

Article

Bayesian Optimization for Fine-Tuning EKF Parameters in UAV Attitude and Heading Reference System Estimation

Assefinew Wondesen , Yisak Debele , Seung-Ki Kim , Ha-Young Shi, Bedada Endale and Beom-Soo Kang *

Department of Aerospace Engineering, Pusan National University, Busan 46241, Republic of Korea; wondebly@pusan.ac.kr (A.W.); yisaktol@pusan.ac.kr (Y.D.); akkim7631@pusan.ac.kr (S.-K.K.); shy621@pusan.ac.kr (H.-Y.S.); endale@pusan.ac.kr (B.E.)

* Correspondence: bskang@pusan.ac.kr; Tel.: +82-51-510-2310

Abstract: In various applications, the extended Kalman filter (EKF) has been vital in estimating a vehicle's translational and angular motion in 3-dimensional (3D) space. It is also essential for the fusion of data from multiple sensors. However, for the EKF to perform effectively, the optimal process noise covariance matrix (Q) and measurement noise covariance matrix (R) must be chosen correctly. The use of EKF has been challenging due to the need for an easy mechanism to select Q and R values. As a result, this research focused on developing an algorithm that can be easily applied to determine Q and R , allowing us to harness the full potential of EKF. Accordingly, an EKF innovation consistency statistics-driven Bayesian optimization algorithm was employed to achieve this goal. Q and R values were tuned until the expected result met the performance requirement for minimum error through improved measurement innovation consistency. The comprehensive results demonstrate that when the optimum Q and R , as tuned by the suggested technique, were used, the performance of the EKF significantly improved.

Keywords: extended Kalman filter; covariance tuning; Bayesian optimization; attitude and heading reference system; UAV



Citation: Wondesen, A.; Debele, Y.; Kim, S.-K.; Shi, H.-Y.; Endale, B.; Kang, B.-S. Bayesian Optimization for Fine-Tuning EKF Parameters in UAV Attitude and Heading Reference System Estimation. *Aerospace* **2023**, *10*, 1023. <https://doi.org/10.3390/aerospace10121023>

Academic Editor: Jules Simo

Received: 18 October 2023

Revised: 6 December 2023

Accepted: 6 December 2023

Published: 9 December 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The extended Kalman filter (EKF) has been the most widely used algorithm for estimating the three-dimensional (3D) translational and rotational motion of unmanned aerial vehicles (UAVs). In aerospace in particular, the attitude of an air vehicle is commonly estimated from data obtained from multiple sensors such as accelerometers, gyroscopes, and magnetometers using the EKF algorithm. Furthermore, one of the critical features of EKF is its ability to fuse data from various sources based on their trustworthiness, impacting the value of the state under consideration. As a result, EKF has been the standard choice for sensor fusion applications. Nevertheless, the values of the essential EKF parameters are left up to the user's discretion, which makes EKF challenging to utilize. Proper tuning of four parameters must be identified for a Kalman filter to function as needed. Namely, the initial state value, X_0 , and the three noise parameters, which are the initial state estimation error covariance P_0 , the process noise covariance Q , and the measurement noise covariance R [1]. However, in several practical scenarios, Q and R are either unknown or only known approximately [2,3]. In other words, the Kalman formulation of the filtering issue presupposes perfect a priori knowledge of the process, and measurement noise statistics, which are seldom precise in practice [4]. In addition, the typical practice of using the matrices Q and R through trial-and-error selection can be time-consuming and may result in sub-optimal algorithm performance [1,5]. Therefore, the process noise covariance matrix Q and the measurement noise covariance matrix R must be tuned to their proper values for the Kalman filter to function as required by the user [2].

Numerous studies have been conducted on Kalman filter usage. However, there is yet to be a universally accepted approach in the literature for selecting appropriate Q and R

values [6]. Despite this, several positive advancements have been made toward developing easy-to-use methods for tuning the EKF parameters. In a broader sense, the solutions to the problem can be viewed in two ways. The most commonly employed method entails the utilization of initially accessible data, specifically about the non-linear behavior of the mathematical model (process noise) and sensor performance (measurement noise). Furthermore, ground reference data, which are often lacking, are required to evaluate the EKF performance with Q and R. If a need arises to further tune Q and R values, a very time-consuming manual trial and error process has to be followed. Consequently, the main limitation of such an approach lies in its inability to adapt to changes in system or sensor behaviors. Non-adaptive approaches have been applied in several research works; for instance, an auto-covariance least squares-based numerical tool was developed to compute the noise covariance [7–9]. Similarly, other works, like in [9,10], proposed methods to optimize the Kalman gain matrix and noise covariance matrices using the autocorrelation function of the residuals between the measurement and the related predicted state. However, in the aforementioned least-squares-based approaches, the nonlinearity of the system behavior could also limit their robustness. In a study by [11], the genetic algorithm was applied to find the optimal noise covariance matrices, which also relied on ground truth data for tuning.

On the other hand, other researchers devised adaptive noise-covariance matrix tuning techniques [12,13]. In these approaches, covariance matrices (CMs) Q and R values need to be tuned online while the EKF operates to improve the accuracy of the estimation, which can be affected by system and environmental disturbances. Mauro et al. [14] applied the recursive prediction error (RPE) algorithm to update the covariance values online. Machine learning approaches have gained increased attention across various application areas. In [15], an artificial neural network (ANN)-based learning algorithm was proposed to continuously monitor the KF estimation error and adjust the measurement noise covariance accordingly. Furthermore, Lbest particle swarm optimization was also used for power system applications to change the covariance values adaptively with performance changes [16]. Nonetheless, in most cases, achieving robustness and adaptability to a wider range of conditions was still a challenge. Moreover, the auto-tuning of the Kalman filter with Bayesian optimization was proposed by Chen et al. [17,18], where the EKF innovation statistics-based cost function was used for increased robustness, although the approach used to tweak the CM values was not feasible for higher dimensions.

This study suggested a Bayesian optimization method for tuning the R and Q values for the EKF-based UAV attitude and heading state estimation algorithm based on the EKF innovation whiteness analysis, considered an EKF estimation performance evaluation metric. The noise in each sensor measurement and the uncertainty in the process model were considered uncorrelated; therefore, all off-diagonal components of the covariance matrices Q and R are equal to zero. In addition, it is assumed that the variance of the sensor's measurement noise and the uncertainty of the process model, respectively, impact the diagonal elements of Q and R. Moreover, we confirmed that scaling the diagonal elements of the covariance matrices Q and R with different factors had an impact on the performance of the EKF. However, when we simultaneously scaled both Q and R by the same factor, there was only a minor difference in performance. Therefore, a Bayesian optimization technique was designed for tuning the optimum scaling factor pair corresponding to Q and R values. This approach of optimizing the EKF covariance matrix by adjusting scaling factors rather than tuning the individual elements of the covariance matrices independently makes our method scalable to systems of any size.

Bayesian optimization necessitates the definition of a search region to find optimal values. We applied a method that dynamically adjusts the search region size based on the optimization progress. The search region expands toward areas with high expected improvement and contracts from areas with low expected improvement. This approach ensures that the search is focused on the most promising regions.

Generally, at first, the initial R and Q values were computed from the noise statistics of raw sensor measurement data. Then, the suggested approach was employed to find the best scaling factor pair for R and Q matrices to improve the EKF performance to the peak. Our extensive tests on publicly available datasets revealed that the proposed approaches could easily tune the optimum R and Q more robustly and accurately than the conventional method-based approach. This approach is valuable for regular users and specialists who readily optimize EKF parameters for different application areas [19].

2. Mathematical Formulations

This section provides an overview of the mathematical methodologies utilized for estimating the 3D orientation of an object through the implementation of the extended Kalman filter (EKF) algorithm. Additionally, it discusses the advantages and challenges associated with employing the EKF algorithm in this specific area. The subsequent section will delve into the proposed solution for addressing the identified challenges.

2.1. Extended Kalman Filter (EKF) Equations

EKF is an iterative prediction/correction approach used for estimating the state of a discrete-time process or measurement. Consider a discrete-time dynamic system described by Equation (1).

$$\begin{aligned} x_k &= F_k x_{k-1} + \omega_k \\ z_k &= H_k x_k + v_k \end{aligned} \quad (1)$$

where x_k is the $n \times 1$ state vector, F_k is the $n \times n$ state transition matrix, ω_k is the $n \times 1$ process noise vector, z_k is the $r \times 1$ measurement vector, v_k is the $r \times 1$ measurement noise vector, and H_k is the $r \times n$ measurement matrix. In addition, the prediction error covariance, P_k , is presented in Equation (2).

$$P_k = F_k P_{k-1} F_k^T + Q \quad (2)$$

where F_k and Q represent the state transition and process noise covariance matrices, respectively. Then, at the last step of every single iteration of the EKF, the Kalman gain, K, is calculated, and the states and prediction error covariance matrix are updated as shown in Equation (3).

$$\begin{aligned} K &= P_k H_k^T (H_k P_k H_k^T + R)^{-1} \\ x_k^+ &= x_k + K(z_k - h(x_k)) \\ P_k^+ &= (I - K H_k) P_k \end{aligned} \quad (3)$$

where R and I stand for the measurement noise and identity matrices, respectively. The covariance matrices (CMs) Q and R used in Equations (2) and (3) account for the uncertainties due to noise in the process model and sensors measurement, respectively. Consequently, the values of R and Q should be selected cautiously so that the reality is perfectly represented. However, even when the process model and measurement function are precise, estimating the noise effects can be challenging. A combination of different factors often causes the noise [2], such as mis-modeled system and measurement dynamics; the presence of a hidden state in the environment that the EKF does not model; errors from the discretization of time; and approximations in the EKF, like the Taylor approximation commonly used for linearization. As a result, determining the ideal values of R and Q is challenging. To overcome a fundamental problem with the EKF, a fine-tuning Bayesian optimization-based algorithm was applied to identify the best Q and R covariance matrices.

