Multi-Sensor Adaptive Weighted Data Fusion Based on Biased Estimation

Mingwei Qiu and Bo Liu *

School of Information and Intelligent Science and Technology, Hunan Agricultural University, Changsha 410127, China; wf2511574305@163.com

* Correspondence: l-tb@163.com; Tel.: +86-131-1731-9830

Abstract: In order to avoid the loss of optimality of the optimal weighting factor in some cases and to further reduce the estimation error of an unbiased estimator, a multi-sensor adaptive weighted data fusion algorithm based on biased estimation is proposed. First, it is proven that an unbiased estimator can further optimize estimation error, and the reasons for the loss of optimality of the optimal weighting factor are analyzed. Second, the method of constructing a biased estimation value by using an unbiased estimation value and calculating the optimal weighting factor by using estimation error is proposed. Finally, the performance of least squares estimation data fusion, batch estimation data fusion, and biased estimation data fusion is compared through simulation tests, and test results show that biased estimation data fusion has a greater advantage in accuracy, stability, and noise resistance.

Keywords: adaptive weighted data fusion; unbiased estimation; biased estimation

1. Introduction

Data fusion can utilize the information of multiple different data sources to complement each other, so the application of data fusion algorithms can suppress negative environmental factors and get more accurate results [1,2]. Data fusion algorithms in combination with various algorithms can achieve better results [3,4], so they have been widely used in various fields [5,6]. However, data fusion can only be applied to linear systems where all noises are independent of each other; if it is applied to a nonlinear system, it needs to be combined with the extended Kalman filter or the unscented Kalman filter [7]. If noises are correlated, they need to be decoupled by linear changes [8].

It is necessary to first process the original data, which includes removing abnormal data, filtering system deviations, and lossless compression of data. Abnormal data is removed using the Grubbs criterion. Filtering bias uses bias estimation algorithms [9,10], and in distributed fusion, if data need to be restored in the data center, lossless compression is required [11].

Before conducting multi-sensor data fusion, each sensor can also use various algorithms to estimate the true value (called the estimation value), which can improve the fusion accuracy because the estimation value has higher accuracy than the measurement value.

When the measured physical quantity does not change over time, if there is sufficient understanding of the system, linear minimum mean square error estimation is the best method. This algorithm can find a linear function that minimizes mean square error through a single measurement value [12–14]. Otherwise, it is best to use least squares estimation [15,16] and batch estimation [17]. When the measured physical quantity changes over time, it can only be suitable to use Kalman filtering [18,19].
Adaptive weighted data fusion is a widely used multi-sensor data fusion algorithm with a unique optimal weighting factor that enables the fusion estimation value to have a lower estimation variance than the measurement variance of all sensors, but it requires prior knowledge of the measurement variance to run [20]. However, its combination with a partially unbiased estimation algorithm (hereinafter referred to as unbiased estimation data fusion) not only further reduces the estimation variance but also removes the need for prior knowledge.

Unbiased estimation data fusion has two shortcomings. The first one is that unbiased estimation does not mean that the estimation results must be reliable. According to the Gauss-Markov theorem, the unbiased estimation variance has a lower bound, and when the lower bound itself is very large, even if the unbiased estimation variance is minimized, the difference between the estimation results and the true value (hereinafter referred to as estimation error) may not be within the acceptable range [21]. The other one is that the optimal weighting factor is not necessarily optimal; the optimal weighting factor is calculated according to the measurement variance; a smaller measurement variance does not necessarily mean a smaller measurement error, so in some cases, the optimal weighting factor is not optimal. Therefore, this paper proposes a biased estimation data fusion algorithm that also does not require priori knowledge, which uses unbiased estimation values to construct biased estimation values with lower estimation errors and calculates the optimal weighting factor according to the estimation error to ensure its optimality.

This paper is organized as follows: Section 2 reviews some existing unbiased estimation data fusion algorithms, Section 3 describes the shortcomings of the existing algorithms and the feasibility of the improved methods, Section 4 describes the specific implementation of the algorithms in this paper, Section 5 compares the performance of the different algorithms, and Section 6 summarizes the entire content and provides prospects for the future.

2. Unbiased Estimation Data Fusion

When \( N \) sensors are measuring an invariant physical quantity and all the sensor nodes in a network measure the observed physical quantity at the same time, the measurement equation is:

\[
\begin{align*}
 z(k)_i &= H_i x + v(k)_i \\
\end{align*}
\]

where \( x \) is the unknown quantity to be measured, \( \mu \) is the true value of \( x \), and \( k \) is the discrete time indicator, \( z(k)_i \) is the measurement value of the \( i \)th sensor at the moment \( k \), \( H_i \) is the measurement matrix of the \( i \)th sensor (the \( H \) are often unit matrices), \( v(k)_i \) is the measurement noise, \( v(k)_i \sim N(0, v_i^2) \) and are independent of each other, and \( v_i^2 \) be called measurement variance. Moreover, according to the additivity of the normal distribution, \( z(k)_i \sim N(\mu, v_i^2) \).

In this system, only \( z(k)_i \) is a known quantity; how one can use \( z(k)_i \) to calculate the true value of \( \mu \)? Since the measurement equations are designed for multiple sensors and the measurement variances may vary, adaptive weighted data fusion is often used to solve such problems, but adaptive weighted data fusion needs prior knowledge of the measurement variance of all the sensors, and in practice, the measurement covariance can be obtained by experiments.

