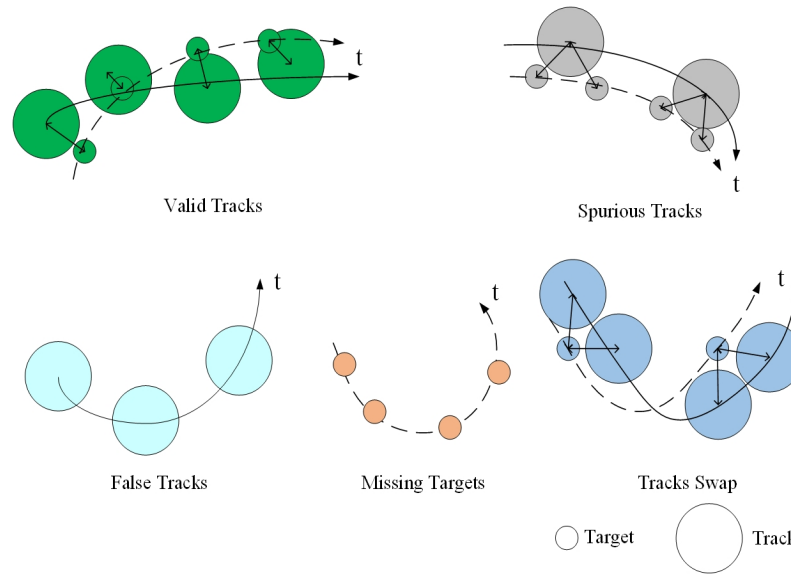

Figure 2. Mapping the tracker hypotheses to objects. In the easiest case, different associations result in evaluation metrics.

- Number of valid tracks (NVT):
  If a track is assigned to a target and the target has only one track, then the track is validated. $N_{valid}(t)$ denotes the NVT at time t;
- Number of missed targets (NMT):
  A target is missed if it is not associated with any track. $N_{missed}(t)$ is the NMT at time t [26];
- Number of false tracks (NFT):
  A track is false if it is not assigned to any target. $N_{false}(t)$ denotes the NFT at time t;
- Number of spurious tracks (NST):
  A track is defined as spurious if it is assigned to more than one target, and the NST is denoted as $N_{spur}(t)$;
- Average number of swaps in tracks (ANST):
  Different confirmed tracks may be assigned to a particular truth. This can happen when crossing targets and targets come close to each other. Assume $N_{swap}(t)$ as the ANST at time t [26];
- Average number of broken tracks (ANBT):
  There is also probable that no track is assigned to the truth for several time steps. If there is no assigned track to the truth, the number of broken tracks is counted at each time step. Reference [26] employed the ANBT to check the track segment associated with the truth;
- Tracks redundancy (TR):
  TR is represented as the ratio of validated tracks and total assigned tracks:

$$TR(t) = \frac{N_{valid}(t)}{N_{valid}(t) + N_{spur}(t)}. \tag{1}$$

2.2. Timeliness Measures

The performance measure provides more information about the track persistence, which is also an indispensable part of the evaluation metrics [37]. Some timeliness metrics for the PE are given as follows:

- Rate of false alarms (RFA):
  The RFA [38] is defined as the NFT per time step, which can be denoted as follows:

$$FLR(t) = \frac{N_{false}(t)}{t}. \tag{2}$$

- Track probability of detection (TPD):
  In the time interval $[t_1, t_2]$, assume t^i_{first} and t^i_{last} as the first and last time that the ith target is present, respectively. According to [39], the TPD of each target is represented as:

$$P^i_d = \frac{t'}{t^i_{last} - t^i_{first}}, \tag{3}$$

where t' denotes the persistent duration where the ith target is assigned to a valid track;

- Rate of track fragmentation (RTF):
  It is likely that the track obtained through some tracking algorithms may not be continuous sometimes. The track segment is assigned to the ith truth, the number of changes that the continuous track becomes fragmental is defined as TFR_i when the track segment is assigned to the ith truth. The smaller the RTF is, the more persistent the tracking estimated by the algorithm is [39];

- Track latency (TL):
  The TL, the delay from the moment that the target arises in the view of the sensor to the moment that target is detected by the tracker in the running period, is a measure of the track timeliness;

- Total execution time (TET):
  The computational cost is another important factor to be considered in the PE of target tracking. Therefore, the total time that is taken to run the tracker is expressed as the TET for each tracking algorithm.

2.3. Accuracy Measures

The measure, favored by the majority of researchers, is a primary choice in evaluating the target tracking, in which several measures can be defined as based on the type of distance between the set of truths and tracks:

- RMSE:
  The RMSE is defined in terms of the estimation error e_k, which is the average difference between the estimated state $\tilde{X}_k$ and the truth state X_k, as:

$$R^l(k) = \sqrt{\frac{1}{n} \sum_{k=1}^{n} \|e_k\|^2}, \tag{4}$$

where n denotes the number of the lth targets detected at the tth time step, $\|e_k\|^2 = e_k^T e_k$. Th MSE/RMSE has long been the dominant quantitative performance metric in the field of signal processing. For the traditional target-tracking algorithms, the aim is to minimize them between the target truth and the estimated track [40], which is not suitable for the track assignments that do not have a one-to-one correspondence. At present, the CE metrics have been widely used in the Hausdorff distance [41], the Wasserstein distance [42,43], and the optimal subpattern assignment (OSPA) distance [20,44–46];

- Hausdorff distance:

The Hausdorff distance is a common method of measuring the distance between two sets of objects, which can be used to measure the similarity between tracks and is given by:

$$d^H(X, \tilde{X}) \triangleq max\{\max_{x \in X} \min_{\tilde{x} \in \tilde{X}} d(x, \tilde{x}), \max_{\tilde{x} \in \tilde{X}} \min_{x \in X} d(x, \tilde{x})\}, \tag{5}$$

where $X = \{x_1, x_2, \cdots, x_k\}$, $\tilde{X} = \{\tilde{x}_1, \tilde{x}_2, \cdots, \tilde{x}_k\}$. x_i and $\tilde{x}_i$ are the ith target and the ith track, respectively. $d(x, \tilde{x})$ is the Euclidean distance between x and $\tilde{x}$. It has been proven that the Hausdorff distance is very useful in assessing the multitarget data fusion algorithms [47–49]. Meanwhile, the distance is relatively insensitive to differences in the numbers of objects;

- Wasserstein distance:

The Wasserstein metric was initially used to measure the similarity of probability distributions [50] and was proposed for the sets of targets in [51]. The Wasserstein distance between X and $\tilde{X}$ is:

$$d^W(X, \tilde{X}) = \min_{\pi \in \Pi_n} \sqrt{\frac{1}{n} \sum_{i=1}^{n} \left\| x_i - \hat{x}_{\pi(i)} \right\|_2^2}, \tag{6}$$

where $X = \{x_1, x_2, \cdots, x_m\}$, $\tilde{X} = \{\tilde{x}_1, \tilde{x}_2, \cdots, \tilde{x}_n\}$. Π_n is the set of all permutations of $\{1, 2, \cdots, n\}$. $\pi \in \Pi_n$ can be written as $\pi = (\pi(1), \pi(2), \cdots, \pi(n))$. $\left\| x_i - \hat{x}_{\pi(i)} \right\|_2^2 = (x_i - \hat{x}_{\pi(i)})^T (x_i - \hat{x}_{\pi(i)})$. The Wasserstein distance extends and provides a rigorous theoretical basis for a natural multitarget miss distance. However, it lacks a physically consistent interpretation when the sets have different cardinalities [52];

- OSPA distance:

The OSPA was proposed to overcome the insensitive shortcoming, whose parameters can deal with the problem that the numbers of elements in the two sets do not match [53]. The OSPA metric between X and $\tilde{X}$ is:

$$d^{OSPA}(X, \tilde{X}) \triangleq \left(\frac{1}{n} \left(\min_{\pi \in \Pi_n} \sum_{i=1}^{m} d^{(c)}(x_i, \tilde{x}_{\pi(i)})^p + c^p(n - m) \right) \right)^{\frac{1}{p}}, \tag{7}$$

where $d^c(x, \tilde{x})$ is the cut-off distance between x and $\tilde{x}$ and $d^c(x, \tilde{x}) = \min\{d(x, \tilde{x}), c\}$, c denotes the truncation parameter, and p is the OSPA metric order parameter. The choice of parameters was given in [53]. The OSPA distance has been used widely in the literature [44,54–56] and has better properties for the multitarget error evaluation than the Hausdorff metric.