2.1.1. Attitude Propagation Model

The attitude of a rigid object can be expressed in terms of Euler's three numbers (roll, pitch, and yaw) or quaternion's four numbers (q_w, q_x, q_y , and q_z) with respect to a given

reference frame. Several researchers have suggested that attitude estimation in quaternion form is commonly preferred because it avoids the gimbal lock problem. Consequently, the quaternion estimation algorithms are presented in this work. A vehicle's attitude and heading values are usually predicted from the angular velocity readings of a gyroscope sensor, which can potentially lead to drifting errors over time. In terms of the quaternion, the approximated and discrete forms of the attitude and heading equation are given in Equation (4).

$$q_k = (I_{4 \times 4} + \frac{1}{2} \Omega_k T_s) q_{k-1} \quad (4)$$

where

$$\Omega_k = \frac{1}{2} \begin{bmatrix} 0 & -\omega_x & -\omega_y & -\omega_z \\ \omega_x & 0 & \omega_z & -\omega_y \\ \omega_y & -\omega_z & 0 & \omega_x \\ \omega_z & \omega_y & -\omega_x & 0 \end{bmatrix}, [\omega_x] = \begin{bmatrix} 0 & -\omega_z & \omega_y \\ \omega_z & 0 & -\omega_x \\ -\omega_y & \omega_x & 0 \end{bmatrix}$$

and the gyroscope reading of angular velocity $\omega = [\omega_x \ \omega_y \ \omega_z]^T$. Also, q_k , T_s , and $I_{4 \times 4}$ represent the current time attitude prediction in the quaternion, sampling time, and identity matrix, respectively. The derivation for Equation (4) can be found in [20]. Due to the Taylor series approximation and discretization process applied in driving Equation (4), inaccuracy in the model is inevitable. Moreover, the noise in the gyroscope measurement aggravates the uncertainty of the prediction model, which imposes difficulty in determining the process noise covariance (Q).

Process Noise Covariance Calculation

Assume that the gyroscope is of the same type along all axes and that the manufacturer guarantees perfect orthogonality and uncorrelatedness between them. Then, the spectral noise covariance matrix is constructed as follows:

$$\Sigma_\omega = k_Q \begin{bmatrix} \sigma_{\omega_x}^2 & 0 & 0 \\ 0 & \sigma_{\omega_y}^2 & 0 \\ 0 & 0 & \sigma_{\omega_z}^2 \end{bmatrix} \quad (5)$$

where σ_{ω_x} , σ_{ω_y} and σ_{ω_z} represent the noise standard deviation of the gyroscope measurements along the X-, Y-, and Z-body axes, respectively. The scaling factor, k_Q , is added for the tuning purpose. The formula for Q is derived from the Jacobian of Equation (4) with respect to the gyroscope readings, as given in Equation (6); finally, Q can be expressed as shown in Equation (7).

$$\begin{aligned} \mathbf{W}_k &= \frac{\partial \mathbf{f}(\mathbf{q}_{k-1}, \omega_k)}{\partial \omega} \\ &= \left[\frac{\partial \mathbf{f}(\mathbf{q}_{k-1}, \omega_k)}{\partial \omega_x} \quad \frac{\partial \mathbf{f}(\mathbf{q}_{k-1}, \omega_k)}{\partial \omega_y} \quad \frac{\partial \mathbf{f}(\mathbf{q}_{k-1}, \omega_k)}{\partial \omega_z} \right] \\ &= \frac{T_s}{2} \begin{bmatrix} -q_1 & -q_2 & -q_3 \\ q_0 & -q_3 & q_2 \\ q_3 & q_0 & -q_1 \\ -q_2 & q_1 & q_0 \end{bmatrix} \end{aligned} \quad (6)$$

$$\begin{aligned} Q &= \frac{k_Q T_s^2}{4} \mathbf{W}_k \Sigma_\omega \mathbf{W}_k^T \\ &= \frac{k_Q T_s^2}{4} \begin{bmatrix} -q_1 & -q_2 & -q_3 \\ q_0 & -q_3 & q_2 \\ q_3 & q_0 & -q_1 \\ -q_2 & q_1 & q_0 \end{bmatrix} \begin{bmatrix} \sigma_{\omega_x}^2 & 0 & 0 \\ 0 & \sigma_{\omega_y}^2 & 0 \\ 0 & 0 & \sigma_{\omega_z}^2 \end{bmatrix} \begin{bmatrix} -q_1 & -q_2 & -q_3 \\ q_0 & -q_3 & q_2 \\ q_3 & q_0 & -q_1 \\ -q_2 & q_1 & q_0 \end{bmatrix}^T \end{aligned} \quad (7)$$

2.1.2. Attitude and Heading Observation Modeling

The gravitational field vector measured in the accelerometer's sensor body axes, $[a_x \ a_y \ a_z]^T$, can be mapped to the vertical gravity field ($[0 \ 0 \ g]^T$) in the NED frame using the rotational matrix, described in terms of quaternion parameters, denoted as $C_b^n(q)$. Similarly, the magnetic field measured by the magnetometer ($[m_x \ m_y \ m_z]^T$) was also mapped to the horizontal and vertical components of the Earth's magnetic field ($[h_x \ h_y \ h_z]^T$). The cosine direction matrix, which is used to rotate the accelerometer and magnetometer body frame readings to align with the Earth's gravitational and magnetic fields in the NED frame, respectively, is presented in Equations (8).

$$C_b^n(q) = \begin{bmatrix} q_0^2 + q_1^2 - q_2^2 - q_3^2 & 2(q_1q_2 + q_0q_3) & 2(q_1q_3 - q_0q_2) \\ 2(q_1q_2 - q_0q_3) & q_0^2 - q_1^2 + q_2^2 - q_3^2 & 2(q_2q_3 + q_0q_1) \\ 2(q_1q_3 + q_0q_2) & 2(q_2q_3 - q_0q_1) & q_0^2 - q_1^2 - q_2^2 + q_3^2 \end{bmatrix} \quad (8)$$

The measurement modeling for EKF-based attitude estimation has been approached in different ways in different research works [21–23]. In the paper by Robert V. et al. [22], the quaternion parameters were considered as the observation states, even though they cannot be directly observed by the sensors. Therefore, an indirect observation of the quaternion from the accelerometer readings (a_x, a_y, a_z) and magnetometer readings (m_x, m_y, m_z) was applied, as given in Equation (9). Consequently, the measurement transition matrix (H) is equal to the identity matrix ($I_{4 \times 4}$).

$$q_k = \begin{bmatrix} q_w \\ q_x \\ q_y \\ q_z \end{bmatrix} = q_{acc} \otimes q_{mag} \quad (9)$$

where

$$q_{acc} = \begin{cases} \begin{bmatrix} \lambda_1 & -\frac{a_y}{2\lambda_1} & \frac{a_x}{2\lambda_1} & 0 \end{bmatrix}^T, & a_z \geq 0 \\ \begin{bmatrix} -\frac{a_y}{2\lambda_2} & \lambda_2 & 0 & \frac{a_x}{2\lambda_2} \end{bmatrix}^T, & a_z < 0 \end{cases}, \quad \lambda_1 = \sqrt{\frac{a_z + 1}{2}}, \quad \lambda_2 = \sqrt{\frac{1 - a_z}{2}}$$

$$q_{mag} = \begin{cases} \begin{bmatrix} \frac{\sqrt{h_N + h_x \sqrt{h_N}}}{\sqrt{2h_N}} & 0 & 0 & \frac{h_y}{\sqrt{2}\sqrt{h_N + h_x \sqrt{h_N}}} \end{bmatrix}^T, & h_x \geq 0 \\ \begin{bmatrix} \frac{h_y}{\sqrt{2}\sqrt{h_N - h_x \sqrt{h_N}}} & 0 & 0 & \frac{\sqrt{h_N - h_x \sqrt{h_N}}}{\sqrt{2h_N}} \end{bmatrix}^T, & h_x < 0 \end{cases},$$

$$\begin{bmatrix} h_x \\ h_y \\ h_z \end{bmatrix} = C_b^n(q_{acc}) \left(\frac{1}{\sqrt{m_x^2 + m_y^2 + m_z^2}} \begin{bmatrix} m_x \\ m_y \\ m_z \end{bmatrix} \right), \quad h_N = \sqrt{h_x^2 + h_y^2}$$

Similarly, in the paper by Guo S. et al. [23], the quaternion parameters were considered as the observation states, but a different mathematical model was used for describing the quaternions in terms of the accelerometer and magnetometer readings, as indicated in Equation (10).