2.1. Unbiased Estimation

Unbiased estimation refers to: if the mathematical expectation of an estimator is equal to the true value of the estimated parameter (this property is called unbiasedness), then this estimator is called the unbiased estimation of the estimated parameter. The mathematical expectation of the estimator constructed by unbiased estimation is equal to the true value of the estimated parameter, so using unbiased estimators instead of
measurement values will not cause an increase in error. Unbiased estimation of sample variance can be used to replace the measurement variance of sensors, allowing adaptive weighted data fusion to operate without knowing the measurement variance of sensors. There are two commonly used unbiased estimation algorithms, as follows:

Least squares estimation:

Assuming that a particular sensor collects n data over a period of time, \( T_i \), the estimation value, and \( \sigma^2 \), estimation variance, are given as follows:

\[
\tilde{T} = \frac{1}{n} \sum_{i=1}^{n} T_i
\]

\[
\tilde{\sigma}^2 = \frac{n-1}{n} \sum_{i=1}^{n} (T_i - \tilde{T})^2
\]

Batch estimation:

Assuming that a particular sensor collects n data over a period of time, these n data are randomly divided into two groups, where the data in the \( i \)th groups are: \( T_{ij}, T_{i2}, \ldots, T_{inn}, i \geq n_i \geq 2 \), \( i = 1, 2 \). \( T_{ij} \), the sample mean, and \( \sigma^2 \), sample variance, are given as follows:

\[
\bar{T}_i = \frac{1}{n_i} \sum_{j=1}^{n_i} T_{ij} \quad i = 1, 2
\]

\[
\sigma^2_i = \frac{n_i-1}{n_i} \sum_{j=1}^{n_i} (T_{ij} - \bar{T}_i)^2
\]

According to the theory of batch estimation in statistics, the optimal estimation value \( \tilde{T} \) and the optimal estimation variance \( \tilde{\sigma}^2 \) of these 2 sets of data are given as follows:

\[
\tilde{T} = \frac{\sigma^2 \bar{T}_2 + \sigma^2 \bar{T}_1}{\Sigma_i \sigma^2_i}
\]

\[
\tilde{\sigma}^2 = \frac{\sigma^2 \sigma^2}{\Sigma_i \sigma^2_i}
\]

2.2. Adaptive Weighted Data Fusion

Adaptive weighted data fusion principle: assume that there are \( N \) sensors working on the same unknown quantity \( x \) (the true value is \( \mu \)) is measured, and the measurement value of the \( i \)th sensor is \( M_i \). The measurement variance of the \( i \)th sensor is \( \sigma^2_i \), and the measurement data of the \( i \)th sensor follows \( N(\mu, \sigma^2_i) \), then the unbiased estimator \( \tilde{M} \) is given as follows:

\[
\tilde{M} = \sum_{i=1}^{N} w_i M_i
\]

\[
\tilde{\sigma}^2 = \sum_{i=1}^{N} w_i^2 \sigma^2_i
\]

Since \( \tilde{M} \) is an unbiased estimator, so that

\[
\sum_{i=1}^{N} w_i = 1
\]

In order to find \( w_i \) that minimizes \( \tilde{\sigma}^2 \) (call this \( w_i \) as the optimal weighting factor), the auxiliary function that is constructed according to the Lagrange multiplier method is given as follows:

\[
f(w_1, w_2, \ldots, w_N, \varphi) = \sum_{i=1}^{N} w_i^2 \sigma^2_i + \varphi (1 - \sum_{i=1}^{N} w_i)
\]

According to multivariate extremum theory, it is known that the solution of the following partial differential equation makes the estimation variance \( \tilde{\sigma}^2 \) minimized, as follows:
The simplified formula is given as follows:

\[
\begin{align*}
\frac{\partial f}{\partial w_1} &= 2w_1 v_1^2 - \varphi = 0 \\
\vdots \\
\frac{\partial f}{\partial w_n} &= 2w_n v_n^2 - \varphi = 0 \\
\frac{\partial f}{\partial \varphi} &= (1 - \sum_{i=1}^{N} w_i) = 0
\end{align*}
\] (12)

Thus, the formula for \( w_i \) are given as follows:

\[
\begin{align*}
w_i &= \frac{\theta}{v_i^2}, \quad i = 1, 2, \ldots, n; \theta = \frac{\varphi}{2} \\
\sum_{i=1}^{N} w_i &= 1
\end{align*}
\] (13)

So, \( \bar{M} \) and \( \sigma^2 \) are given as follows:

\[
\begin{align*}
\bar{M} &= \sum_{i=1}^{N} \left( v_i^2 \sum_{j=1}^{n} \frac{1}{v_j^2} \right)^{-1} M_i \\
\sigma^2 &= \left( \sum_{i=1}^{N} \frac{1}{v_i^2} \right)^{-1}
\end{align*}
\] (15, 16)

When adaptive weighted data fusion is combined with other unbiased estimation algorithms, it is possible to replace \( M_i \) and \( v_i^2 \) in Equation (15) with an unbiased estimation value and an unbiased estimation variance. The Figure 1 shows the flowchart for batch estimation data fusion, and the flowchart for other unbiased estimation data fusion is more or less the same.