In addition, there are various improved methods based on the OSPA [57], which are be enumerated as follows:

- Generalized OSPA (GOSPA):

In the GOSPA metric, we look for an optimal assignment between the truth targets and the estimated tracks, leaving missed and false targets unsigned [58]. The GOSPA metric penalizes localization errors for properly detected targets, the NMT, and the NFT [59]. The GOSPA can be represented as an optimization over assignment sets:

$$d_p^{(c,2)}(X, \tilde{X}) = \left[\min_{\gamma \in \Gamma} \left(\sum_{(i,j) \in \gamma} d^p(x_i, \tilde{x}_j) + \frac{c^p}{\alpha}(n + m - 2|\gamma|) \right) \right]^{\frac{1}{p}}, \tag{8}$$

where $\gamma \subseteq \{1, 2, \cdots, n\} \times \{1, 2, \cdots, m\}$, $(i, j), (i, j') \in \gamma \rightarrow j = j'$, and $(i, j), (i', j) \in \gamma \rightarrow i = i'$. Γ denotes the set of all possible assignment sets γ. α is the additional parameter to control the cardinality mismatch penalty; in general, $\alpha = 2$. The terms $\frac{c^p}{\alpha}(m - |\gamma|)$ and $\frac{c^p}{\alpha}(n - |\gamma|)$ represent the costs (to the pth power) for the NMT and the NFT, respectively;

- OSPA-on-OSPA (OSPA$^{(2)}$) metric:

  The OSPA$^{(2)}$ metric [60,61] is the distance between two sets of tracks, which establishes an assignment between the real and the estimated trajectories that is not allowed to change with time and enables capturing the tracking errors of fragmentation and track switching. It is also simple to compute and flexible enough to capture many important aspects of the tracking performance. The OSPA$^{(2)}$ distance is defined as follows:

$$d_{p,q}^{(c,2)}(X, \tilde{X}; \omega) \triangleq \left(\frac{1}{n} \left(\min_{\pi \in \Pi_n} \sum_{i=1}^{m} d_q^{(c)}(x_i, \tilde{x}_{\pi(i)}; \omega)^p + c^p(n-m) \right) \right)^{\frac{1}{p}}, \quad (9)$$

  where q is the order of the base distance and w is a collection of convex weights.

Finally, we introduce the single integrated air picture (SIAP) metric. Despite the terminology, it is applicable to tracking in general and not just in relation to an air picture. It is made up of multiple individual metrics [62,63]. The SIAP metric requires an association between tracks and targets. We used a unique association in this paper. A description of the key metrics of the SIAP is given in Table 1.

Table 1. The key metrics of the SIAP.

Metric	Description
Ambiguity	A measure of the number of tracks assigned to each true object
Completeness	The percentage of live objects with tracks on them
LS	The percentage of time spent tracking true objects across the dataset
LT	1/R, where R is the average number of excess tracks assigned; the higher this value, the better
Positional Accuracy	Given by the average positional error of the track to the truth
Spuriousness	The percentage of tracks unsigned to any object
Velocity Accuracy	The average error in the velocity of the track to the truth
Number of Targets	The total number of targets
Number of Tracks	The total number of tracks

3. CE Approaches

The above metrics provide the criteria for the PE of the tracking algorithm, which are combined to give the overall performance by the CE model. In this section, we review several CE approaches to analyze and judge the performance of target tracking.

3.1. The Weight of Each Evaluation Metric Set

The analytic hierarchy process (AHP) is given to ascertain the proportion of each metric, which was used to model and analyze the evaluation metric of the PE according to layers [64–66]. The specific steps are as follows.

Step 1 The assessment metric system:

According to Section 2, all levels of evaluation metrics are established. $U_i(i \in [1, m])$ in $U = \{U_1, U_2, \cdots, U_m\}$ is the ith metric of the primary metric in the system; $U_{ij}(j \in [1, l])$ in $U_i = \{U_{i1}, U_{i2}, \cdots, U_{im}\}$ is the jth metric of U_i; $U_{ij} = \{U_{ij1}, U_{ij2}, \cdots, U_{ijm}\}$, U_{ijk} is the kth metric of U_{ij}. The rest can be established in the same manner;

Step 2 The comparison matrix A_{ij}:

Citing the numbers 1-9 as a scale, each influencing metric U_i in the above metric set U is determined according to the importance of the element to its corresponding quantized value A_{ij}. $A = (A_{ij})_{n \times n}$, $A_{ji} = \frac{1}{A_{ij}}$, $A_{ii} = 1$;

Step 3 The maximum eigenvalue λ_{max} of A and the corresponding normalized eigenvector W:

W is denoted as $W = (W_1, W_2, \cdots W_n)$, $\sum_{i=1}^{n} W_i = 1$, where W_i denotes the weight of the ith evaluation metric.

3.2. Cloud Barycenter Evaluation

Based on the traditional fuzzy set and probability theory, the cloud theory provides a powerful method by combining the qualitative information with the quantitative data [67]. As a kind of mathematical model, cloud theory describes the mapping relationship between quality and quantity through a fuzzy and stochastic relation completely [68]. The cloud barycenter evaluation developed from cloud theory is a CE method that has been extensively used in numerous complex systems, especially in the military field [67]. The cloud barycenter evaluation method is a qualitative and quantitative method to achieve the transformation between the conception and the data.

The cloud is represented by three digital characteristics, including the expected value E_x, the entropy E_n, and the hyper entropy H_e [67,69], where E_x is the central value of the fuzzy concept in the defined domain, E_n represents the degree of the fuzziness of the qualitative concept, and H_e is the fuzzy measurement and the entropy of E_n, which is a mapping of the uncertainty of the qualitative concept.

The cloud barycenter evaluation is realized by establishing the cloud model of each metric. The specific evaluation processes are as follows:

Step 1 The cloud model of the comment set:

The comment set of metrics was ascertained by experts. For example, we set $S=$ { excellent, good, fair, worse, poor } to denote the comment set of target tracking, which is shown in Table 2. We set the comment as the corresponding continuous number field interval [0,1]. The formula of the cloud model is represented as:

$$\begin{cases} E_{xi0} = \frac{c_{min} + c_{max}}{2}, \\ E_{ni0} = \frac{c_{max} - c_{min}}{6}, \end{cases} \tag{10}$$

where E_{xi0}, E_{ni0} are the expected values and entropy of some qualitative comments, respectively.

Table 2. Number field variation interval of the comments set.

Comments	Excellent	Good	Fair	Worse	Poor
Number field interval	[1,c1]	[c1,c2]	[c2,c3]	[c3,c4]	[c4,0]

Step 2 The quantitative and the qualitative variables for the given metric set;

(a) The cloud model of quantitative metrics:
The corresponding quantitative metrics values were established by n experts as $E_{x11}, E_{x21}, \cdots, E_{xn1}$, which can be denoted by the cloud model:

$$E_{x1} = \frac{E_{x11} + E_{x21} + \cdots + E_{xn1}}{n}, \tag{11}$$

$$E_{n1} = \frac{\max(E_{x11}, E_{x21}, \cdots, E_{xn1}) - \min(E_{x11}, E_{x21}, \cdots, E_{xn1})}{6}. \tag{12}$$

(b) The cloud model of qualitative metrics:
In the same way, every qualitative metric, which are represented by the linguistic value, can also be described by the cloud model:

$$E_{x2} = \frac{E_{x12}E_{n12} + E_{x22}E_{n22} + \cdots + E_{xn2}E_{nn2}}{E_{n12} + E_{n22} + \cdots + E_{nn2}}, \tag{13}$$

$$E_{n2} = E_{n12} + E_{n22} + \cdots + E_{nn2}, \tag{14}$$

where $E_{x12}, E_{x22}, \cdots, E_{xn2}$ and $E_{n12}, E_{n22}, \cdots, E_{nn2}$ denote the expected values and entropy of the cloud model, respectively;

Step 3 The weighted departure degree:

$S = (S_1, S_2, \cdots, S_n)$ is the n-dimensional integrated barycenter vector, each dimension value of which is calculated by $S_i = g_i \times h_i (i = 1, 2, \cdots, n)$, where $g_i = (E_{x1}, E_{x2}, \cdots, E_{xn})$ is the cloud barycenter position and $h_i = (W_1, W_2, \cdots, W_n)$ is the cloud barycenter height calculated by the AHP. $S^0 = (S_1{}^0, S_2{}^0, \cdots, S_n{}^0)$ denotes the ideal cloud vector. The synthesized vector is normalized as follows:

$$
S_i^T = \begin{cases} \frac{S_i^0 - S_i}{S_i^0}, S_i < S_i^0. \\ \frac{S_i - S_i^0}{S_i^0}, S_i \geq S_i^0. \end{cases} \tag{15}
$$

Finally, the weighted departure degree is given by:

$$
\theta = \sum_{j=1}^{n} (S_j^T \times W_j)(0 \leq \theta \leq 1); \tag{16}
$$

Step 4 Result analysis:

The comment set is put in a consecutive interval. Meanwhile, each comment value is realized by a cloud model. The cloud-generator model can be established as Figure 3 shows. The comment set can be divided into five categories: excellent, good, fair, worse, and poor. For a specific case, assessment results can be output by inputting $1 + \theta$ into the cloud-generator model.
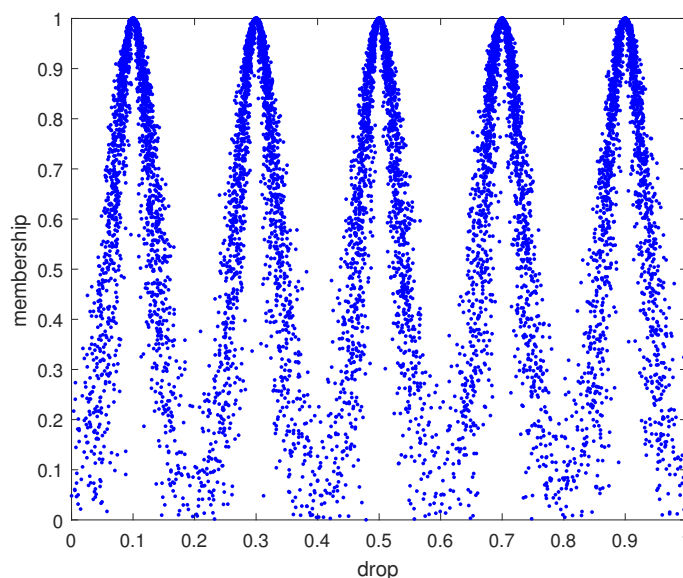

Figure 3. Qualitative evaluation of the cloud-generator model.