$$q_k = \frac{1}{4}(W_a + I_{4 \times 4})(W_m + I_{4 \times 4})q_{k-1} \quad (10)$$

where

$$W_a = \begin{pmatrix} a_z & a_y & -a_x & 0 \\ a_y & -a_z & 0 & a_x \\ -a_x & 0 & -a_z & a_y \\ 0 & a_x & a_y & a_z \end{pmatrix},$$

$$W_m = \begin{pmatrix} h_N m_x + h_z m_z & h_z m_y & h_N m_z - h_z m_x & -h_N m_y \\ h_z m_y & h_N m_x - h_z m_z & h_N m_y & h_N m_z + h_z m_x \\ h_N m_z - h_z m_x & h_N m_y & -h_N m_x - h_z m_z & h_z m_y \\ -h_N m_y & h_N m_z + h_z m_x & h_z m_y & -h_N m_x + h_z m_z \end{pmatrix},$$

$$h_z = \left(\frac{1}{\sqrt{a_x^2 + a_y^2 + a_z^2}} \begin{bmatrix} a_x \\ a_y \\ a_z \end{bmatrix} \right) \cdot \left(\frac{1}{\sqrt{m_x^2 + m_y^2 + m_z^2}} \begin{bmatrix} m_x \\ m_y \\ m_z \end{bmatrix} \right), \quad h_N = \sqrt{1 - h_z^2},$$

and $I_{4 \times 4}$ is the identity matrix. In another paper [21], the sensor's measurement output values were directly used as the observation states. In this case, the measurement model transition matrix is a non-linear function in terms of the quaternions, as shown in Equation (11).

$$\begin{bmatrix} a_x \\ a_y \\ a_z \\ m_x \\ m_y \\ m_z \end{bmatrix} = \underbrace{\begin{bmatrix} C_b^n(q)^T & 0_{3 \times 3} \\ 0_{3 \times 3} & C_b^n(q)^T \end{bmatrix}}_H \begin{bmatrix} 0 \\ 0 \\ g \\ h_N \\ 0 \\ h_z \end{bmatrix} \quad (11)$$

Measurement Noise Covariance Calculation

In most scientific articles, the measurement noise covariance is assumed to be time-invariant and freely accessible from the measuring device's datasheet [24]. If σ is the measurement's known time-invariant nominal standard deviation, a straightforward approach to express the measurement noise covariance matrix R is

$$R = k_R J \Sigma_{a,m} J^T$$

$$\Sigma_{a,m} = \text{diag}(\sigma_{a_x}^2 \quad \sigma_{a_y}^2 \quad \sigma_{a_z}^2 \quad \sigma_{m_x}^2 \quad \sigma_{m_y}^2 \quad \sigma_{m_z}^2) \quad (12)$$

where k_R and J represent the scaling factor value and the Jacobian of the observation states with respect to the sensors' measurement (i.e., $\mathbf{u} = [a_x, a_y, a_z, m_x, m_y, m_z]$). Alternatively, experimental data can also be used to compute the noise variance. If measurements of the accelerometer and magnetometer were taken at rest conditions, the variance could be calculated using Equations (13) and (14).

$$\sigma_a^2 = \frac{1}{N-1} \begin{bmatrix} \sum_{i=1}^N a_{x,i}^2 \\ \sum_{i=1}^N a_{y,i}^2 \\ \sum_{i=1}^N (a_{z,i} - g)^2 \end{bmatrix} \quad (13)$$

where σ_a , a_x , a_y , a_z , g , and N represent the accelerometer standard deviation; acceleration along the X-, Y-, and Z-body axes, and gravity and the total number of samples, respectively. And

$$\sigma_m^2 = \frac{1}{N-1} \begin{bmatrix} \sum_{i=1}^N (m_{x,i} - m_{x,\mu})^2 \\ \sum_{i=1}^N (m_{y,i} - m_{y,\mu})^2 \\ \sum_{i=1}^N (m_{z,i} - m_{z,\mu})^2 \end{bmatrix} \quad (14)$$

where σ_m , $m_{x,y,z}$, and m_{μ} represent the magnetometer measurement's standard deviation, the Earth's magnetic field along the X-, Y-, and Z-body axes, and the mean of the measured magnetic field, respectively.

However, the function to find the value of J depends on the chosen observation states. For the measurement modeling approaches presented in Equations (9) and (10), J was calculated using Equation (15).

$$J = \frac{\partial q_k}{\partial u} = \begin{bmatrix} \frac{\partial q_w}{\partial a_x} & \frac{\partial q_w}{\partial a_y} & \frac{\partial q_w}{\partial a_z} & \frac{\partial q_w}{\partial m_x} & \frac{\partial q_w}{\partial m_y} & \frac{\partial q_w}{\partial m_z} \\ \frac{\partial q_x}{\partial a_x} & \frac{\partial q_x}{\partial a_y} & \frac{\partial q_x}{\partial a_z} & \frac{\partial q_x}{\partial m_x} & \frac{\partial q_x}{\partial m_y} & \frac{\partial q_x}{\partial m_z} \\ \frac{\partial q_y}{\partial a_x} & \frac{\partial q_y}{\partial a_y} & \frac{\partial q_y}{\partial a_z} & \frac{\partial q_y}{\partial m_x} & \frac{\partial q_y}{\partial m_y} & \frac{\partial q_y}{\partial m_z} \\ \frac{\partial q_z}{\partial a_x} & \frac{\partial q_z}{\partial a_y} & \frac{\partial q_z}{\partial a_z} & \frac{\partial q_z}{\partial m_x} & \frac{\partial q_z}{\partial m_y} & \frac{\partial q_z}{\partial m_z} \end{bmatrix} \quad (15)$$

where the J value corresponding to Equation (11) is equal to the identity matrix ($I_{6 \times 6}$) since the observation states chosen were identical to the sensors' measurement output.

2.2. Bayesian Optimization

Bayesian optimization is an iterative process that begins with some prior beliefs about the objective function to be estimated, such as its smoothness and other characteristics. Over time, it collects more evidence through an acquisition function to update its initial beliefs about the objective function. These prior beliefs are encoded in a probability distribution function. The most widely used distribution function in Bayesian optimization is the Gaussian process (GP) function. A Gaussian process is a collection of random variables, each of which, when taken in any finite linear combination, follows a multivariate normal distribution.

2.2.1. Gaussian Process (GP) Regression

A Gaussian process is a (potentially infinite) collection of random variables, where the joint distribution of every finite subset of these random variables is a multivariate Gaussian distribution:

$$f(\mathbf{x}) \sim \mathcal{GP}(m(\mathbf{x}), k(\mathbf{x}, \mathbf{x}')). \quad (16)$$

Initially, the mean was assumed to be zero for simplicity. The covariance (kernel) function, represented by $k(\mathbf{x}, \mathbf{x}')$, models the joint variability of the Gaussian process random variables. Preceding any data points observation, an infinite number of candidate functions (prior) can fit the initial mean ($m(\mathbf{x})$) and covariance($k(\mathbf{x}, \mathbf{x}')$). The prior distribution represented the expected outputs of the function over inputs \mathbf{x} without any observation. When we start to have observations, only functions that fit the observed data points are retained, referred to as a posterior distribution. With new additional observations, the current posterior was updated continuously toward improving the model function until a stopping criterion was met. The joint distribution of the training output, f , and test data output, f_* , with noisy observation, is described as follows:

$$\begin{bmatrix} y \\ f_* \end{bmatrix} = \mathcal{N}\left(0, \begin{bmatrix} K(X, X) + \sigma_n^2 I & K(X, X_*) \\ K(X_*, X) & K(X_*, X_*) \end{bmatrix}\right) \quad (17)$$

A conditional probability is applied to obtain the posterior distribution over a function that agrees with the observed data points.

$$f_* | X, y, X_* \sim \mathcal{N}(\bar{f}_*, cov(f_*)), \quad \text{where}$$

$$\bar{f}_* \stackrel{\Delta}{=} \mathbb{E}[f_* | X, y, X_*] = K(X_*, X)[K(X, X) + \sigma_n^2 I]^{-1}y, \quad (18)$$

$$cov(f_*) = K(X_*, X_*) - K(X_*, X)[K(X, X) + \sigma_n^2 I]^{-1}K(X, X_*).$$

One of the most common kernels in modeling Gaussian processes is the exponentiated quadratic kernel, also known as the Gaussian kernel, mathematically presented in Equation (19).

$$K(X_1, X_2) = \sigma^2 \exp\left(-\frac{\|X_1 - X_2\|^2}{2\ell^2}\right) \quad (19)$$

where σ and l represent the variance and length scale, respectively.

2.2.2. Acquisition Function

The Bayesian optimization algorithm starts with preliminary information about the function to be modeled and optimized. It updates the objective function progressively as new knowledge comes in from the newly sampled data. The sampling point was bravely selected using the expected improvement acquisition function. The expected improvement acquisition function is one of the most widely selected methods among other acquisition function techniques for Bayesian optimization implementation. A good acquisition function should trade off exploration and exploitation. The acquisition function considers the expected mean and variance at each point along the objective function domain to compute the value that indicates how desirable it is to sample next at that position. Suppose we would like to minimize $f(x)$, and the best solution so far is x^* , then, the equation of expected improvement (EI) is defined as Equation (20).

$$EI(x; \xi) = \begin{cases} (f(x^*) - \mu(x) - \xi)\Phi\left(\frac{f(x^*) - \mu(x) - \xi}{\sigma(x)}\right) + \sigma(x)\phi\left(\frac{f(x^*) - \mu(x) - \xi}{\sigma(x)}\right), & \sigma(x) > 0 \\ 0, & \sigma(x) \leq 0 \end{cases} \quad (20)$$

Equation (20) can be represented in a simplified form as in Equation (21).

$$EI(z) = \begin{cases} \sigma(x)(z\Phi(z) + \phi(z)), & \sigma(x) > 0 \\ 0, & \sigma(x) \leq 0 \end{cases} \quad (21)$$

where $z = \frac{d}{\sigma(x)}$ and $d = f(x^*) - \mu(x) - \xi$. The symbols d , σ , μ , ξ , Φ , and ϕ represent the difference between the predicted and observed mean, the expected mean uncertainty, expected mean, trade-off control for exploration and exploitation, normal probability distribution function (pdf), and normal cumulative distribution function (cdf), respectively. The EI algorithm was further elaborated graphically in Figure 1.