Figure 1. Flowchart of batch estimation data fusion.

3. Rationale

Statistical inference is a statistical method of inferring the whole through samples, and there are two kinds of errors in statistical inference: systematic error and random error. Unbiased estimation has no systematic error because of its unbiased nature, but there is still random error, so its estimation results are not necessarily reliable.

The optimal weighting factor is calculated based on the measurement variance alone, but a smaller measurement variance does not mean a smaller measurement error.
Although the measurement variance of the data measured by different sensors varies, the mathematical expectations are all equal, so, in practice, it is easy to have a smaller measurement variance but a larger measurement error.

### 3.1. Biased Estimators

**Theorem 1.** The error of the unbiased estimator can be further optimized.

**Proof of Theorem 1.** Assuming that the true value of the measured unknown quantity $x$ is $\mu$, the sensor measurement data follows $N(\mu, \sigma^2)$. $E$ is an unbiased estimator constructed using a linear combination of the sensor measurement data, and letting the estimation variance of $E$ be $\sigma^+$, then $E \sim N(\mu, \sigma^2)$, defining the estimation error $\text{est}$ as follows:

$$\text{est} = E - \mu \quad (17)$$

$E(\text{est}) = 0, D(\text{est}) = \sigma^2$ is easy to obtain, and $\text{est}$ follows $N(0, \sigma^2)$. Assuming that the estimation error is tolerable when $\text{est} \in [-z, z]$, the probability that the estimation error is tolerable is:

$$P(\text{est} \in [-z, z]) = 2 \int_0^z \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x^2}{2\sigma^2}} dx \quad (18)$$

If $z = \sigma$, we obtain: $P(\text{est} \in [-\sigma, \sigma]) \approx 68.3\%$, which means that for any $\sigma$, there is about a 31.7% chance that $\text{est}$ is larger than the estimation standard deviation $\sigma$, so that even if $\sigma$ is very small, $\text{est}$ may be able to be reduced further. $\Box$

**Theorem 2.** A biased estimation value with a smaller error can be constructed from an unbiased estimation value.

**Proof of Theorem 2.** Assuming unbiased estimator $A$ follows $N(\mu, \sigma^2)$, the biased estimator corresponding to $A$ is defined as $B$ and $B = A + s$, where $s$ is called offset, which is a nonzero variable. The following results can be obtained:

$$E(B) = E(A) + E(s) = \mu + s \quad (s \neq 0) \quad (19)$$

Because $E(B)$ is not equal to the true value $\mu$, $B$ is a biased estimator. Assuming that $A_1$ is a specific estimate of $A$, its estimation error is $d_1$ and $d_1 = |A_1 - \mu|$. $s_1$ is the offset corresponding to $A_1$ and $s_1 \propto d_1$, and the biased estimator corresponding to $A_1$ is $B_1$ and $B_1 = A_1 + s_1$. The estimation error of $B_1$ is $e_1$ and $e_1 = |A_1 - \mu + s_1|$, so:

$$e_1 = \begin{cases} d_1 + |s_1| & \text{if } (A_1 - \mu)s_1 > 0 \\ d_1 - |s_1| & \text{if } (A_1 - \mu)s_1 < 0 \text{ and } |s_1| < d_1 \\ |s_1| - d_1 & \text{if } (A_1 - \mu)s_1 < 0 \text{ and } |s_1| > d_1 \end{cases} \quad (20)$$

So, when $(A_1 - \mu)s_1 < 0$ and $|s_1| < 2d_1$, $e_1 < d_1$. In other words, when the sign of $s_1$ is opposite to $(A_1 - \mu)$ and $|s_1| < 2d_1$, the estimation error of biased estimation is smaller than that of unbiased estimation. $\Box$

### 3.2. Optimal Weighting Factors

**Theorem 3.** In some cases, the optimal weighting factor is not optimal.

**Proof of Theorem 3.** Suppose there are two sensors $S_1$ and $S_2$ that simultaneously measure the unknown quantity $x$ (the true value is $\mu$). Their measurement variances are $\nu_1^2$ and $\nu_2^2$, and they satisfy $\nu_1^2 < \nu_2^2$. Their measurement data are $z_1$ and $z_2$. For
adaptive weighted data fusion, \( w_1 \), the weights of \( S_1 \), will always be greater than \( w_2 \), the weights of \( S_2 \), since \( v_1^2 < v_2^2 \).

Define \( m_i \), the measurement error (unlike \( est \), \( m_i \) only considers the magnitude of the error, not the sign of the error), as follows:

\[
m_i = |z_i - \mu|
\]  

\label{eq:21}

So \( m_1 = |z_1 - \mu|, m_2 = |z_2 - \mu| \). If \( m_1 > m_2 \), then \( w_1 \times m_1 + w_2 \times m_2 > w_1 \times m_2 + w_2 \times m_1 \), which implies that the estimation error of the fusion estimation value computed from optimal weighting factors is not minimal. So, when \( v_1^2 < v_2^2 \) but \( m_1 > m_2 \), the optimal weighting factor is not optimal, and the probability is given as follows:

\[
P(m_1 > m_2) = 2 \int_{-\infty}^{+\infty} \int_{x}^{X} \frac{1}{\sqrt{2\pi}v_2} e^{-\frac{x^2}{2v_2^2}} dx_1
\]  