3.3. Fuzzy CE Method

The fuzzy CE is based on fuzzy mathematics, which quantitatively expresses the objective attributes of some uncertain things [70–72]. The specific process is as follows:

Step 1 The metric set:

We analyzed the result of target tracking and establish the evaluation metric set U as follows:

$$
U = \{U_1, U_2, \cdots, U_n\}, \tag{17}
$$

where U_i is the *i*th evaluation metric;

Step 2 The evaluation level set:

The evaluation level set is given by $V = \{V_1, V_2, \cdots, V_m\}$, where V_i is the ith grey category. V is the remark collection, which is made up of remarks of the research object;

Step 3 The evaluation matrix:

Starting from a single factor for the evaluation, we determine the degree of membership about evaluation objects to the evaluation level set and make the fuzzy evaluation. Then, combining the single-factor set, a multi-factor evaluation set is given by:

$$R = \begin{pmatrix} R_1 \\ R_2 \\ \cdots \\ R_m \end{pmatrix} = \begin{pmatrix} r_{11} & r_{12} & \cdots & r_{1m} \\ r_{21} & r_{22} & \cdots & r_{2m} \\ \cdots & \cdots & \cdots & \cdots \\ r_{n1} & r_{n2} & \cdots & r_{nm} \end{pmatrix}, \tag{18}$$

where r_{ij} denotes the membership degree of U_i corresponding to V_j;

Step 4 The fuzzy CE value:

$$C = BR = \begin{pmatrix} b_1 & b_2 & \cdots & b_n \end{pmatrix} \begin{pmatrix} r_{11} & r_{12} & \cdots & r_{1m} \\ r_{21} & r_{22} & \cdots & r_{2m} \\ \cdots & \cdots & \cdots & \cdots \\ r_{n1} & r_{n2} & \cdots & r_{nm} \end{pmatrix} = \begin{pmatrix} c_1 & c_2 & \cdots & c_j \end{pmatrix}, \tag{19}$$

where C is the fuzzy CE set and B is the weight of the metric.

According to the principle of the maximum membership degree, the comprehensive value of the PE is obtained; thereby, the corresponding performance levels [70] are calculated.

3.4. Grey Clustering

Grey theory is a useful methodology for incomplete information systems. Grey relational analysis can be used to analyze relationships between the uncertainty and the gray category [73,74]. The main steps of the method are as follows:

Step 1 Triangular whitenization weight functions are established and obtained as follows:

$$f_{1j}^k(d_{ij}) = f_{1j}^k(c_{1j}^k, \infty) = \begin{cases} \frac{d_{ij}}{c_{1j}^k}, d_{ij} \in [0, c_{1j}^k] \\ 1, d_{ij} \in (c_{1j}^k, \infty) \\ 0, d_{ij} \notin [0, \infty) \end{cases}, \tag{20}$$

$$f_{2j}^k(d_{ij}) = f_{2j}^k(-, c_{2j}^k, +) = \begin{cases} \frac{d_{ij}}{c_{2j}^k}, d_{ij} \in [0, c_{2j}^k] \\ \frac{d_{ij} - 2c_{2j}^k}{-c_{2j}^k}, d_{ij} \in (c_{2j}^k, 2c_{2j}^k] \\ 0, d_{ij} \notin [0, 2c_{2j}^k] \end{cases}, \tag{21}$$

$$f_{3j}^k(d_{ij}) = f_{3j}^k(0, c_{3j}^k) = \begin{cases} 1, d_{ij} \in [0, c_{3j}^k] \\ \frac{d_{ij} - 2c_{3j}^k}{-c_{3j}^k}, d_{ij} \in (c_{3j}^k, 2c_{3j}^k] \\ 0, d_{ij} \notin [0, 2c_{3j}^k] \end{cases}, \tag{22}$$

where $d_{ij}(i = 1, 2, \cdots, n; j = 1, 2, \cdots, m)$ is the sample of the ith algorithm about jth evaluation metric and c_j^k is the midpoints of the jth clustering metric belonging to the kth grey category. The type of measure determines the choice of three functions: If the measure is extremely large data (preferably larger), we select $f_{1j}^k(d_{ij})$. If it is

the moderate measure, $f_{2j}^k(d_{ij})$ is selected. If it is extremely small, then $f_{3j}^k(d_{ij})$ is the first choice;

Step 2 The clustering coefficient:

$$\sigma_i^k = \sum_{j=1}^m f_j^k(d_{ij})w_j,$$ (23)

$$\delta_i^k = \frac{\sigma_i^k}{\sum\limits_{k=1}^s \sigma_i^k},$$ (24)

where w_j is the weight of the jth clustering and is determined by the AHP. σ_i^k is the weight cluster coefficient of the ith algorithm about the kth grey category, and δ_i^k is the normalized weight cluster coefficient, respectively [75];

Step 3 The integrated clustering coefficient η_i for each monitoring point with respect to the grey classes k can be calculated with the following equation [76]:

$$\eta_i = \sum_{k=1}^s k \cdot \delta_i^k, (i = 1, 2, \cdots, n);$$ (25)

Step 4 According to the integrated clustering coefficient, the evaluation result is determined. The value range of the integrated clustering coefficient is divided into s intervals of the same length, which are: $[1, 1 + s - 1/s]$, $[1 + s - 1/s, 1 + 2(s - 1)/s]$, $\cdots$, $[s - s - 1/s, s]$. The track algorithm is judged as the kth grey category, when η_i belongs to $[1 + (k-1)(s-1)/s, 1 + k(s-1)/s]$.

4. Rating and Overall Performance

Most simulations are run in the Monte Carlo scenario to describe the characteristics of the performance metrics. In [21], the analysis and assessment of the tracking algorithm were performed both with simulated and real data, where the real data were measured with the Multi-Static Primary Surveillance Radar (MSPSR) L-Band demonstrator, and the metrics were calculated for the performance evaluation such as the mean variance statistic of the NMT, RTF, RMSE, etc. Refs. [25,77] calculated the GOSPA varying values of c and γ using a multiple target tracking example in the MATLAB code. The COCO 2017 validation set and the MOTChallenge (MOT17) dataset were used in terms of the Hausdorff, Wasserstein, and OSPA metrics [78]. In the paper, the metrics mentioned in Section 2 were combined to give a score or a membership value by the aforementioned CE approaches. Three measures were taken together to judge the efficiency of target tracking in the cloud barycenter evaluation (the synthetic measures do not involve simulations). At the same time, the realization of the fuzzy theory and the grey clustering is in the absence of the correctness and the timeliness scenario. In order to demonstrate it more vividly, our simulation was performed by using a graphical user interface (GUI).

4.1. Application of Cloud Theory for Target Tracking

In this section, we discuss the application of cloud theory for target tracking. In this scenario, three categories of the performance metrics are involved, and the last is the accuracy, which is divided into the TE, TPE, and TVE. The judgment matrix was given by experts and displayed in Figure 4, and then, the metric weights were ascertained. W_1 = (0.5436, 0.1634, 0.2970) is the weight of three primary metrics. Then, W_2 = (0.3746, 0.1835, 0.0943, 0.1184, 0.0983, 0.0503, 0.0806) is the weight of the correctness measure; W_3 = (0.0854, 0.5424, 0.2133, 0.1588) is the weight of the timeliness measure; W_4 = (0.1692, 0.4434, 0.3874) is the weight of the accuracy measure. We can draw some conclusions in which the tracking accuracy and correctness are more important than the timeliness in judging the efficiency of target tracking, where the sequential ratio of those is 0.2970, 0.5346, and 0.1634.

According to the evaluation of experts, $S = \{\text{excellent}, \text{good}, \text{fair}, \text{worse}, \text{poor}\}$ was placed in [0,1]. Therefore, Table 3 represents the results of the cloud model of comments.

Table 3. Cloud model of the comments.