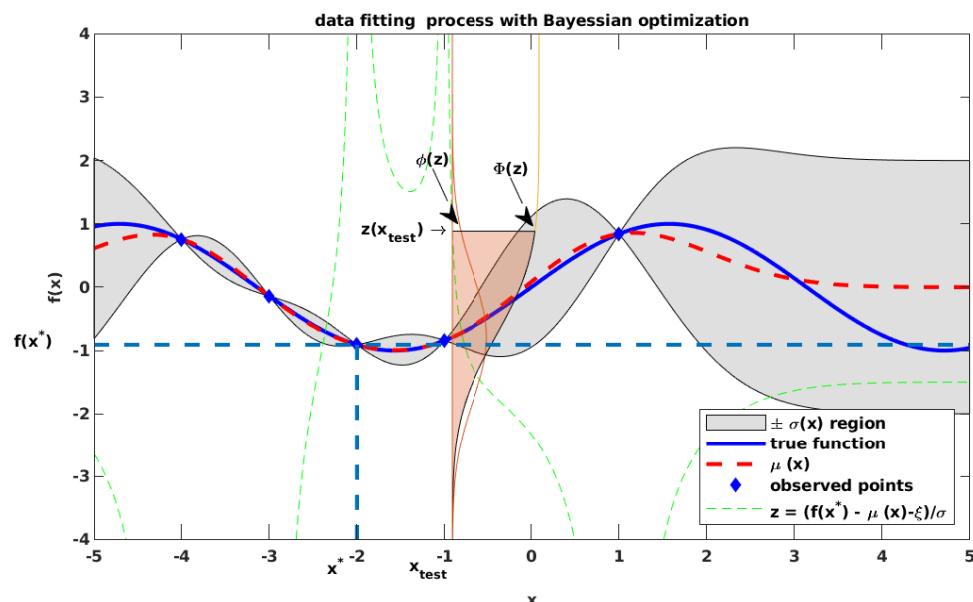


Figure 1. Cont.

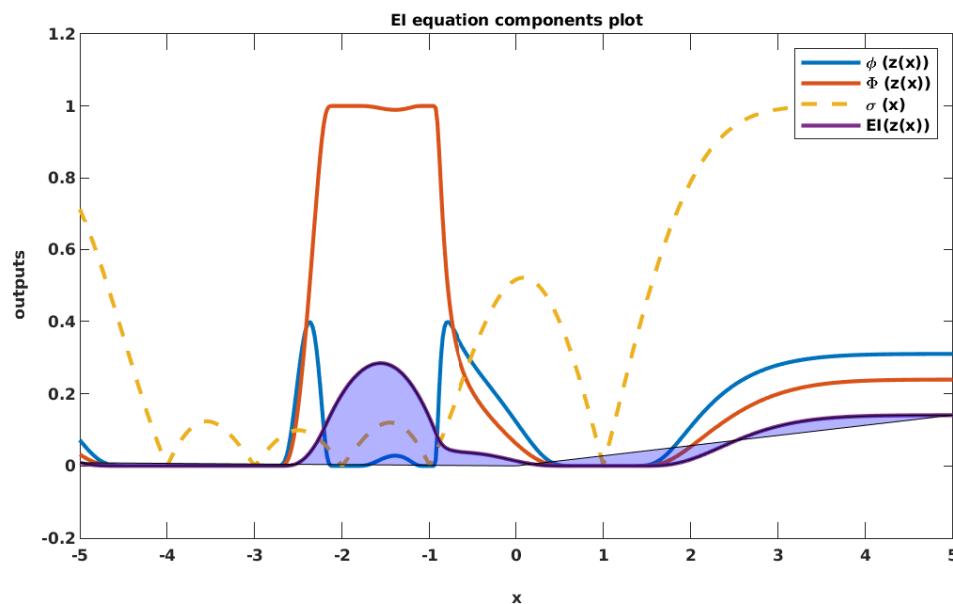


Figure 1. The graphical illustration of the Bayesian optimization algorithm, showing the sample data points, expected function, uncertainty region shown in a gray area, true function, normal probability distribution function indicated by $\phi(z)$ corresponding to the test point x_{test} at the top image, normal cumulative distribution function shaded in light red and indicated by $\Phi(z)$ corresponding to the test point at the top image, EI shown in a light blue area, and the functions plot of EI's component computed corresponding to the search space shown at the bottom image.

3. Process and Measurement Noise Covariance Tuning With Bayesian Optimization

The system block diagram for tuning R and Q is shown in Figure 2. The system consists of three sub-blocks: the EKF sub-block, the performance evaluator sub-block, and the Bayesian-based optimization sub-block.

The EKF sub-block continuously predicted the attitude and heading of the UAV using the gyroscope's readings. The prediction was then updated with the indirect measurement of attitude and heading from the accelerometer and magnetometer readings. The performance evaluator sub-block buffered the deviation between the sensor's measurements and the predicted referred to as innovation.

The proposed algorithm's implementation necessitates buffering a sequence of EKF innovation values for a specific duration N. Then, the autocorrelation of the innovation data was computed for several time lags, which helped to find the optimization cost function that is explained more in Section 3.2. Furthermore, the Bayesian optimization algorithm begins by considering the initially trained data of the k_R and k_Q pair along with the cost values to compute the prior expected function and its uncertainty. Then, new data sampling was conducted through the expected improvement (EI) acquisition function guidance within the allowed k_R and k_Q search domains. Further discussion about the implementation of EI is found in Section 2.2.2. Similarly, the posterior expected mean and the corresponding uncertainty computation followed by incorporating the new sample data. This process iteration continued until the optimum k_R and k_Q values were found.

3.1. Adaptive Search Region for Bayesian Optimization

Bayesian optimization is a powerful technique for optimizing functions, but the size of the search space can significantly impact its efficiency. To address this challenge, a method was developed that leverages historical information about the objective function's performance over a specific interval of iterations. This method intelligently adapts the search space to improve the optimization process.

The approach identifies regions within the search space where promising optimal values have previously been discovered. It then makes precise adjustments to the search

space size based on these insights. These adjustments can take the form of expansion, contraction, or zooming, depending on certain conditions.

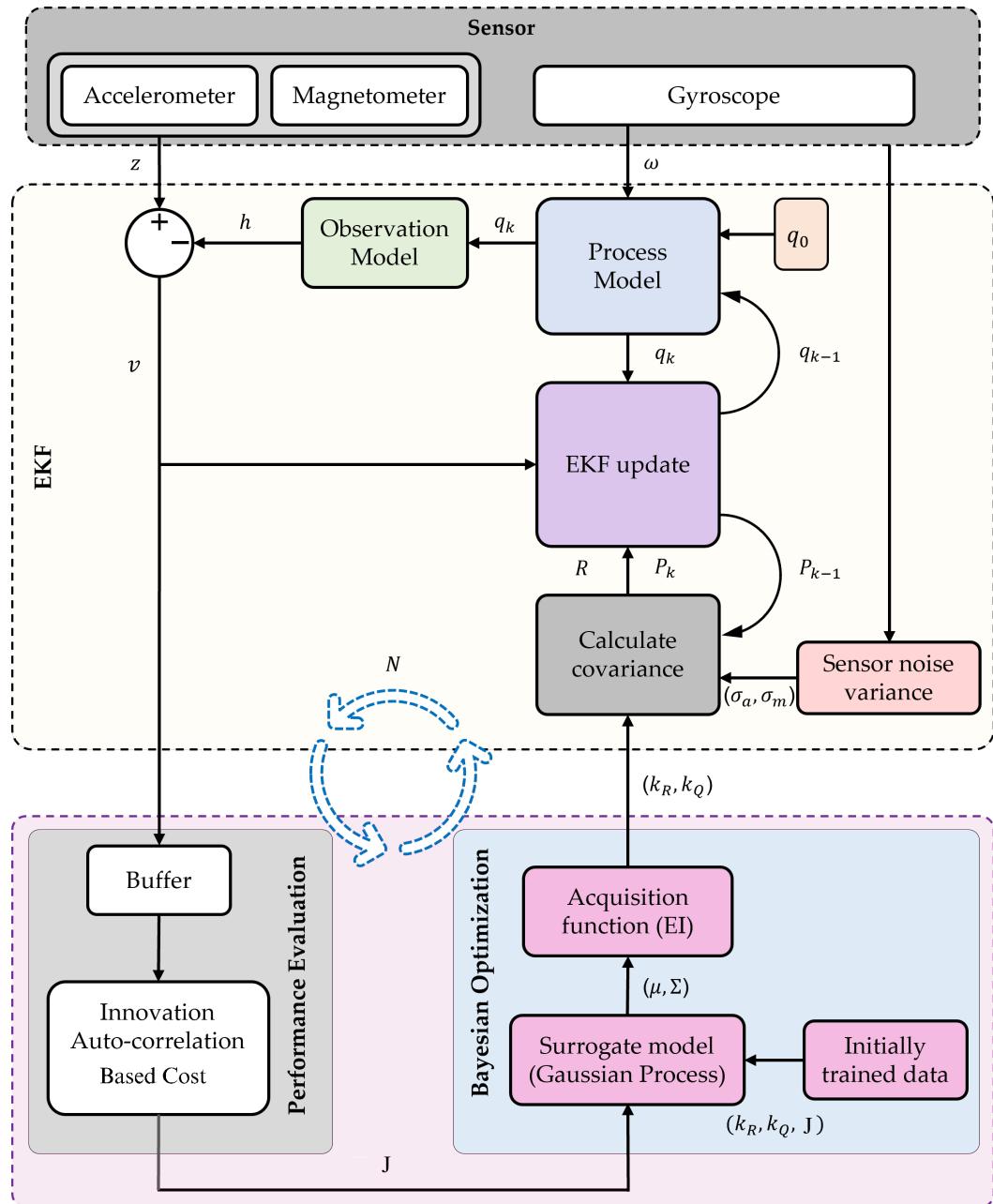


Figure 2. EKF noise covariance matrix tuning algorithm block diagram.