\label{eq:22}

It can be shown that \( P(m_1 > m_2) > 0 \), which indicates that the weights of any two sensors may not be optimal. The above proof uses measured values and measured variances, but replacing them with estimated values and estimated variances is equally valid. \( \square \)

**Theorem 4.** It is optimal to compute the optimal weighting factor using the measurement error.

**Proof of Theorem 4.** Suppose there are \( N \) sensors measuring the unknown quantity \( x \) (the true value is \( \mu \)), their measurement values are \( M_1, M_2 \ldots M_n \), and their measurement errors are \( m_1, m_2 \ldots m_n \); moreover, there are \( 0 < m_i < m_{i+1} \). The optimal weighting factor calculated from the measurement error is \( w_1, w_2 \ldots w_n \), because \( m_i < m_{i+1} \), then \( w_i > w_{i+1} \), so the corresponding fusion estimation value \( \bar{M}_1 \) is as follows:

\[
\bar{M}_1 = \sum_{i=1}^{N} w_i M_i
\]  

\label{eq:23}

In this paper, only one case is proved, which has two weights different from \( w_1, w_2 \ldots w_n \), and the proof can be easily generalized to other cases.

Counterfactual: suppose that when the optimal weighting factor is \( v_1, v_2, v_3 \ldots v_n \) and \( v_1 < v_2 \), the fusion estimation value \( \bar{M}_2 \) is more accurate than \( \bar{M}_1 \), where \( \bar{M}_2 \) is given as follows:

\[
\bar{M}_2 = v_1 M_1 + v_2 M_2 + \sum_{i=3}^{N} w_i M_i
\]  

\label{eq:24}

The common term of \( \bar{M}_1 \) and \( \bar{M}_2 \) is \( \bar{M}_3 \), and \( \bar{M}_3 = \sum_{i=3}^{N} w_i M_i, E(\bar{M}_3) = \alpha \mu \) where \( \alpha = \sum_{i=3}^{N} w_i \). So \( \bar{M}_1 \) and \( \bar{M}_2 \) become this form:

\[
\bar{M}_1 = w_1 M_1 + w_2 M_2 + \bar{M}_3, \hspace{1cm} \bar{M}_2 = v_1 M_1 + v_2 M_2 + \bar{M}_3
\]  

\label{eq:25}

Because \( w_1 + w_2 \ldots + w_n = v_1 + v_2 + w_3 \ldots + w_n = 1 \), so \( w_1 + w_2 = v_1 + v_2 \) and set \( w_1 + w_2 = \theta \). Substituting \( M_i = \mu + m_i \) for \( \bar{M}_1 \) and \( \bar{M}_2 \), then we get the following equation:

\[
\bar{M}_1 = (w_1 + w_2)\mu + w_1 m_1 + w_2 m_2 + \bar{M}_3
\]  

\label{eq:26}

\[
\bar{M}_2 = (v_1 + v_2)\mu + v_1 m_1 + v_2 m_2 + \bar{M}_3
\]  

\label{eq:27}

Because \( \bar{M}_3 = (w_1 + w_2)\mu \), so the larger \( |w_1 m_1 + w_2 m_2| \) and \( |v_1 m_1 + v_2 m_2| \) are the less accurate \( \bar{M}_1 \) and \( \bar{M}_2 \) are. Because \( m_1 > 0 \) and \( m_2 > 0 \), so \( |w_1 m_1 + w_2 m_2| = w_1 m_1 + w_2 m_2 \) and \( |v_1 m_1 + v_2 m_2| = v_1 m_1 + v_2 m_2 \). Set \( f(\gamma) \) is given as follows:

\[
f(\gamma) = \gamma m_1 + (\theta - \gamma)m_2
\]  

\label{eq:28}
So \( f(w) = w_1m_1 + w_2m_2 \) and \( f(v) = v_1m_1 + v_2m_2 \). The derivative of \( f(y) \) is given as follows:

\[
\frac{df(y)}{dy} = m_1 - m_2 < 0
\] (29)

So, the larger \( y \) is, the smaller \( f(y) \). Because \( w_1 > w_2 \) and \( v_1 < v_2 \), so \( w_1 > \frac{\theta}{2} \), \( v_1 < \frac{\theta}{2} \) and \( w_1m_1 + w_2m_2 < v_1m_1 + v_2m_2 \). \( \bar{M}_1 \) is more accurate than \( \bar{M}_2 \). Contradicting the hypothesis, the optimal weighting factor calculated from the estimation error is better. The above proof uses measured values and measured variances, but replacing them with estimated values and estimated variances is equally valid. □

4. Algorithm Implementation

The algorithm in Section 4.1 only uses measurement data from a single sensor, so it can be implemented in a distributed manner. The algorithms in Sections 4.2 and 4.3 include data fusion algorithms, so they can only be implemented using centralized methods.