Comments	Number Field Interval	Numeral Characteristics
excellent	[1,0.8]	(0.9,0.033)
good	[0.8,0.6]	(0.7,0.033)
fair	[0.6,0.4]	(0.5,0.033)
worse	[0.4,0.2]	(0.3,0.033)
poor	[0.2,0]	(0.1,0.033)

$$
S = \begin{bmatrix}
0.9 & 0.7 & 0.9 & 0.7 & 0.5 & 0.5 & 0.3 & 0.3 & 0.5 & 0.3 & 0.1 & 0.7 & 0.7 & 0.7 \\
0.9 & 0.9 & 0.5 & 0.7 & 0.5 & 0.5 & 0.5 & 0.3 & 0.5 & 0.1 & 0.3 & 0.5 & 0.5 & 0.5 \\
0.9 & 0.5 & 0.7 & 0.5 & 0.5 & 0.7 & 0.3 & 0.5 & 0.9 & 0.5 & 0.5 & 0.7 & 0.7 & 0.7 \\
0.7 & 0.7 & 0.7 & 0.5 & 0.5 & 0.5 & 0.7 & 0.7 & 0.7 & 0.7 & 0.5 & 0.5 & 0.9 & 0.7 \\
0.9 & 0.9 & 0.5 & 0.7 & 0.7 & 0.5 & 0.5 & 0.3 & 0.5 & 0.5 & 0.7 & 0.9 & 0.9 & 0.9 \\
0.7 & 0.5 & 0.9 & 0.7 & 0.5 & 0.1 & 0.9 & 0.1 & 0.9 & 0.9 & 0.3 & 0.5 & 0.9 & 0.7 \\
0.9 & 0.7 & 0.5 & 0.7 & 0.3 & 0.7 & 0.5 & 0.7 & 0.7 & 0.7 & 0.5 & 0.9 & 0.7 & 0.9 \\
0.9 & 0.7 & 0.5 & 0.5 & 0.3 & 0.5 & 0.7 & 0.7 & 0.3 & 0.3 & 0.5 & 0.7 & 0.7 & 0.7 \\
0.9 & 0.5 & 0.7 & 0.7 & 0.3 & 0.3 & 0.5 & 0.5 & 0.3 & 0.3 & 0.9 & 0.9 & 0.9 & 0.9 \\
0.9 & 0.5 & 0.5 & 0.7 & 0.9 & 0.5 & 0.7 & 0.5 & 0.5 & 0.5 & 0.7 & 0.7 & 0.7 & 0.9
\end{bmatrix}
$$

S is the decision matrix, which denotes the integrated cloud gravity center. Combining Equations (13) and (14) and Table 3, the cloud model of the parameter status is calculated and given by Table 4.

Table 4. The cloud model of the parameter status.

Parameter	Expectations	Entropy
C1	0.86	0.33
C2	0.66	0.33
C3	0.64	0.33
C4	0.64	0.33
C5	0.5	0.33
C6	0.48	0.33
C7	0.56	0.33
C8	0.46	0.33
C9	0.68	0.33
C10	0.48	0.33
C11	0.5	0.33
C12	0.7	0.33
C13	0.76	0.33
C14	0.76	0.33

Then, the conclusion can be obtained in the part of "performance evaluation". $W = [0.2022, 0.0990, 0.0509, 0.0639, 0.0530, 0.0271, 0.0435, 0.0140, 0.0886, 0.0349, 0.0260, 0.0277, 0.0725, 0.0633]$. In the ideal state, $E = [0.9, 0.9, 0.9, 0.9, 0.9, 0.9, 0.9, 0.9, 0.9, 0.9, 0.9, 0.9, 0.9, 0.9]$; the integrated evaluation vector of the cloud barycenter is $S^0 = W \times E = (0.1819, 0.0891, 0.0458, 0.0575, 0.0477, 0.0244, 0.0391, 0.0126, 0.0798, 0.0314, 0.0234, 0.0249, 0.0652, 0.0570)$. The actual integrated vector is $S = (W_1, W_2, \cdots, W_{14}) \times (E_{x1}, E_{x2}, \cdots, E_{x14}) = (0.1739, 0.0653, 0.0326, 0.0409, 0.0265, 0.0130, 0.0244, 0.0064, 0.0603, 0.0167, 0.0130, 0.0194, 0.0551, 0.0481)$. S is normalized as $S^T = (S_1^T, S_2^T, \cdots, S_{14}^T) = (-0.0444, -0.2667, -0.2889, -0.2889, -0.4444, -0.4667, -0.3778, -0.4889, -0.2444, -0.4667, -0.4444, -0.2222, -0.1556, -0.1556)$. Finally, the weighted departure degree can be acquired as $\theta = -0.20477$.
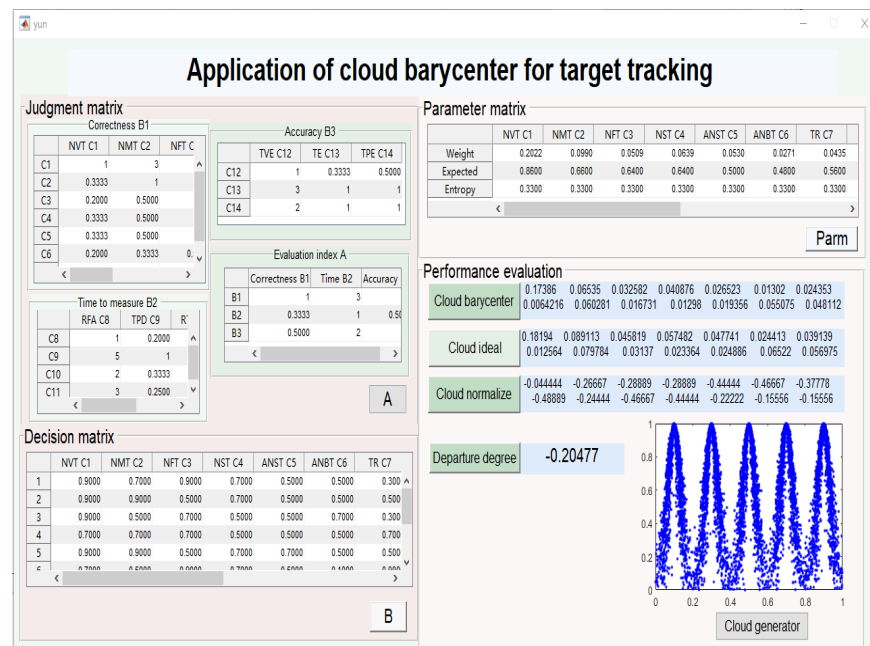
Figure 4. PE of cloud gravity for target tracking.

In the ideal state, θ for PE is zero; however, the actual θ is -0.20477. When $1 + \theta = 0.79523$ is input into the cloud generator model, the cloud drops close to the "good" cloud object. Then, the PE result of target tracking is deemed to be good. We came up with a conjecture that the measure of the accuracy and the correctness are more effective than the timeliness in target tracking. Then, we verified it in the following simulation.

4.2. Application of Fuzzy CE for Target Tracking

Fuzzy CE uses the fuzzy mathematics tool to depict the influence of various metrics, which was applied in the situation without the timeliness measure. Here, $U = \{$ NMT U_1, NVT U_2, NST U_3, NFT U_4, TSE U_5, TBE U_6, TE $U_7 \}$. $B = (0.1621, 0.1284, 0.0773, 0.1120, 0.0836, 0.2000, 0.2365)$ is the weight of the seven metrics. $V = \{V_1, V_2, V_3, V_4\}$ denotes the comment set, corresponding four levels: "excellent", "good", "medium", and "poor". Here, the subordinate degree was determined by the expert assessment. We obtain R, C as follows.

$$
R = \begin{bmatrix} R1 \\ R2 \\ R3 \\ R4 \\ R5 \\ R6 \\ R7 \end{bmatrix} = \begin{bmatrix} 0.2 & 0.4 & 0.4 & 0 \\ 0.3 & 0.4 & 0.3 & 0 \\ 0.1 & 0.4 & 0.5 & 0 \\ 0.2 & 0.4 & 0.4 & 0 \\ 0.1 & 0.5 & 0.4 & 0 \\ 0.2 & 0.3 & 0.5 & 0 \\ 0 & 0.5 & 0.5 & 0 \end{bmatrix}
$$

$$
C = BR = \begin{bmatrix} 0.1621 \\ 0.1284 \\ 0.0773 \\ 0.1120 \\ 0.2000 \\ 0.1922 \\ 0.2365 \end{bmatrix}^T \begin{bmatrix} 0.2 & 0.4 & 0.4 & 0 \\ 0.3 & 0.4 & 0.3 & 0 \\ 0.1 & 0.4 & 0.5 & 0 \\ 0.2 & 0.4 & 0.4 & 0 \\ 0.1 & 0.5 & 0.4 & 0 \\ 0.2 & 0.3 & 0.5 & 0 \\ 0 & 0.5 & 0.5 & 0 \end{bmatrix} = \begin{bmatrix} 0.1494 & 0.4120 & 0.4386 & 0 \end{bmatrix}
$$