If the minimum sampled values found so far have been at and near the boundary of the search area, the method would expand the search space from that boundary side, given that relatively higher sampled values have been observed at the opposite boundary of the search space. On the other hand, if the sampled values have been relatively high at and near the edge of the search space, the algorithm would contract the search space from that side. In cases where the minimum sampled values are concentrated around the middle of the search space, and both of the opposite edges exhibit maximum values, the method would employ a “zooming” strategy. This proactive adjustment ensures that the algorithm explores potentially fruitful regions more extensively.

In summary, the method offers a dynamic approach to adapting the search space in Bayesian optimization, harnessing historical knowledge to guide the optimization process

effectively. By expanding, contracting, or zooming the search space intelligently, it can expedite the discovery of optimal solutions while conserving computational resources.

3.2. Optimization Cost Function

In most cases, knowing the underlying truth about the system state for performance comparison is difficult or impossible. As a result, in this research, the EKF performance measures were analyzed in terms of innovation, which is always available and easy to access. If proper EKF parameters have been set, innovation denoted by ν_k should be zero mean and white with a covariance of S_k . Using Equation (22), we can test the whiteness of the innovation by checking its autocorrelation (R_{xx}) which should be zero at all lag times except where $\tau = 0$, within an allowable statistical error.

$$R_{xx}(\tau) = \frac{1}{R_{xx}(0)} \sum_{k=0}^{N-\tau-1} \nu_k^T \nu_{k+\tau} \quad (\tau \geq 0) \quad (22)$$

where τ and N represent the time lag and total number of EKF innovation samples. In practice, the autocorrelation of a white EKF innovation at zero lag should be unity for the normalized case and nearly zero as the lag increases from zero in both directions symmetrically. As a consequence, the sum of autocorrelation values at zero and near-zero lags will be significantly larger compared to the sum at the remaining lags for signals that are more white than for those that are less white.

Therefore, the optimization effort is equivalent to pushing all the autocorrelation values at non-zero lags to zero, as illustrated in Figure 3, to narrow and center the function graph at zero lag. As a result, the good cost function (J) for the Bayesian optimization should be the ratio of the area under the autocorrelation function within the time lag ranges $[-m, -L]$ and $(L, m]$ to the area under the curve within the time lag range $[-L, L]$, which is described in Equation (23). The L value can be selected to map the cost value at a manageable range.

$$J = \frac{A_2 + A_3}{A_1} = \frac{\sum_{\tau=-m}^{-L-1} R_{xx}(\tau) + \sum_{\tau=L+1}^m R_{xx}(\tau)}{\sum_{\tau=-L}^L R_{xx}(\tau)} \quad (23)$$

where A and m represent the shaded area under the curve at the corresponding time range and the choice of maximum lag for whiteness testing, respectively. Minimizing the cost function corresponds to choosing covariance matrices R and Q that produce whiter (narrower) EKF innovations.

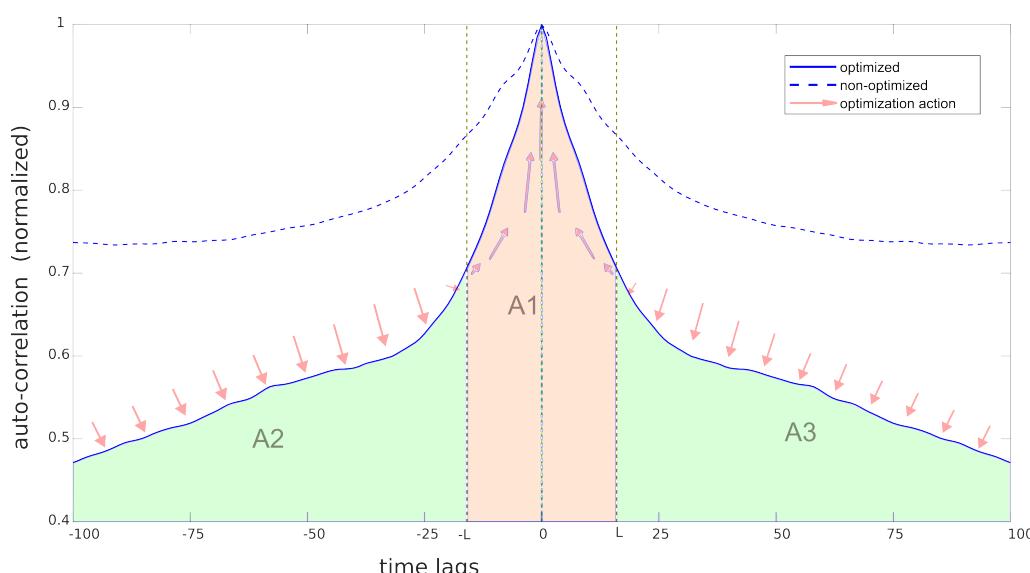


Figure 3. Autocorrelation function graph shape comparison and reshaping action illustration.

4. Testing

Our algorithm's effectiveness was tested based on two publicly available datasets published by Laidig et al. [25] and Caruso M. et al. [26]. The specifications of the sensors used to obtain the datasets are summarized in Table 1.

Table 1. IMU sensors datasheet noise specification.

Sensors Specs	Datasets		
	BOARD [25]		SASARI [26]
Accel noise ($\sigma_{a_{std}}$)	[0.044 0.050 0.074] m/s ²		[0.86 0.80 0.85] m/s ²
Gyro noise ($\sigma_{\omega_{std}}$)	[0.10 0.09 0.12] deg/s		[0.38 0.39 0.37] deg/s
Mag noise ($\sigma_{m_{std}}$)	[0.71 0.70 0.68] μ T		[0.06 0.04 0.04] μ T
Model	myon aktos-t		Xsens-MTx

Based on the noise information from the given sensors, Σ_w is represented as shown in Equation (24).

$$\begin{aligned}\Sigma_w^{\text{BOARD}} &= k_Q \left(\frac{\pi}{180} \begin{bmatrix} 0.10 & 0 & 0 \\ 0 & 0.09 & 0 \\ 0 & 0 & 0.12 \end{bmatrix} \right) \\ \Sigma_w^{\text{SASARI}} &= k_Q \left(\frac{\pi}{180} \begin{bmatrix} 0.38 & 0 & 0 \\ 0 & 0.39 & 0 \\ 0 & 0 & 0.37 \end{bmatrix} \right)\end{aligned}\quad (24)$$

Similarly, the value of $\Sigma_{a,m}$ mentioned in Equation (12) can be obtained by referencing Table 1, applying the necessary conversions to ensure compatibility with the measurement units and accounting for any scaling caused by normalization.

4.1. Test Results

This section presents the results that justify the selection of the proposed method and its performance. As stated in Section 3.2, testing the performance of the EKF-based AHRS estimation method for every process noise and measurement noise covariance combination, which in this study is controlled by k_R and k_Q values, would be exceedingly time- and resource-intensive. Additionally, ground truth data may only be available in some real-time operations for comparison. Therefore, we employed the Bayesian optimization technique to efficiently explore the parameter space defined by k_R and k_Q . By sampling a limited number of data points within this range, a mathematical model was constructed that accurately captured the underlying trends. Through iterative and guided sampling, the optimal point was quickly found.

Furthermore, as reference data would typically not be available to calculate the attitude and heading estimation error in most real-time applications, the attitude and heading RMS error values were not used to optimize the measurement and process noise parameters. Instead, our algorithm relied on innovation data, which are always available. To test the proposed algorithm, an AHRS developed by [23] and the datasets presented in Table 1 were used. The results obtained for the two datasets are presented as follows.

4.1.1. Case BOARD Dataset:

Figures 4–6 show the Bayesian-based data fitting and optimization results to determine the ideal k_R and k_Q values for iteration numbers twenty-one, seventy-one, and one hundred seventy-four, respectively.

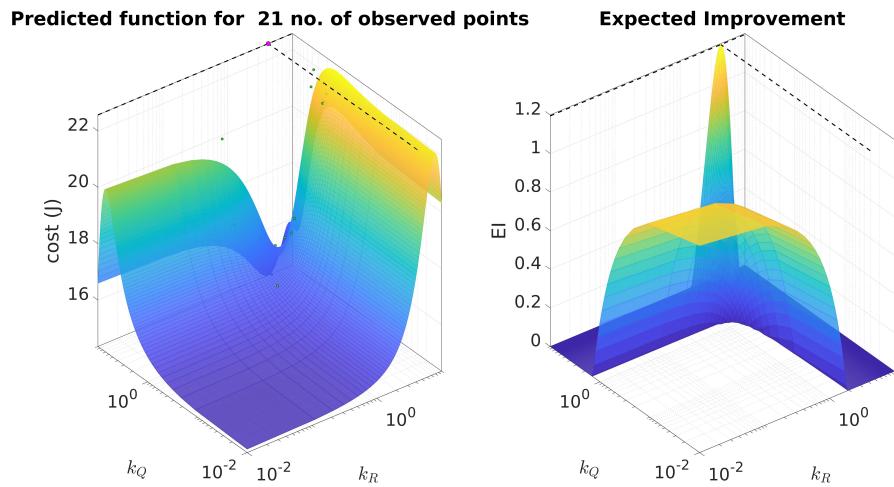


Figure 4. Bayesian optimization progress for the initial 21 samples, displaying the 3D plots of the objective function prediction and corresponding expected improvement results for the specified ranges of k_R and k_Q values.