4.1. Unbiased Estimation

Suppose that \( N \) sensors simultaneously measure the unknown quantity \( x \) (the true value is \( \mu \)), and all the sensor nodes in a network measure the observed physical quantity at the same time. For any of the sensors, if it collects 10 data over a period of time, in order from smallest to largest: \( a_1, a_2, \cdots, a_{10} \), then construct the unbiased estimator \( E \) as follows:

\[
E_1 = \frac{a_1 + a_2}{2}, \quad E_2 = \frac{a_3 + a_4 + a_5 + a_{10}}{4}, \quad E_3 = \frac{a_3 + a_4 + a_5 + a_{9}}{4}, \quad E_4 = \frac{1}{10}\sum_{i=1}^{10} a_i
\] (30)

Assuming that the sensor follows \( N(\mu, \sigma^2) \), theoretically, \( E_1 \sim N\left( \mu, \frac{\sigma^2}{2} \right) \) and \( E_2, E_3 \sim N\left( \mu, \frac{\sigma^2}{4} \right) \). Because \( a_1 \sim a_{10} \) are data sorted in order of size, these estimators have lower estimated variances. Through experiments, it was found that the estimated variances of \( E_1 \sim E_3 \) are approximately \( \frac{\sigma^2}{7.5}, \frac{\sigma^2}{8.5} \). Therefore, \( E_1 \sim E_3 \) is independent of each other and has an estimated variance smaller than the measured variance.

\( \bar{E} \), the mean of estimation values, and \( \sigma_f^2 \), the estimation variance, are given as follows:

\[
\bar{E} = \frac{\sum_{i=1}^{10} E_i}{10}, \quad \sigma_f^2 = (E_i - \bar{E})^2
\] (31)

An adaptive weighted data fusion algorithm is used to obtain \( \hat{x} \), the unbiased estimation value, and \( \sigma^2 \), the unbiased estimation variance. And the specific formula is as follows:

\[
w_i = (\sigma^2 \sum_{i=1}^{10} \frac{1}{\sigma_i})^{-1}, \quad \hat{x} = \sum_{i=1}^{10} w_i E_i, \quad \delta^2 = (\sum_{i=1}^{10} \frac{1}{\sigma_i})^{-1}
\] (32)

Because \( E_4 \) is the least squares estimator, theoretically, the accuracy of the fused estimate \( \hat{x} \) will be higher than that of the least squares estimate.

If the number of measurement data collected by a sensor within a certain period is \( m \) and \( m \) is not equal to 10, the unbiased estimator can be constructed according to the following rules:

1. Each group consists of at least two measurement values, and each group does not use duplicate measurement values. The unbiased estimator corresponding to each group is the mean of the measurement values within the group.
2. When \( m \) is even (\( m = 2k \)), for any measurement value \( a_i \), it is necessary to ensure that \( a_i \) and \( a_{m-i} \) are in the same group.
3. When \( m \) is odd (\( m = 2k + 1 \)), for any measurement value \( a_i \), it is necessary to ensure that \( a_i \) and \( a_{m-i} \) are in the same group. In addition, it is also necessary to ensure that \( a_{k}, a_{k+1}, \) and \( a_{k+2} \) are in the same group.
4. Add an additional group that includes all measured values, such as $E_4$ in Equation (30).

4.2. Biased Estimation

$|s_i|$, the size of the offset $s_i$, needs to be determined first. For the $i$th sensor, define its unbiased estimation value and unbiased estimation variance as $\bar{A}_i$ and $\sigma^2_i$. The fusion estimation value $\bar{A}$ and the fusion estimation variance of these unbiased estimation values $\bar{\sigma}^2$ are computed by adaptive weighted data fusion algorithm.

Define $g_i$ as the degree of deviation of the $i$th unbiased estimation value from the true value. And $g_i$ is given as (use $\bar{A}$ replaces true value) follows:

$$g_i = \frac{|A_i - \bar{A}|}{\bar{\sigma}}$$

(33)

Define $\omega$ as the coefficient of $|s_i|$, where the size of $s_i$, and the value of $\omega_i$ are determined by $g_i$. When $g_i < 0.1$, $\omega_i = 0$, when $1 > g_i > 0.1$, $\omega_i = 0.5$. Otherwise, $\omega_i = 1$. $\omega$ is used to control the size of the offset and avoid excessive offset.

$|s_i|$ is given as follows:

$$|s_i| = \frac{g_i \cdot \sigma_i}{\bar{\sigma}} \cdot \omega_i$$

(34)

Then, the sign of $s_i$ needs to be determined. $s_i$ takes a negative sign when $A_i > \bar{A}$ and a positive sign otherwise. So $B_i$ is given as follows:

$$B_i = \begin{cases} \frac{g_i \cdot \sigma_i}{\bar{\sigma}} \cdot \bar{\sigma} \cdot \omega_i & \text{if } A_i > \bar{A} \\ -\frac{g_i \cdot \sigma_i}{\bar{\sigma}} \cdot \bar{\sigma} \cdot \omega_i & \text{if } A_i < \bar{A} \end{cases}$$

(35)

4.3. Data Fusion

The application of the adaptive weighted data fusion algorithm requires measuring variance, which is replaced by estimated variance in unbiased estimation data fusion. In biased estimation data fusion, the estimated variance is calculated based on the estimation error, and the design concept of the calculation formula comes from the Kalman gain. Define $e_i$ as the biased estimation error of $B_i$, and $e_i$ is given as follows:

$$e_i = B_i - \mu$$

(36)

The true value is unknown; use $\bar{A}$ instead of the true value. So $e_i^2$ is given as:

$$e_i^2 = (B_i - \bar{A})^2$$

(37)

Then, through 100,000 experiments, calculate $\bar{e}_i^2$, mean of $e_i^2$, and $\bar{g}_i$, mean of $g_i$, for different $g_i$-value intervals. The $\bar{e}_i^2$ and $\bar{g}_i$ over the three $g_i$ intervals are shown in the Table 1.