The simulation result is given in Figure 5. According to the metric information, we could obtain the fuzzy CE of all the evaluation information. The fuzzy CE set was calculated and performed the "medium" in terms of the maximum membership degree on the basis of

the principle. To combine the analysis above, the total score of the PE for target tracking can be given by: D = 0.1494 × 90 + 0.4120 × 80 + 0.4386 × 70 = 77.108, and the result showed that the performance for target tracking performed medium.
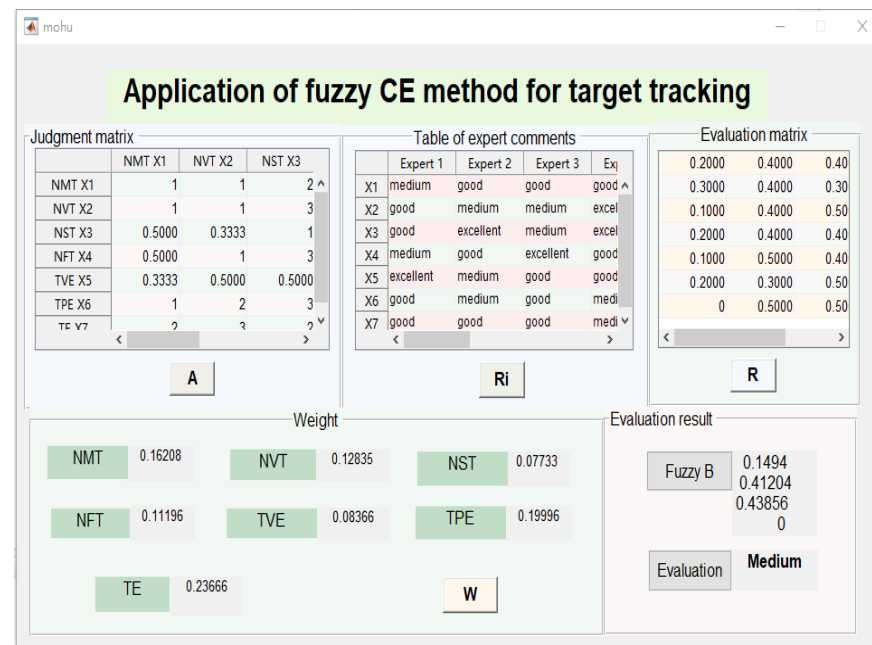

Figure 5. PE of target tracking based on fuzzy CE.

4.3. PE in Target Tracking Using Grey Clustering

In this section, representative metrics of the timeliness and the accuracy measures were chosen, which included the TPE, TL, TVE, TPD, and RFA. w = { 0.29783, 0.088788, 0.15777, 0.29783, 0.15777 } is the weight of these. Four grey categories were established, where K = {1, 2, 3, 4} denotes "excellent", "good", "medium", and "poor". Based on the scenario [79], the following tracking methods were used for the performance evaluation where results were available for 200 Monte Carlo runs:

(1) Pseudo-observation (PS) [80];
(2) Projection (PRO) [81,82];
(3) Karush–Kuhn–Tucker (KKT) [83];
(4) Karush–Kuhn–Tucker–Kalman filter (KKT_KF) [79];
(5) Unconstrained Kalman filter (UKF) [84];
(6) Trajectory function of time (T-FoT) [85,86].

Performance metrics were now found based on the given tracking results, which are given in Table 5. According to Table 5, we can determine the whitenization weight functions in Table 6. Taking the first metric (TPE) as an example, the four whitenization weight functions are:

$$f_1^1(6,\infty) = \begin{cases} \frac{d_{ij}}{6}, d_{ij} \in [0,6] \\ 1, d_{ij} \in (6,\infty) \\ 0, d_{ij} \notin [0,\infty) \end{cases}$$

$$f_1^2(-,4.5,+) = \begin{cases} \frac{d_{ij}}{c_{2j}^k}, d_{ij} \in [0,4.5] \\ \frac{d_{ij}-9}{-4.5}, d_{ij} \in (4.5,9] \\ 0, d_{ij} \notin [0,9] \end{cases}$$

$$f_1^3(-,3,+) = \begin{cases} \frac{d_{ij}}{3}, d_{ij} \in [0,3] \\ \frac{d_{ij}-6}{-3}, d_{ij} \in (3,6] \\ 0, d_{ij} \notin [0,6] \end{cases}$$

$$f_1^4(0,1) = \begin{cases} 1, d_{ij} \in [0,1] \\ \frac{d_{ij}-2}{-1}, d_{ij} \in (1,2] \\ 0, d_{ij} \notin [0,2] \end{cases}$$

Table 5. Metric statistics of each tracking algorithm.

Arithmetic	TPE	TL	TVE	TPD	RFA
PS	5.2545	1	2.1056	0.958	0.00042
PRO	4.4997	1	2.8736	0.973	0.0085
KKT	4.4997	2	3.0789	0.965	0.014
KKT_KF	3.9048	1	2.505	0.966	0.0007
UKF	6.3551	1	2.505	0.968	0.001
T-FoT	5.5309	2	3.0789	0.961	0.005

Table 6. Whitenization weight functions of four grey categories.

Excellent Grey Category	Good Grey Category
$f_1^1(c_1^1, \infty) = f_1^1(6, \infty)$	$f_1^2(-, c_1^2, +) = f_1^2(-, 4.5, +)$
$f_2^1(c_2^1, \infty) = f_2^1(1.5, \infty)$	$f_2^2(-, c_2^2, +) = f_2^2(-, 1.25, +)$
$f_3^1(c_3^1, \infty) = f_3^1(3, \infty)$	$f_3^2(-, c_3^2, +) = f_3^2(-, 2.6, +)$
$f_4^1(c_4^1, \infty) = f_4^1(0.965, \infty)$	$f_4^2(-, c_4^2, +) = f_4^2(-, 0.95, +)$
$f_5^1(c_5^1, \infty) = f_5^1(0.01, \infty)$	$f_5^2(-, c_5^2, +) = f_5^2(-, 0.001, +)$

Medium Grey Category	Poor Grey Category
$f_1^3(-, c_1^3, +) = f_1^3(-, 3, +)$	$f_1^4(0, c_1^4) = f_1^4(0, 1)$
$f_2^3(-, c_2^3, +) = f_2^3(-, 1, +)$	$f_2^4(0, c_2^4) = f_2^4(0, 0.5)$
$f_3^3(-, c_3^3, +) = f_3^3(-, 2, +)$	$f_3^4(0, c_3^4) = f_3^4(0, 1)$
$f_4^3(-, c_4^3, +) = f_4^3(-, 0.5, +)$	$f_4^4(0, c_4^4) = f_4^4(0, 0.5)$
$f_5^3(-, c_5^3, +) = f_5^3(-, 0.0005, +)$	$f_5^4(0, c_5^4) = f_5^4(0, 0.0001)$

$f_{1j}^k(c_{1j}^k, \infty)$ was applied in the excellent grey category, which is an upper measure. The good grey category and medium grey category were selected by $f_{2j}^k(-, c_{2j}^k, +)$. $f_{3j}^k(0, c_{3j}^k)$ is a whitenization weight function of lower measure, which was used in the poor grey category next. From the weight of the evaluation metric and the grey clustering, the weighted clustering coefficient matrix σ can be determined.

$$\sigma = (\sigma_i^k) = \begin{vmatrix} 0.7331 & 0.8083 & 0.4698 & 0 \\ 0.8656 & 0.8006 & 0.3427 & 0 \\ 0.9255 & 0.7552 & 0.2425 & 0 \\ 0.6936 & 0.8847 & 0.5296 & 0 \\ 0.8024 & 0.8481 & 0.2258 & 0 \\ 0.8966 & 0.6882 & 0.1425 & 0 \end{vmatrix}$$

Furthermore, the normalized clustering coefficient matrix can be obtained:

$$\delta_i^k = \begin{vmatrix} 0.3645 & 0.4019 & 0.2336 & 0 \\ 0.4309 & 0.3985 & 0.1706 & 0 \\ 0.4813 & 0.3927 & 0.1261 & 0 \\ 0.3290 & 0.4197 & 0.2513 & 0 \\ 0.4277 & 0.4520 & 0.1203 & 0 \\ 0.5191 & 0.3984 & 0.0825 & 0 \end{vmatrix}$$

The simulation result is shown in Figure 6. The integrated clustering coefficients of the six algorithms were calculated. For example, the integrated clustering coefficient of the PS in the first algorithm was $\eta_1 = \sum\limits_{k=1}^{4} k \cdot \delta_i^k = 1 \times 0.3645 + 2 \times 0.4019 + 3 \times 0.2336$