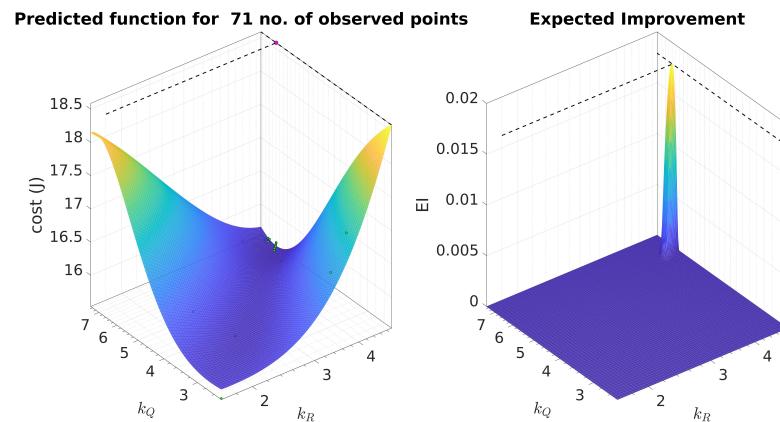


Figure 5. Bayesian optimization progress for the 71 samples, displaying the 3D plots of the objective function prediction and corresponding expected improvement results for the specified ranges of k_R and k_Q values.

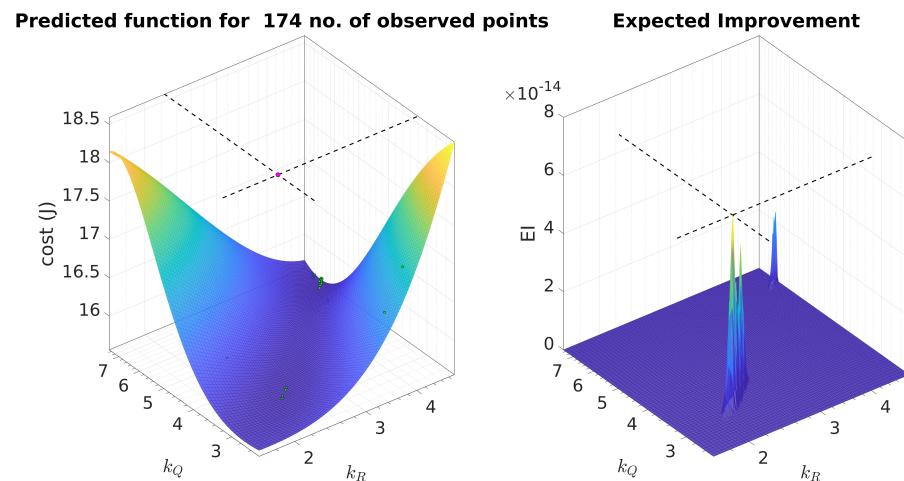


Figure 6. Bayesian optimization progress for the 174 samples, displaying the 3D plots of the objective function prediction and corresponding expected improvement results for the specified ranges of k_R and k_Q values.

The attitude and heading estimation accuracy improvement when optimized EKF parameters were used, in terms of angle and quaternion errors, are also depicted in Figures 7 and 8, respectively.

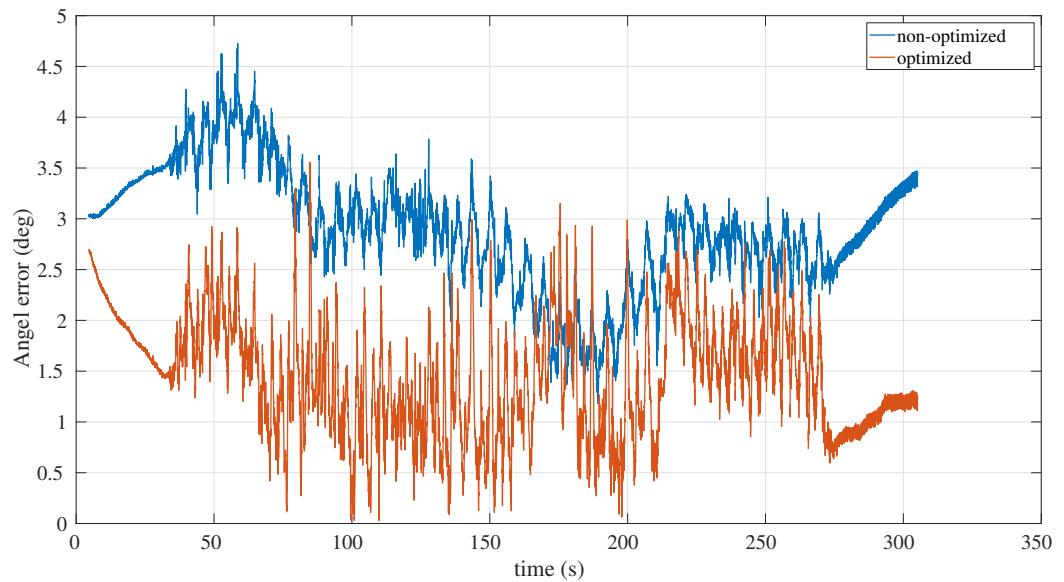


Figure 7. The angle estimation error with respect to the reference for non-optimized and optimized EKF parameters.

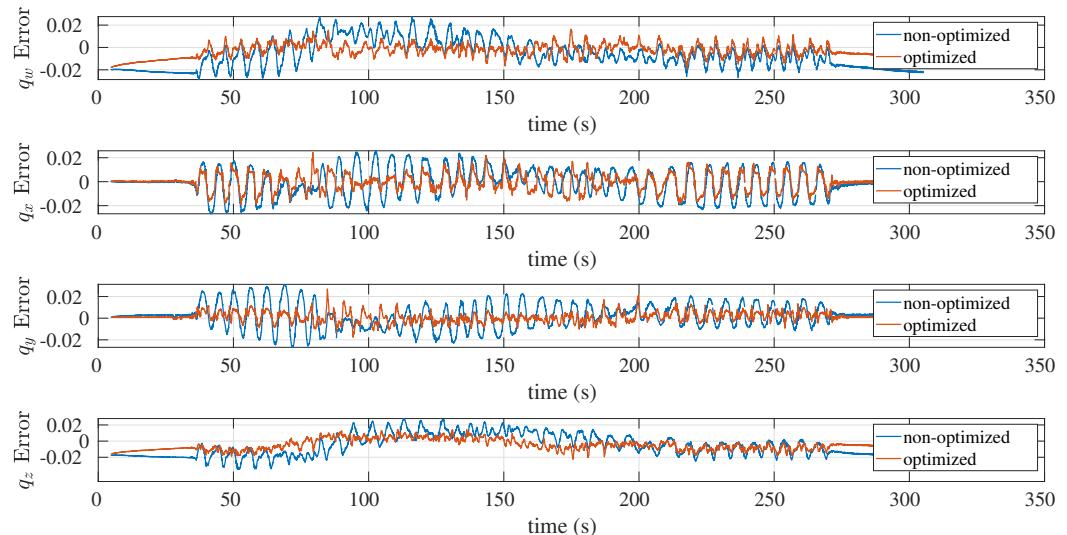


Figure 8. Quaternion estimation errors with respect to the references for non-optimized and optimized EKF parameters.

4.1.2. Case SASARI Dataset:

Figures 9–11 show the Bayesian-based data fitting and optimization results to determine the ideal k_R and k_Q values for iteration numbers twenty-one, seventy-one, and eighty-five, respectively.

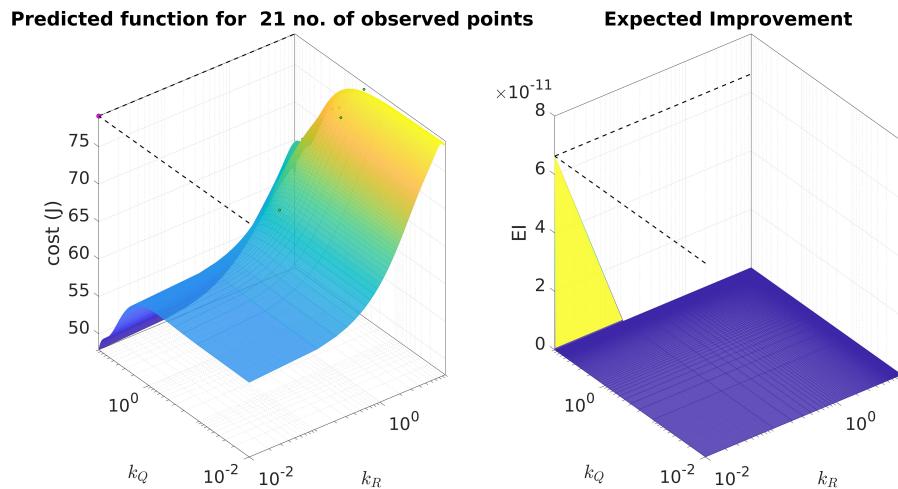


Figure 9. Bayesian optimization progress for the initial 21 samples, displaying the 3D plots of the objective function prediction and corresponding expected improvement results for the specified ranges of k_R and k_Q values.