**Table 1.** Mean estimation error and mean $g$ for different $g$ intervals.

<table>
<thead>
<tr>
<th>$g_i$</th>
<th>$\bar{e}_i^2$</th>
<th>$\bar{g}_i$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$g_i &lt; 0.1$</td>
<td>0.0176</td>
<td>0.0493</td>
</tr>
<tr>
<td>$0.1 \leq g_i &lt; 1$</td>
<td>0.0203</td>
<td>0.5268</td>
</tr>
<tr>
<td>$g_i \geq 1$</td>
<td>0.0293</td>
<td>1.6735</td>
</tr>
</tbody>
</table>

$\bar{e}_i^2$ and $\bar{g}_i$ can be represented as follows:

$$\begin{cases} \bar{e}_i^2 = 0.0176, \bar{g}_i = 0.0493 & \text{if } g_i < 0.1 \\ \bar{e}_i^2 = 0.0203, \bar{g}_i = 0.5268 & \text{if } 0.1 \leq g_i < 1 \\ \bar{e}_i^2 = 0.0293, \bar{g}_i = 1.6735 & \text{if } g_i \geq 1 \end{cases}$$

(38)

So, $\bar{\sigma}_i^2$, which represents the biased estimation variance of $B_i$, is given as follows:
\[
\sigma^2_i = e^2_i + \frac{\sigma_i}{B_i + 8_i} \cdot (e^2_i - e^2_i)
\]  

(39)

Thus, \(w_i\), the optimal weighting factor, and \(\hat{M}\), the fusion estimation value, are given as follows:

\[
w_i = (\sigma^2_i \sum_{i=1}^{N} \frac{1}{B_i})^{-1}, \quad \hat{M} = \sum_{i=1}^{N} w_i B_i
\]  

(40)

The flowchart of biased estimation data fusion is shown in Figure 2:

**Figure 2.** Flowchart of biased estimation data fusion.

5. Results

The simulation test is divided into two parts as follows: 1. a normal distribution white noise test (all noise follows a normal distribution), and 2. A uniform distribution white noise test (environmental noise is uniformly distributed white noise). The test items for the two tests are shown in Table 2:

**Table 2.** The test items for the two tests.

<table>
<thead>
<tr>
<th>Test Items</th>
<th>Test Value 1</th>
<th>Test Value 2</th>
<th>Test Value 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>True value</td>
<td>20</td>
<td>null</td>
<td>null</td>
</tr>
<tr>
<td>Measurement variance</td>
<td>0.25</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Noise variance</td>
<td>0.2</td>
<td>0.5</td>
<td>1</td>
</tr>
</tbody>
</table>

The testing mode is as follows: ten rounds of experiments are conducted for each combination of test items, and ten tests are conducted for each round. The average of the ten test results is used as the result of this round of experiments. Three sensors are used for each test, with 10 measurement data used for each sensor.
The test indicators are mean relative error (mre) and mean square error (mse). mre is used to compare the average performance of the three algorithms, and mse is used to compare the stability of the three algorithms. Their calculation formulas are as follows:

$$mre_i = \frac{\sum_{j=1}^{10} (m_{ij} - \mu)}{10}$$

$$mse_i = \frac{\sum_{j=1}^{10} (m_{ij} - \mu)^2}{10}$$

where \(i\) represents the number of test rounds and \(m_{ij}\) represents the estimated value of the \(j\)th test in the \(i\)th round. Relative improvement (ri) is used to compare the performance differences between two algorithms, and the calculation formula for ri is as follows:

$$ri = \frac{V_\alpha - V_\beta}{V_\alpha} \times 100\%$$

where \(V_\alpha\) and \(V_\beta\) represent the same indicator value (e.g., mre) for algorithms \(\alpha\) and algorithms \(\beta\). And, if the indicator has multiple values, calculate the mean of the indicator first, and then calculate ri. If \(ri\) is greater than 0, it means that algorithm \(\beta\) has improved \(ri\) compared to algorithm \(\alpha\).

5.1. Normal Distribution White Noise Test

The test results are as follows (Figures 3–5 show the test results when the measurement noise is 0.25, and the remaining result graphs are Figures A1–A6 in Appendix A):

![Graphs showing test results](image)

**Figure 3.** mre and mse when measurement variance = 0.25 and normal distribution noise variance = 0.2.
From Table 3, it can be seen that the \( m_{re} \) of biased estimation data fusion (BED) decreased by an average of 9.25% and 10.26% compared to least squares estimation data fusion (LSD) and batch estimation data fusion (BTD). Therefore, BED has higher fusion accuracy.

From Table 4, it can be seen that the \( m_{se} \) of BED decreased by an average of 16.24% and 20.28% compared to LSD and BTD. Therefore, the bed has higher stability.