+ 4×0 = 1.8691, and the other coefficients were η_2 = 1.7397 (PRO), η_3 = 1.6448 (KKT), η_4 = 1.9222 (KKT_KF), η_5 = 1.6927 (UKF), η_6 = 1.5634 (T-FoT), respectively. The value range of the integrated clustering coefficient was divided into four intervals of the same length. According to **Step 4** in Section 3.4, $\eta_2, \eta_3, \eta_5, \eta_6 \in [1, 1 + 3/4]$, $\eta_1, \eta_4 \in [1 + 3/4, 1 + 6/4]$, and the evaluated result was that the methods of the PRO, KKT, UKF, and T-FoT had "excellent" performance and the methods of the PS and KKT_KF "good" performance. It is not difficult to find that the CE method for target tracking achieved a satisfactory result by comparing the above algorithms.
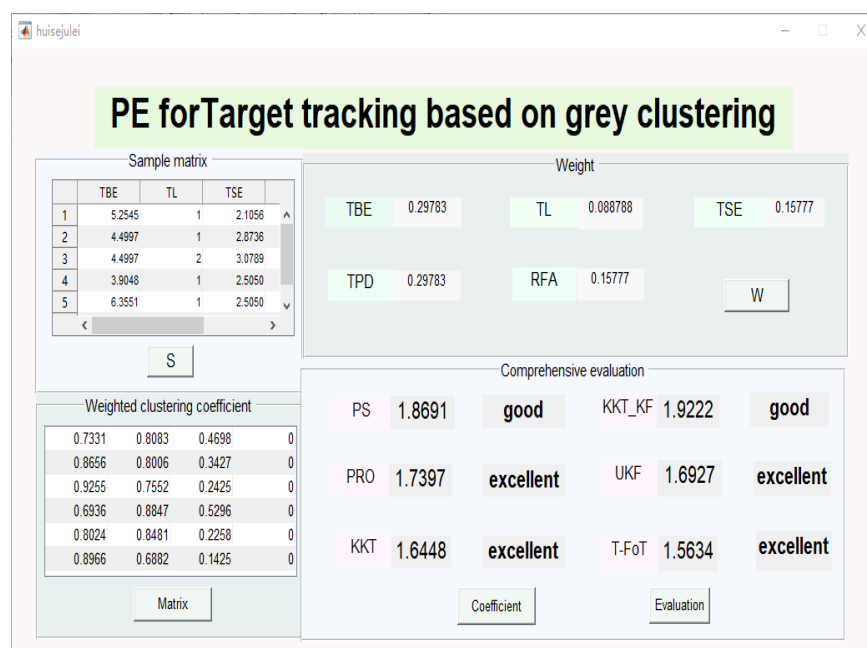

Figure 6. PE for target tracking based on grey clustering.

The simulation showed that the three measures contained sufficient information for evaluation by the comparative analysis. However, there were slight differences in the scenarios in which the one lacks correctness and the other one lacks timeliness. The results of grey clustering were consistent with cloud theory, and the fuzzy CE showed poor results. In other words, for this case, it can be concluded that the timeliness measure was less informative than the correctness and accuracy measures.

So far, on the one hand, PE has a tendency to focus on improving OSPA. On the other hand, PE metrics have been redefined using different association algorithms. A more effective PE method needs to be explored to enhance the algorithm efficiency.

5. Conclusions and Remaining Challenges

This paper reviewed PE metrics for target tracking and some CE approaches. The measures were divided into three categories and described for each category. Finally, the simulation result showed that a combination of metrics from different classes can provide a criterion for PE in target tracking.

Instead of estimating the discrete-time state of the target, it is actually of greater interest to estimate the continuous-time trajectory of the target, which contains more information than the discrete time series of point estimates. The T-FoT framework [85,87,88] is promising and powerful as it completely describes the movement pattern/dynamic behavior of the targets over time and enables the use of many curve-fitting tools such as the Gauss process and neural networks, in addition to various flexible parametric regression analysis methods. However, based on this framework, the output of the estimator or tracker is a spatio-temporal trajectory function, or functions in the case of multiple targets. How to evaluate

the quality of the trajectory functions/curves in the presence of missing detection and false alarms remains an open problem.

Author Contributions: This paper is a collaborative work by all the authors. Y.S.: methodology, investigation, writing—original draft, formal analysis, and software. Z.H.: resources, formal analysis, and validation. T.L.: conceptualization, methodology, and writing—review and editing. H.F.: writing—review and editing. All authors have read and agreed to the published version of the manuscript.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data will be made available upon request.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

CE	Comprehensive evaluation
PE	Performance evaluation
TPE	Track position error
TE	Trajectory error
TVE	Track velocity error
NVT	Number of valid tracks
NMT	Number of missed targets
NFT	Number of false tracks
NST	Number of spurious tracks
ANST	Average number of swaps in tracks
ANBT	Average number of broken tracks
TR	Tracks redundancy
RFA	Rate of false alarms
TPD	Track probability of detection
RTF	Rate of track fragmentation
TL	Track latency
TET	Total execution time
OSPA	Optimal subpattern assignment
AHP	Analytic hierarchy process
SIAP	Single integrated air picture
MOPT	Multiple-object-tracking precision
MOTA	Multiple-object-tracking accuracy
GUI	Graphical user interface
PS	Pseudo-observation
PRO	Projection
KKT	Karush–Kuhn–Tucker
KKT_KF	Karush–Kuhn–Tucker–Kalman filter
UKF	Unconstrained Kalman filter
T-FoT	Trajectory function of time

References

1. Chong, C.Y. Tracking and data fusion: A handbook of algorithms (bar-shalom, y. et al; 2011)[bookshelf]. *IEEE Control. Syst. Mag.* **2012**, *32*, 114–116.
2. Blasch, E. Target Tracking Toolbox lecture notes and software to support EE716. Ph.D. Thesis, Wright State University, Dayton, OH, USA, 2001.
3. Available online: https://www.pdx.edu/biomedical-signal-processing-lab/signal-point-kalman-filters-and-the-rebel-toolkit (accessed on 10 September 2021).