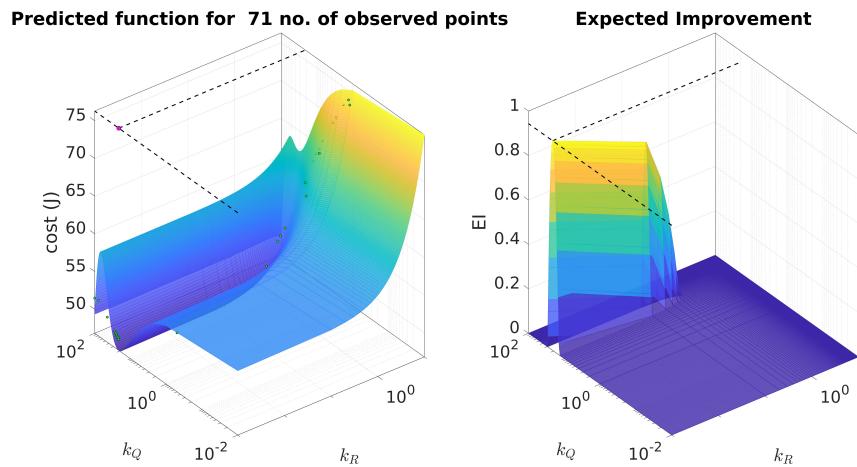


Figure 10. Bayesian optimization progress for the 71 samples, displaying the 3D plots of the objective function prediction and corresponding expected improvement results for the specified ranges of k_R and k_Q values.

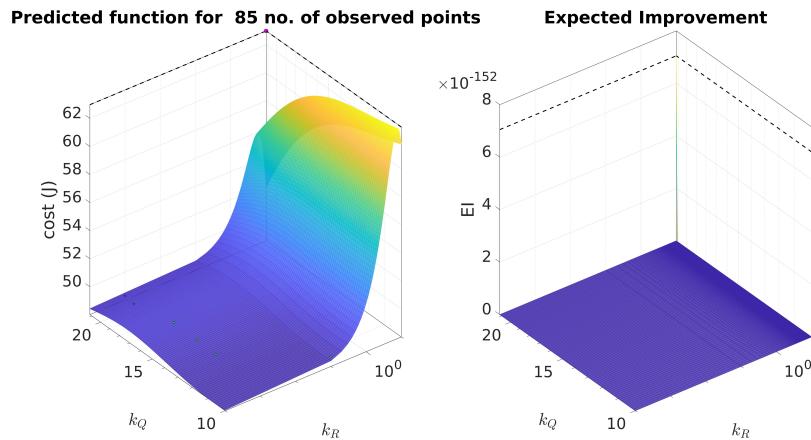


Figure 11. Bayesian optimization progress for the 85 samples, displaying the 3D plots of the objective function prediction and corresponding expected improvement results for the specified ranges of k_R and k_Q values.

The attitude and heading estimation accuracy improvement when optimized EKF parameters were used, in terms of angle and quaternion errors, are depicted in Figures 12 and 13, respectively.

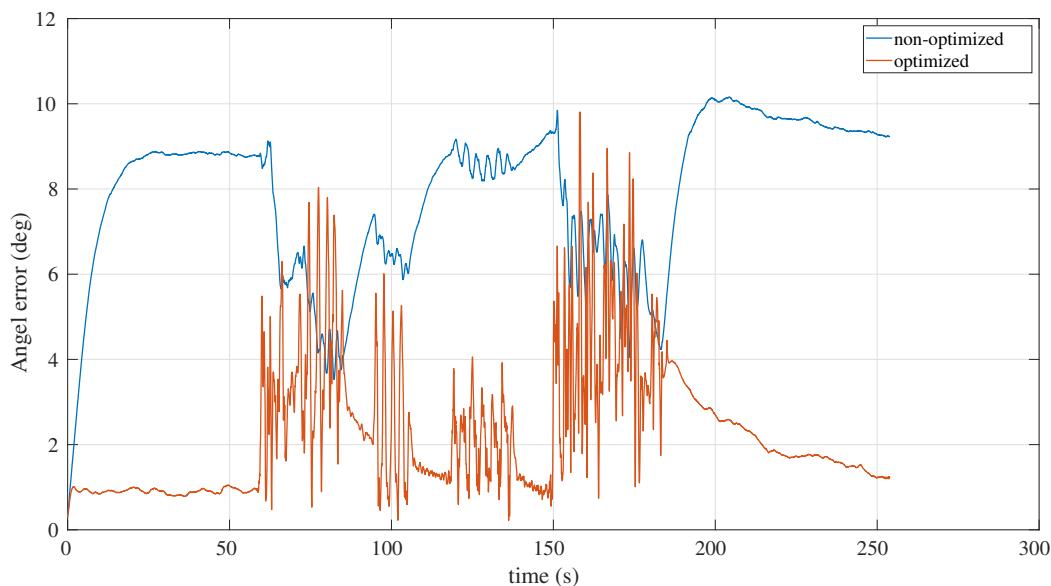


Figure 12. Angle estimation error with respect to the references for non-optimized and optimized EKF parameters.

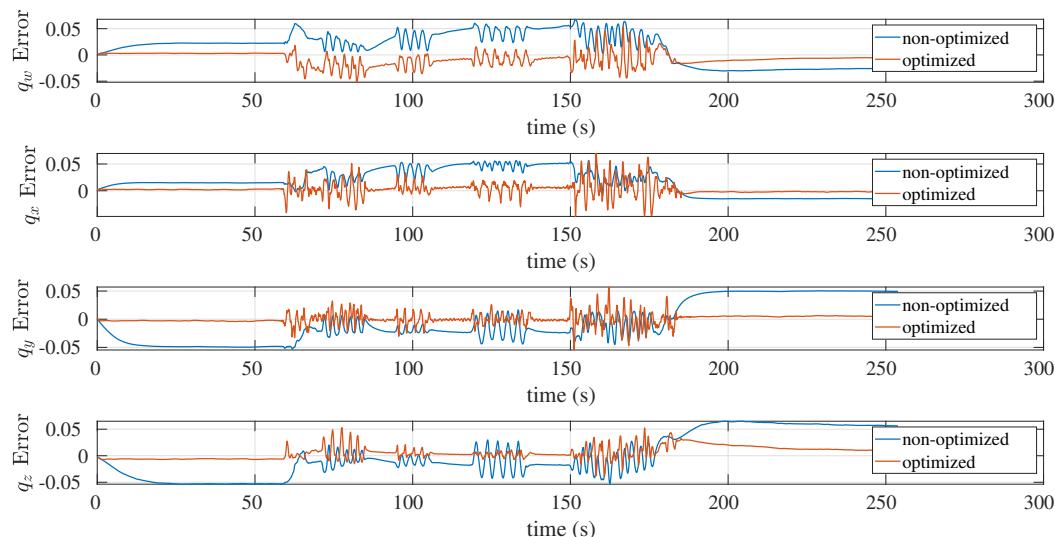


Figure 13. Quaternion estimation errors with respect to the references for non-optimized and optimized EKF parameters.

5. Discussion

To validate and assess the effectiveness of our proposed approach in computing the optimum values of k_R and k_Q , we utilized openly available published datasets, along with the recently released attitude and heading reference system (AHRS) algorithm by Guo S. et al. [23]. The rationale behind employing different datasets lies in recognizing the inherent variability introduced by data originating from different devices and recorded under varying conditions, leading to distinct noise behaviors. By subjecting our optimization algorithm to such heterogeneous datasets, we aim to enhance the robustness and credibility of the results obtained. For this purpose, we consider the BOARD [25] and SASARI [26] datasets, each acquired using the sensor specifications, as shown in Table 1.

At the core of our research lies a Bayesian-based optimization algorithm. The algorithm starts by generating twenty initial samples, randomly selected from the specified range of k_R and k_Q values. Subsequently, the acquisition function intelligently explores the parameter space to identify the most optimistic k_R and k_Q pair for the next test point. Throughout the iterative process, the algorithm progressively narrows the search space, ultimately converging on the optimal k_R and k_Q pair in less than two hundred iterations. This systematic approach ensures that the best pair is efficiently determined, thereby contributing to the overall effectiveness of our proposed methodology.

Figures 4–6 depict the progress of the objective function prediction along with the output of the expected improvement (EI) data acquisition function at the given search space for iteration numbers twenty-one, seventy-one, and one hundred seventy-four, respectively, while utilizing the dataset BOARD. The crossing-dotted line in each figure indicates the maximum value of EI, signifying the next sampling point. In particular, Figures 5 and 6 showcase the continuous improvement of the estimated model function as the number of samples gradually increases. Remarkably, even with a relatively small number of iterations, a nearly accurate objective function model is obtained, which closely represents the observed data. Notably, the model obtained at the seventy-one iteration demonstrates an excellent representation of the observed data. Similarly, Figures 9–11 show the progress of the optimization process when the SASARI dataset was utilized. The findings illustrate how the application of Bayesian optimization allowed us to efficiently accomplish a task that typically required a long time and considerable effort, using only minimal samples. The optimal k_R and k_Q values discovered at the eighty-fifth iteration have excellent EKF performance, as was validated with experimental test results depicted in Figures 7, 8, 12, and 13 for two dataset utilization cases. The accuracy of attitude and heading estimation improved markedly when using the optimized k_R and k_Q values, compared to when non-optimized parameters (i.e., $k_R = 1$ and $k_Q = 1$) were used.

The performance comparison of EKF for optimized and non-optimized values of k_R and k_Q are graphically presented in Figures 7, 8, 12, and 13, and summarized in Table 2.