**Table 3.** \( m_{re} \) of all algorithms when noise follows normal distribution.

<table>
<thead>
<tr>
<th>Measure Variance</th>
<th>Noise Variance</th>
<th>Least Squares Estimation Data Fusion</th>
<th>Batch Estimation Data Fusion</th>
<th>Biased Estimation Data Fusion</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.25</td>
<td>0.2</td>
<td>0.506%</td>
<td>0.518%</td>
<td>0.445%</td>
</tr>
<tr>
<td></td>
<td>0.5</td>
<td>0.615%</td>
<td>0.597%</td>
<td>0.555%</td>
</tr>
<tr>
<td></td>
<td>1.0</td>
<td>0.872%</td>
<td>0.887%</td>
<td>0.816%</td>
</tr>
</tbody>
</table>
Table 4. \( mse \) of all algorithms when noise follows normal distribution.

<table>
<thead>
<tr>
<th>Measure Variance</th>
<th>Noise Variance</th>
<th>Least Squares Estimation Data Fusion</th>
<th>Batch Estimation Data Fusion</th>
<th>Biased Estimation Data Fusion</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.25</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.2</td>
<td>0.017</td>
<td>0.017</td>
<td>0.014</td>
</tr>
<tr>
<td></td>
<td>0.5</td>
<td>0.024</td>
<td>0.025</td>
<td>0.018</td>
</tr>
<tr>
<td></td>
<td>1.0</td>
<td>0.047</td>
<td>0.050</td>
<td>0.042</td>
</tr>
<tr>
<td></td>
<td>mean</td>
<td>0.029</td>
<td>0.030</td>
<td>0.024</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.2</td>
<td>0.068</td>
<td>0.067</td>
<td>0.053</td>
</tr>
<tr>
<td></td>
<td>0.5</td>
<td>0.075</td>
<td>0.081</td>
<td>0.061</td>
</tr>
<tr>
<td></td>
<td>1.0</td>
<td>0.065</td>
<td>0.066</td>
<td>0.054</td>
</tr>
<tr>
<td></td>
<td>mean</td>
<td>0.069</td>
<td>0.071</td>
<td>0.056</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.2</td>
<td>0.073</td>
<td>0.079</td>
<td>0.055</td>
</tr>
<tr>
<td></td>
<td>0.5</td>
<td>0.103</td>
<td>0.104</td>
<td>0.091</td>
</tr>
<tr>
<td></td>
<td>1.0</td>
<td>0.127</td>
<td>0.136</td>
<td>0.111</td>
</tr>
<tr>
<td></td>
<td>mean</td>
<td>0.101</td>
<td>0.106</td>
<td>0.085</td>
</tr>
</tbody>
</table>

And, as the noise variance increases, the average increases in \( mre \) and \( mse \) of LSD are 0.297% and 0.036, the average increases in \( mre \) and \( mse \) of BTD are 0.306% and 0.038, and the average increases in the \( mre \) and \( mse \) of BED are 0.269% and 0.03. So, \( mre \) and \( mse \) of BED decreased by an average of 11.59% and 18.85% compared to LSD and BTD, respectively. Therefore, BED has higher noise resistance.

5.2. Uniform Distribution White Noise Test

The test results are as follows (Figures 6–8 show the test results when the measurement noise is 0.25, and the remaining result graphs are shown in Figures A7–A12):
Figure 6. $mre$ and $mse$ when measurement variance = 0.25 and uniformly distributed noise variance = 0.2.

Figure 7. $mre$ and $mse$ when measurement variance = 0.25 and uniformly distributed noise variance = 0.5.
From Table 5, it can be seen that the \( mre \) of BED decreased by an average of 8.56% and 6.58% compared to LSD and BTD. Therefore, BED has higher fusion accuracy, but the advantage of \( mre \) is reduced by 22.39% compared to normal distribution white noise testing.

From Table 6, it can be seen that the \( mse \) of BED decreased by an average of 17.42% and 21.01% compared to LSD and BTD. Therefore, BED has higher fusion accuracy, but the advantage of \( mse \) is increased by 5.23% compared to normal distribution white noise testing.

**Table 5.** \( mre \) of all algorithms when noise follows uniform distribution.

<table>
<thead>
<tr>
<th>Measure Variance</th>
<th>Noise Variance</th>
<th>Least Squares Estimation Data Fusion</th>
<th>Batch Estimation Data Fusion</th>
<th>Biased Estimation Data Fusion</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.2</td>
<td>0.463%</td>
<td>0.494%</td>
<td>0.430%</td>
</tr>
<tr>
<td></td>
<td>0.5</td>
<td>0.455%</td>
<td>0.458%</td>
<td>0.424%</td>
</tr>
<tr>
<td></td>
<td>1.0</td>
<td>0.564%</td>
<td>0.578%</td>
<td>0.518%</td>
</tr>
<tr>
<td></td>
<td>mean</td>
<td>0.494%</td>
<td>0.510%</td>
<td>0.457%</td>
</tr>
<tr>
<td>0.25</td>
<td>0.2</td>
<td>0.762%</td>
<td>0.830%</td>
<td>0.705%</td>
</tr>
<tr>
<td></td>
<td>0.5</td>
<td>0.769%</td>
<td>0.794%</td>
<td>0.685%</td>
</tr>
<tr>
<td></td>
<td>1.0</td>
<td>0.962%</td>
<td>0.947%</td>
<td>0.865%</td>
</tr>
<tr>
<td></td>
<td>mean</td>
<td>0.831%</td>
<td>0.857%</td>
<td>0.751%</td>
</tr>
<tr>
<td>1</td>
<td>0.2</td>
<td>1.075%</td>
<td>1.084%</td>
<td>0.981%</td>
</tr>
<tr>
<td></td>
<td>0.5</td>
<td>1.058%</td>
<td>1.071%</td>
<td>0.982%</td>
</tr>
<tr>
<td></td>
<td>1.0</td>
<td>1.180%</td>
<td>1.167%</td>
<td>1.078%</td>
</tr>
<tr>
<td></td>
<td>mean</td>
<td>1.104%</td>
<td>1.107%</td>
<td>1.013%</td>
</tr>
<tr>
<td>2</td>
<td>0.2</td>
<td>0.012</td>
<td>0.013</td>
<td>0.011</td>
</tr>
<tr>
<td></td>
<td>0.5</td>
<td>0.012</td>
<td>0.012</td>
<td>0.011</td>
</tr>
<tr>
<td></td>
<td>1.0</td>
<td>0.018</td>
<td>0.021</td>
<td>0.016</td>
</tr>
<tr>
<td></td>
<td>mean</td>
<td>0.014</td>
<td>0.015</td>
<td>0.012</td>
</tr>
</tbody>
</table>