4. Wan, E.A.; Van Der Merwe, R.; Haykin, S. The unscented Kalman filter. *Kalman Filter. Neural Net.* **2001**, *5*, 221–280.
5. Paul, A.S. Sigma-point Kalman Smoothing: Algorithms and Analysis with Applications to Indoor Tracking. Ph.D. Thesis, Oregon Health & Science University, Portland, OR, USA, 2010.
6. Straka, O.; Flídr, M.; Duník, J.; Ŝimandl, M. A software framework and tool for nonlinear state estimation. *IFAC Proc. Vol.* **2009**, *42*, 510–515. [CrossRef]
7. Straka, O.; Flídr, M.; Duník, J.; Simandl, M.; Blasch, E. Nonlinear estimation framework in target tracking. In Proceedings of the 2010 13th International Conference on Information Fusion, Edinburgh, UK, 26–29 July 2010.
8. Blasch, E.P.; Straka, O.; Duník, J.; Šimandl, M. Multitarget tracking performance analysis using the non-credibility index in the nonlinear estimation framework (NEF) toolbox. In Proceedings of the IEEE 2010 National Aerospace & Electronics Conference, Dayton, OH, USA, 14–16 July 2010; pp. 107–115.
9. Crouse, D.F. The tracker component library: Free routines for rapid prototyping. *IEEE Aerosp. Electron. Syst. Mag.* **2017**, *32*, 18–27. [CrossRef]
10. Thomas, P.A.; Barr, J.; Balaji, B.; White, K. An open source framework for tracking and state estimation ('Stone Soup'). In *Signal Processing, Sensor/Information Fusion, and Target Recognition XXVI*; International Society for Optics and Photonics: Bellingham, WA, USA, 2017; Volume 10200, p. 1020008.
11. Last, D.; Thomas, P.; Hiscocks, S.; Barr, J.; Kirkland, D.; Rashid, M.; Li, S.B.; Vladimirov, L. Stone Soup: Announcement of beta release of an open-source framework for tracking and state estimation. In *Signal Processing, Sensor/Information Fusion, and Target Recognition XXVIII*; International Society for Optics and Photonics: Bellingham, WA, USA, 2019; Volume 11018, p. 1101807.
12. Costa, P.C.; Laskey, K.B.; Blasch, E.; Jousselme, A.L. Towards unbiased evaluation of uncertainty reasoning: The URREF ontology. In Proceedings of the 2012 15th International Conference on Information Fusion, Singapore, 9–12 July 2012; pp. 2301–2308.
13. Xu, Z. Performance evaluation of business administration training room in application-oriented universities. In Proceedings of the 2020 2nd International Conference on Computer Science Communication and Network Security (CSCNS2020), Sanya, China, 22–23 December 2021; Volume 336, p. 09015.
14. Zhang, G.; Hui, G.; Zhang, G.; Hu, Y.; Zhao, Z. A Novel Comprehensive Evaluation Method of Forest State Based on Unit Circle. *Forests* **2019**, *10*, 5. [CrossRef]
15. Li, X. Application Research on the Model of the Performance Evaluation of Enterprise Informatization. *J. Inf.* **2008**, *12*, 15–17.
16. Bar-Shalom, Y.; Li, X.R.; Kirubarajan, T. *Estimation with Applications to Tracking and Navigation: Theory, Algorithms and Software*; John Wiley & Sons: Hoboken, NJ, USA, 2004; p. 584.
17. Popp, R.L.; Kirubarajan, T.; Pattipati, K.R. Survey of assignment techniques for multitarget tracking. *Multitarg.-Multisens. Tracking Appl. Adv.* **2000**, *3*, 77–159.
18. Colegrove, S.B.; Cheung, B.; Davey, S.J. Tracking system performance assessment. In Proceedings of the Sixth International Conference of Information Fusion, Cairns, Australia, 8–10 July 2003; Volume 2, pp. 926–933.
19. Sheng, X.; Chen, Y.; Guo, L.; Yin, J.; Han, X. Multitarget Tracking Algorithm Using Multiple GMPHD Filter Data Fusion for Sonar Networks. *Sensors* **2018**, *18*, 3193. [CrossRef]
20. Ristic, B.; Vo, B.N.; Clark, D.; Vo, B.T. A Metric for Performance Evaluation of Multi-Target Tracking Algorithms. *IEEE Trans. Signal Process.* **2011**, *59*, 3452–3457. [CrossRef]
21. Kulmon, P.; Stukovska, P. Assessing Multiple-Target Tracking Performance Of GNN Association Algorithm. In Proceedings of the 2018 19th International Radar Symposium (IRS), Bonn, Germany, 20–22 June 2018.
22. Evirgen, E.A. Multi sensor track fusion performance metrics. In Proceedings of the 2016 24th Signal Processing and Communication Application Conference (SIU), Zonguldak, Turkey, 16–19 May 2016; pp. 97–100.
23. Gorji, A.A.; Tharmarasa, R.; Kirubarajan, T. Performance measures for multiple target tracking problems. In Proceedings of the 14th International Conference on Information Fusion, Chicago, IL, USA, 5–8 July 2011.
24. de Villiers, J.P.; Focke, R.W.; Pavlin, G.; Jousselme, A.L.; Dragos, V.; Laskey, K.B.; Costa, P.; Blasch, E. Evaluation metrics for the practical application of URREF ontology: An illustration on data criteria. In Proceedings of the 2017 20th International Conference on Information Fusion (Fusion), Xi'an, China, 10–13 July 2017.
25. García-Fernández, Á.F.; Rahmathullah, A.S.; Svensson, L. A time-weighted metric for sets of trajectories to assess multi-object tracking algorithms. *arXiv* **2021**, arXiv:2110.13444.
26. Rothrock, R.L.; Drummond, O.E. Performance metrics for multiple-sensor multiple-target tracking. In Proceedings of the SPIE Conference on Signal and Data Processing of Small Targets, San Diego, CA, USA, 30 July–2 August 2001; Volume 4048, pp. 521–531.
27. Drummond, O.; Fridling, B. Ambiguities in evaluating performance of multiple target tracking algorithms. *Proc. Spie Int. Soc. Opt. Eng.* **1992**, *1698*, 326–337.
28. Li, X.; Zhao, Z. Evaluation of estimation algorithms part I: Incomprehensive measures of performance. *Aerosp. Electron. Syst. IEEE Trans.* **2006**, *42*, 1340–1358. [CrossRef]
29. Blackman, S.; Popoli, R. *Design and Analysis of Modern Tracking Systems*; Artech House Publishers: London, UK, 1999; p. 1015.
30. Bernardin, K.; Stiefelhagen, R. Evaluating multiple object tracking performance: The clear mot metrics. *Eurasip J. Image Video Process.* **2008**, *2008*, 246309. [CrossRef]
31. Milan, A.; Schindler, K.; Roth, S. Challenges of ground truth evaluation of multitarget tracking. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, Portland, OR, USA, 23–28 June 2013; pp. 735–742.

32. Pelletier, M.; Sivagnanam, S.; Blasch, E.P. A track scoring MOP for perimeter surveillance radar evaluation. In Proceedings of the 2012 15th International Conference on Information Fusion, Singapore, 9–12 July 2012; pp. 2028–2034.

33. Blasch, E.P.; Straka, O.; Yang, C.; Qiu, D.; Šimandl, M.; Ajgl, J. Distributed tracking fidelity-metric performance analysis using confusion matrices. In Proceedings of the 2012 15th International Conference on Information Fusion, Singapore, 9–12 July 2012.

34. Evers, C.; Löllmann, H.W.; Mellmann, H.; Schmidt, A.; Barfuss, H.; Naylor, P.A.; Kellermann, W. LOCATA challenge-overview of evaluation measures. *Trans. Signal Process.* **2008**, *56*, 3447–3457.

35. Mori, S.; Chang, K.C.; Chong, C.Y.; Dunn, K.P. Tracking performance evaluation-prediction of track purity. In *Signal and Data Processing of Small Targets 1989*; International Society for Optics and Photonics: Bellingham, WA, USA, 1989; Volume 1096, pp. 215–223.

36. Blasch, E.P.; Valin, P. Track purity and current assignment ratio for target tracking and identification evaluation. In Proceedings of the 14th International Conference on Information Fusion, Chicago, IL, USA, 5–8 July 2011.

37. Blasch, E. Fusion Evaluation Tutorial. In Proceedings of the International Conference on Information Fusion, Stockholm, Sweden, 28 June–1 July 2004.

38. Coraluppi, S.; Grimmett, D.; de Theije, P. Benchmark Evaluation of Multistatic Trackers. In Proceedings of the 2006 9th International Conference on Information Fusion, Florence, Italy, 10–13 July 2006.

39. Grimmett, D.; Coraluppi, S.; Cour, B.; Hempel, C.G.; Lang, T.; Theije, P.; Willett, P. MSTWG multistatic tracker evaluation using simulated scenario data sets. In Proceedings of the 2008 11th International Conference on Information Fusion, Cologne, Germany, 30 June–3 July 2008.

40. Guerriero, M.; Svensson, L.; Svensson, D.; Willett, P. Shooting two birds with two bullets: How to find Minimum Mean OSPA estimates. In Proceedings of the 2010 13th International Conference on Information Fusion, Edinburgh, UK, 26–29 July 2010.

41. Chang, F.; Chen, Z.; Wang, W.; Wang, L. The Hausdorff distance template matching algorithm based on Kalman filter for target tracking. In Proceedings of the 2009 IEEE International Conference on Automation and Logistics, Shenyang, China, 5–7 August 2009; pp. 836–840.

42. Da, K.; Li, T.; Zhu, Y.; Fu, Q. A Computationally Efficient Approach for Distributed Sensor Localization and Multitarget Tracking. *IEEE Commun. Lett.* **2020**, *24*, 335–338. [CrossRef]

43. Hoffman, J.R.; Mahler, R. Multitarget miss distance and its applications. In Proceedings of the Fifth International Conference on Information Fusion. FUSION 2002, Annapolis, MD, USA, 8–11 July 2002; Volume 1, pp. 149–155.

44. Schuhmacher, D.; Vo, B.T.; Vo, B.N. A Consistent Metric for Performance Evaluation of Multi-Object Filters. *IEEE Trans. Signal Process.* **2008**, *56*, 3447–3457. [CrossRef]

45. Ristic, B.; Vo, B.N.; Clark, D. Performance evaluation of multitarget tracking using the OSPA metric. In Proceedings of the 2010 13th International Conference on Information Fusion, Edinburgh, UK, 26–29 July 2010.

46. Nagappa, S.; Clark, D.E.; Mahler, R. Incorporating track uncertainty into the OSPA metric. In Proceedings of the 14th International Conference on Information Fusion, Chicago, IL, USA, 5–8 July 2011.

47. El-Fallah, A.I.; Ravichandran, R.B.; Mehra, R.K.; Hoffman, J.R.; Alford, M.G. Scientific performance evaluation for distributed sensor management and adaptive data fusion. In *Signal Processing, Sensor Fusion, and Target Recognition X*; International Society for Optics and Photonics: Bellingham, WA, USA, 2001; Volume 4380, pp. 328–338.

48. Hoffman, J.R.; Mahler, R.; Zajic, T. User-defined information and scientific performance evaluation. In *Signal Processing, Sensor Fusion, and Target Recognition X*; International Society for Optics and Photonics: Bellingham, WA, USA, 2001; Volume 4380, pp. 300–311.

49. Mahler, R. Scientific performance metrics for data fusion: New results. In *Signal Processing, Sensor Fusion, and Target Recognition IX*; International Society for Optics and Photonics: Bellingham, WA, USA, 2000; Volume 4052, pp. 172–182.

50. Villani, C. *Optimal Transport. Old and New*; Springer: Berlin/Heidelberg, Germany, 2009; p. 976.

51. Hoffman, J.R.; Mahler, R. Multitarget Miss Distance via Optimal Assignment. *IEEE Trans. Syst. Man Cybern. Part Syst. Hum.* **2004**, *34*, 327–336. [CrossRef]

52. García-Femández, Á.F.; Svensson, L. Spooky effect in optimal OSPA estimation and how GOSPA solves it. In Proceedings of the 2019 22th International Conference on Information Fusion (FUSION), Ottawa, ON, Canada, 2–5 July 2019.