Table 2. RMS error of attitude and heading estimation in quaternion and axis angle representations analyzed using BOARD and SASARI datasets for non-optimized and optimized EKF parameter selection cases.

Datasets	k_R	k_Q	Quaternion Estimation				Angle Axis Representation	
			Error (RMS)					
			q_w	q_x	q_y	q_z		
BOARD	1	1	0.0140	0.0115	0.0110	0.0147	2.8941	
	1.22	7.3	0.0074	0.0067	0.0047	0.0087	1.4832	
	Improvement		47.14 %	41.74%	57.27%	40.82%	48.75%	
SASARI	1	1	0.0326	0.0287	0.0372	0.0413	7.8811	
	0.1	10.0	0.0138	0.0113	0.0091	0.0129	2.2604	
	Improvement		57.67%	60.63%	75.54%	68.77%	71.32%	

6. Conclusions

In conclusion, we presented a novel Bayesian-based optimization algorithm for tuning the process and measurement noise covariance in EKF-based UAV AHRS estimation applications. The algorithm is built upon EKF-innovation consistency analysis, which is always accessible. Our approach fine-tunes the scaling factor values of the covariance matrices as a whole, allowing scalability to higher-dimensional systems. Furthermore, we

designed an objective function that is both well-behaved and easy to optimize, along with an adaptive search space adjustment strategy. These combined features facilitate rapid and efficient convergence of our algorithm.

The performance of the proposed method was evaluated using openly available datasets. The results showed that the EKF-based AHRS estimation performance significantly improved with the measurement and process noise covariance matrices determined by the proposed algorithm. The method is also general and can be applied to a wide range of AHRS applications.

Author Contributions: Conceptualization, A.W.; methodology, A.W.; software, A.W.; validation, A.W. and S.-K.K.; formal analysis, A.W.; investigation, A.W.; resources, H.-Y.S., B.-S.K. and S.-K.K.; data curation, A.W., B.E. and Y.D.; writing—original draft preparation, A.W. and S.-K.K.; writing—review and editing, A.W., H.-Y.S., Y.D., B.E. and B.-S.K.; visualization, A.W. and Y.D.; supervision, B.-S.K. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by the Development of Social Complex Disaster Response Technology through the Korea Planning & Evaluation Institute of Industrial Technology funded by the Ministry of the Interior and Safety in 2023. (Project Name: Development of risk analysis and evaluation technology for high reliability stampede accidents using CCTV and Drone imaging, project number: 20024403, contribution rate: 50%). This research was also supported by the BK21 FOUR (Fostering Outstanding Universities for Research) funded by the Ministry of Education (MOE, Korea) and the National Research Foundation of Korea (NRF).

Data Availability Statement: Data are contained within the article.

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

Abbreviations

The following abbreviations are used in this manuscript:

KF	Kalman Filter
EKF	extended Kalman filter
UAV	unmanned aerial vehicle
AHRS	attitude and heading reference system
3D	3-dimension
CM	covariance matrix
ANN	artificial neural network
RPE	recursive prediction error
NED	north east down
GP	Gaussian process
EI	expected improvement
PDF	probability density function
CDF	cumulative distribution function
IMU	inertial measurement unit
NIS	normalized innovation squared
RMS	root mean square

References

1. Saha, M.; Goswami, B.; Ghosh, R. Two novel costs for determining the tuning parameters of the Kalman Filter. *arXiv* **2011**, arXiv:1110.3895. [[CrossRef](#)]
2. Pieter, A.; Adam, C.; Michael, M.; Andrew, N.; Sebastian, T. Discriminative Training of Kalman Filters. *Robot. Sci. Syst.* **2005**, 52, 1401–1406.
3. Carew, B.; Belanger, P. Identification of optimum filter steady-state gain for systems with unknown noise covariances. *IEEE Trans. Autom. Control* **1973**, 18, 582–587. [[CrossRef](#)]
4. Mehra, R. Approaches to adaptive filtering. *IEEE Trans. Autom. Control* **1972**, 17, 693–698. [[CrossRef](#)]
5. Duník, J.; Straka, O.; Kost, O.; Havlík, J. Noise covariance matrices in state-space models: A survey and comparison of estimation methods—Part I. *Int. J. Adapt. Control. Signal Process.* **2017**, 31, 1505–1543. [[CrossRef](#)]
6. Gelen, A.G.; Atasoy, A. A New Method for Kalman Filter Tuning. In Proceedings of the 2018 International Conference on Artificial Intelligence and Data Processing (IDAP), Malatya, Turkey, 28–30 September 2018. [[CrossRef](#)]

7. Åkesson, B.M.; Jørgensen, J.B.; Poulsen, N.K.; Jørgensen, S.B. A tool for Kalman filter tuning. *Comput. Aided Chem. Eng.* **2007**, *24*, 859–864. [[CrossRef](#)]
8. Odelson, B.J.; Rajamani, M.R.; Rawlings, J.B. A new autocovariance least-squares method for estimating noise covariances. *Automatica* **2006**, *42*, 303–308. [[CrossRef](#)]
9. Jindrich, D.; Oliver, K.; Ondřej, S. Design of measurement difference autocovariance method for estimation of process and measurement noise covariances. *Automatica* **2018**, *90*, 16–24. [[CrossRef](#)]
10. Mehra, R. On the identification of variances and adaptive Kalman filtering. *IEEE Trans. Autom. Control* **1970**, *15*, 175–184. [[CrossRef](#)]
11. Zhang, A.; Atia, M.M. An efficient tuning framework for Kalman filter parameter optimization using design of experiments and genetic algorithms. *J. Inst. Navig.* **2020**, *67*, 775–793. [[CrossRef](#)]
12. Chhabra, A.; Venepally, J.R.; Kim, D. Measurement Noise Covariance-Adapting Kalman Filters for Varying Sensor Noise Situations. *Sensors* **2021**, *21*, 8304. [[CrossRef](#)]
13. Yuen, K.-V.; Liang, P.-F.; Kuok, S.-C. Online estimation of noise parameters for Kalman filter. *Struct. Eng. Mech.* **2013**, *47*, 361–381. [[CrossRef](#)]
14. Riva, M.H.; Beckmann, D.; Dagen, M.; Ortmaier, T. Online Parameter and Process Covariance Estimation using adaptive EKF and SRCuKF approaches. In Proceedings of the 2015 IEEE Conference on Control Applications (CCA), Sydney, Australia, 21 September 2015. [[CrossRef](#)]
15. Ullah, I.; Fayaz, M.; Kim, D. Improving Accuracy of the Kalman Filter Algorithm in Dynamic Conditions Using ANN-Based Learning Module. *Symmetry* **2019**, *11*, 94. [[CrossRef](#)]
16. Ayyarao, S.T.; Ramanarao P. Tuning of extended Kalman filter for power systems using two lbest particle swarm optimization. *IJCTA* **2017**, *10*, 197–206.
17. Chen, Z.; Heckman, C.; Julier, S.; Ahmed, N. Weak in the NEES?: Auto-tuning Kalman Filters with Bayesian Optimization. *arXiv* **2018**, arXiv:1807.08855. [[CrossRef](#)]
18. Chen, Z.; Ahmed, N.; Julier, S.; Heckman, C. Kalman Filter Tuning with Bayesian Optimization. *arXiv* **2019**, arXiv:1912.08601. [[CrossRef](#)].
19. Park, S.; Gil, M.-S.; Im, H.; Moon, Y.-S. Measurement Noise Recommendation for Efficient Kalman Filtering over a Large Amount of Sensor Data. *Sensors* **2019**, *19*, 1168. [[CrossRef](#)] [[PubMed](#)]
20. Wondosen, A.; Jeong, J.-S.; Kim, S.-K.; Debele, Y.; Kang, B.-S. Improved Attitude and Heading Accuracy with Double Quaternion Parameters Estimation and Magnetic Disturbance Rejection. *Sensors* **2021**, *21*, 5475. [[CrossRef](#)]
21. Sabatini, A.M. Kalman-Filter-Based Orientation Determination Using Inertial/Magnetic Sensors: Observability Analysis and Performance Evaluation. *Sensors* **2011**, *11*, 9182. [[CrossRef](#)]
22. Valenti, R.G.; Dryanovski, I.; Xiao, J. A Linear Kalman Filter for MARG Orientation Estimation Using the Algebraic Quaternion Algorithm. *IEEE Trans. Instrum. Meas.* **2016**, *65*, 467–481. [[CrossRef](#)]
23. Guo, S.; Wu, J.; Wang, Z.; Qian, J. Novel MARG-Sensor Orientation Estimation Algorithm Using Fast Kalman Filter. *J. Sens.* **2017**, *2017*, 8542153. [[CrossRef](#)]
24. René, S.; Christos, G. How To NOT Make the Extended Kalman Filter Fail. *Ind. Eng. Chem. Res.* **2013**, *52*, 3354–3362. [[CrossRef](#)]
25. Laidig, D.; Caruso, M.; Cereatti, A.; Seel, T. BROAD—A Benchmark for Robust Inertial Orientation Estimation. *Data* **2021**, *6*, 72. [[CrossRef](#)]
26. Caruso, M.; Sabatini, A.M.; Laidig, D.; Seel, T.; Knaflitz, M.; Della Croce, U.; Cereatti, A. Analysis of the Accuracy of Ten Algorithms for Orientation Estimation Using Inertial and Magnetic Sensing under Optimal Conditions: One Size Does Not Fit All. *Sensors* **2021**, *21*, 2543. [[CrossRef](#)]

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.