**Table 6.** \( mse \) of all algorithms when noise follows uniform distribution.

<table>
<thead>
<tr>
<th>Measure Variance</th>
<th>Noise Variance</th>
<th>Least Squares Estimation Data Fusion</th>
<th>Batch Estimation Data Fusion</th>
<th>Biased Estimation Data Fusion</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.25</td>
<td>0.2</td>
<td>0.012</td>
<td>0.013</td>
<td>0.011</td>
</tr>
<tr>
<td></td>
<td>0.5</td>
<td>0.012</td>
<td>0.012</td>
<td>0.011</td>
</tr>
<tr>
<td></td>
<td>1.0</td>
<td>0.018</td>
<td>0.021</td>
<td>0.016</td>
</tr>
<tr>
<td></td>
<td>mean</td>
<td>0.014</td>
<td>0.015</td>
<td>0.012</td>
</tr>
</tbody>
</table>
As the noise variance increases, the average increases in \( mre \) and \( mse \) of LSD are 0.305\% and 0.03, the average increases in \( mre \) and \( mse \) of BTD are 0.298\% and 0.031, and the average increases in \( mre \) and \( mse \) of BED are 0.278\% and 0.024. So, \( mre \) and \( mse \) of BED decreased by an average of 8.21\% and 21.29\% compared to LSD and BTD. BED has higher noise resistance, and the comprehensive noise resistance is equivalent to that of normal distribution white noise testing.

In summary, even if the noise does not conform to a normal distribution, as long as the mean noise is 0, the algorithm can perform well.

6. Discussion

In this paper, an adaptive weighting data fusion algorithm based on biased estimation is proposed to address the shortcomings of unbiased estimation data fusion. The algorithm in this paper first further reduces the estimation error by biased estimation and then uses the estimation error instead of the measurement variance to compute the optimal weighting factor, which avoids the fact that the optimal weighting factor is not optimal in some cases. The final results of the simulation experiments show that this algorithm outperforms the original algorithm in terms of fusion accuracy, stability, and noise resistance, which are better than other algorithms.

The essence of this algorithm is to sacrifice unbiasedness for higher estimation accuracy, so the overall accuracy of biased estimation is significantly higher than that of unbiased estimation. However, due to the loss of unbiasedness, the performance of the algorithm is not as good as expected when using adaptive weighted data fusion, so there is still much room for improvement.

Author Contributions: Writing—original draft preparation, M.Q.; writing—review and editing, B.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Major Science and Technology Projects in Yunnan Province (202202AE090032).

Data Availability Statement: Data are contained within the article.

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

Appendix A
Figure A1. \( m_{re} \) and \( m_{se} \) when measurement variance = 1 and normal distribution noise variance = 0.2.

Figure A2. \( m_{re} \) and \( m_{se} \) when measurement variance = 1 and normal distribution noise variance = 0.5.
Figure A3. $mre$ and $mse$ when measurement variance = 1 and normal distribution noise variance = 1.

Figure A4. $mre$ and $mse$ when measurement variance = 2 and normal distribution noise variance = 0.2.
Figure A5. \( m_{re} \) and \( m_{se} \) when measurement variance = 2 and normal distribution noise variance = 0.5.

Figure A6. \( m_{re} \) and \( m_{se} \) when measurement variance = 2 and normal distribution noise variance = 1.
Figure A7. $m_{re}$ and $m_{se}$ when measurement variance = 1 and uniformly distributed noise variance = 0.2.

Figure A8. $m_{re}$ and $m_{se}$ when measurement variance = 1 and uniformly distributed noise variance = 0.5.
Figure A9. $mre$ and $mse$ when measurement variance = 1 and uniformly distributed noise variance = 1.

Figure A10. $mre$ and $mse$ when measurement variance = 2 and uniformly distributed noise variance = 0.2.
Figure A11. $m_{re}$ and $m_{se}$ when measurement variance = 2 and uniformly distributed noise variance = 0.5.

Figure A12. $m_{re}$ and $m_{se}$ when measurement variance = 2 and uniformly distributed noise variance = 1.

References


Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.