53. Vu, T.; Evans, R. A new performance metric for multiple target tracking based on optimal subpattern assignment. In Proceedings of the 17th International Conference on Information Fusion (FUSION), Salamanca, Spain, 7–10 July 2014, 2014.

54. Mei, L.; Li, H.; Zhou, Y.; Li, D.; Long, W.; Xing, F. Output-Only Damage Detection of Shear Building Structures Using an Autoregressive Model-Enhanced Optimal Subpattern Assignment Metric. *Sensors* **2020**, *20*, 2050. [CrossRef]

55. Lian, F.; Zhang, G.H.; Duan, Z.S.; Han, C.Z. Multi-Target Joint Detection and Estimation Error Bound for the Sensor with Clutter and Missed Detection. *Sensors* **2016**, *16*, 169. [CrossRef] [PubMed]

56. Li, W.; Han, C. Dual Sensor Control Scheme for Multi-Target Tracking. *Sensors* **2018**, *18*, 1653. [CrossRef]

57. Schubert, R.; Klöden, H.; Wanielik, G.; Kälberer, S. Performance evaluation of Multiple Target Tracking in the absence of reference data. In Proceedings of the 2010 13th International Conference on Information Fusion, Edinburgh, UK, 26–29 July 2010.

58. Rahmathullah, A.S.; García-Fernández, Á.F.; Svensson, L. Generalized optimal sub-pattern assignment metric. In Proceedings of the 2017 20th International Conference on Information Fusion (Fusion), Xi'an, China, 10–13 July 2017.

59. Xia, Y.; Granstrcom, K.; Svensson, L.; García-Fernández, A.F. Performance evaluation of multi-bernoulli conjugate priors for multitarget filtering. In Proceedings of the 2017 20th International Conference on Information Fusion (Fusion), Xi'an, China, 10–13 July 2017.

60. Beard, M.; Vo, B.T.; Vo, B.N. OSPA$^{(2)}$: Using the OSPA metric to evaluate multitarget tracking performance. In Proceedings of the 2017 International Conference on Control, Automation and Information Sciences (ICCAIS), Chiang Mai, Thailand , 31 October–1 November 2017; pp. 86–91.

61. Beard, M.; Ba, T.V.; Vo, B.N. A Solution for Large-Scale Multi-Object Tracking. *IEEE Trans. Signal Process.* **2020**, *68*, 2754–2769. [CrossRef]

62. Votruba, P.; Nisley, R.; Rothrock, R.; Zombro, B. *Single Integrated Air Picture (SIAP) Metrics Implementation*; Technical Report; Single Integrated Air Picture System Engineering Task Force: Arlington, VA, USA, 2001.

63. Available online: https://stonesoup.readthedocs.io/en/v0.1b7/auto_examples/Metrics.html (accessed on 16 September 2021).

64. Shang, J.; Vargas, L. New Concepts and Applications of AHP in the Internet Era. *J. Multi-Criteria Decis. Anal.* **2012**, *19*, 1–2. [CrossRef]

65. Cho, J.; Lee, J. Development of a new technology product evaluation model for assessing commercialization opportunities using Delphi method and fuzzy AHP approach. *Expert Syst. Appl.* **2013**, *40*, 5314–5330. [CrossRef]

66. Song, W.; Wen, W.; Guo, Q.; Chen, H.; Zhao, J. Performance evaluation of sensor information fusion system based on cloud theory and fuzzy pattern recognition. In Proceedings of the 2020 IEEE International Conference on Information Technology, Big Data and Artificial Intelligence (ICIBA), Chongqing, China, 6–8 November 2020; Volume 1, pp. 299–303.

67. Ma, L.Y.; Shi, Q.S. Method of Safety Assessment about the Electric Power Supply Company Based on Cloud Gravity Center Theory. *Adv. Mater. Res.* **2012**, *354*, 1149–1156. [CrossRef]

68. Zhang, C.; Shi, Q.; Liu, T.L. Air Base Damage Evaluation Using Cloud Barycenter Evaluation Method. *Adv. Mater. Res.* **2012**, *430*, 803–807. [CrossRef]

69. Gu, H.; Cheng, Z.; Quan, S. Equipment maintenance support capability evaluation using cloud barycenter evaluation method. *Telkomnika Indones. J. Electr. Eng.* **2013**, *11*, 599–606. [CrossRef]

70. Liu, H.; Li, B.; Sun, Y.; Dou, X.; Zhang, Y.; Fan, X. Safety Evaluation of Large-size Transportation Bridges Based on Combination Weighting Fuzzy Comprehensive Evaluation Method. *Iop Conf. Ser. Earth Environ. Sci.* **2021**, *787*, 012194. [CrossRef]

71. Zhang, L.; Pan, Z. Fuzzy Comprehensive Evaluation Based on Measure of Medium Truth Scale. In Proceedings of the 2009 International Conference on Artificial Intelligence and Computational Intelligence, Shanghai, China, 7–8 November 2009; Volume 2, pp. 83–87.

72. Wang, J.; Zhang, Y.; Wang, Y.; Gu, L. Assessment of Building Energy Efficiency Standards Based on Fuzzy Evaluation Algorithm. *Eng. Sustain.* **2019**, *173*, 1–14. [CrossRef]

73. Delgado, A.; Cuadra, D.; Simon, K.; Bonilla, K.; Lee, E. Evaluation of Water Quality in the Lower Huallaga River Watershed using the Grey Clustering Analysis Method. *Int. J. Adv. Comput. Sci. Appl.* **2021**, *12*. [CrossRef]

74. Delgado, A.; Fernandez, A.; Chirinos, B.; Barboza, G.; Lee, E. Impact of the Mining Activity on the Water Quality in Peru Applying the Fuzzy Logic with the Grey Clustering Method. *Int. J. Adv. Comput. Sci. Appl.* **2021**, *12*. [CrossRef]

75. Dang, Y.G.; Liu, S.F.; Liu, B. Study on the Integrated Grey Clustering Method under the Clustering Coefficient with Non-Distinguished Difference. *Chin. J. Manag. Sci.* **2005**, *13*, 69–73.

76. Jiskani, I.M.; Han, S.; Rehman, A.U.; Shahani, N.M.; Brohi, M.A. An Integrated Entropy Weight and Grey Clustering Method-Based Evaluation to Improve Safety in Mines. *Min. Metall. Explor.* **2021**, *38*, 1773–1787. [CrossRef]

77. Rahmathullah, A.S.; García-Fernández, Á.F.; Svensson, L. A metric on the space of finite sets of trajectories for evaluation of multitarget-tracking algorithms. *IEEE Trans. Signal Process.* **2016**, *68*, 3917–3908.

78. Rezatofighi, H.; Nguyen, T.; Vo, B.N.; Vo, B.T.; Reid, I. How trustworthy are the existing performance evaluations for basic vision tasks? *arXiv* **2020**, arXiv:2008.03533.

79. Zhou, J.; Li, T.; Wang, X. State Estimation with Linear Equality Constraints Based on Trajectory Function of Time and Karush-Kuhn-Tucker Conditions. In Proceedings of the 2021 International Conference on Control, Automation and Information Sciences (ICCAIS), Xi'an, China, 14–17 October 2021; pp. 438–443.

80. Tahk, M.; Speyer, J.L. Target tracking problems subject to kinematic constraints. *IEEE Trans. Autom. Control.* **1990**, *35*, 324–326. [CrossRef]

81. Ko, S.; Bitmead, R.R. State estimation for linear systems with state equality constraints. *Automatica* **2007**, *43*, 1363–1368. [CrossRef]

82. Simon, D.; Chia, T.L. Kalman filtering with state equality constraints. *IEEE Trans. Aerosp. Electron. Syst.* **2002**, *38*, 128–136. [CrossRef]

83. Boyd, S.; Boyd, S.P.; Vandenberghe, L. *Convex Optimization*; Cambridge University Press: Cambridge, UK, 2004.

84. Xu, L.; Li, X.R.; Duan, Z.; Lan, J. Modeling and state estimation for dynamic systems with linear equality constraints. *IEEE Trans. Signal Process.* **2013**, *61*, 2927–2939. [CrossRef]

85. Li, T.; Chen, H.; Sun, S.; Corchado, J.M. Joint Smoothing and Tracking Based on Continuous-Time Target Trajectory Function Fitting. *IEEE Trans. Autom. Sci. Eng.* **2019**, *16*, 1476–1483. [CrossRef]

86. Zhou, J.; Li, T.; Wang, X.; Zheng, L. Target Tracking With Equality/Inequality Constraints Based on Trajectory Function of Time. *IEEE Signal Process. Lett.* **2021**, *28*, 1330–1334. [CrossRef]

87. Li, T. Single-Road-Constrained Positioning Based on Deterministic Trajectory Geometry. *IEEE Comm. Lett.* **2019**, *23*, 80–83. [CrossRef]
88. Li, T.; Fan, H. From Target Tracking to Targeting Track: A Data-Driven Approach to Non-cooperative Target Detection and Tracking. *arXiv* **2021**, arXiv:2104.11122.