

# Kernel Entropy Based Extended Kalman Filter for GPS Navigation Processing

Dah-Jing Jwo\* and Jui-Tao Lee

Department of Communications, Navigation and Control Engineering, National Taiwan Ocean University, Keelung, 202301, Taiwan

\*Corresponding Author: Dah-Jing Jwo. Email: djjwo@mail.ntou.edu.tw

Received: 12 January 2021; Accepted: 14 February 2021

**Abstract:** This paper investigates the kernel entropy based extended Kalman filter (EKF) as the navigation processor for the Global Navigation Satellite Systems (GNSS), such as the Global Positioning System (GPS). The algorithm is effective for dealing with non-Gaussian errors or heavy-tailed (or impulsive) interference errors, such as the multipath. The kernel minimum error entropy (MEE) and maximum correntropy criterion (MCC) based filtering for satellite navigation system is involved for dealing with non-Gaussian errors or heavy-tailed interference errors or outliers of the GPS. The standard EKF method is derived based on minimization of mean square error (MSE) and is optimal only under Gaussian assumption in case the system models are precisely established. The GPS navigation algorithm based on kernel entropy related principles, including the MEE criterion and the MCC will be performed, which is utilized not only for the time-varying adaptation but the outlier type of interference errors. The kernel entropy based design is a new approach using information from higher-order signal statistics. In information theoretic learning (ITL), the entropy principle based measure uses information from higher-order signal statistics and captures more statistical information as compared to MSE. To improve the performance under non-Gaussian environments, the proposed filter which adopts the MEE/MCC as the optimization criterion instead of using the minimum mean square error (MMSE) is utilized for mitigation of the heavy-tailed type of multipath errors. Performance assessment will be carried out to show the effectiveness of the proposed approach for positioning improvement in GPS navigation processing.

**Keywords:** GPS; satellite navigation; extended Kalman filter; entropy; correntropy; multipath; non-Gaussian

## 1 Introduction

Non-Gaussian noise is often encountered in many practical environments where the estimation performance deteriorates dramatically. Multipath [1] is known to be one of the dominant error sources in high accuracy global navigation satellite systems (GNSS) positioning systems, such as the Global Positioning System (GPS) [1,2]. Multipath effects occur when GPS signals arrive at



This work is licensed under a Creative Commons Attribution 4.0 International License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

a receiver site via multiple paths due to reflections from nearby objects, such as the ground and water surfaces, buildings, vehicles, hills, trees, etc. Many estimation algorithms have been studied to eliminate the positioning error caused by multipath. Since multipath errors are among uncorrelated errors that are not cancelled out during observation differencing, the performance of high precision GPS receivers are mostly limited by the multipath induced errors. One of the most important issues in GPS system performance improvement is the interference suppression techniques.

Due to its simple structure, stable performance and low computational complexity, the conventional adaptive filtering algorithm where the least mean square error (MSE) is involved has been widely used in a variety of applications in the fields of adaptive signal processing and machine learning. However, the MSE criterion is limited to the assumption of linearity and Gaussianity while most of the noise in real world is non-Gaussian. The performance deteriorates significantly in the non-Gaussian noise environment. The well-known Kalman filter (KF) [2,3] provides optimal (minimum MSE) estimate of the system state vector and has been recognized as one of the most powerful state estimation techniques. The traditional Kalman-type filter provides the best filter estimate when the noise is Gaussian, but most noise in real life is unknown, uncertain and non-Gaussian. Since the Kalman filter uses only second order signal information, it is not optimal when the system is disturbed by heavy-tailed (or impulsive) non-Gaussian noises. The extended Kalman filter (EKF) is a nonlinear version of the KF and has been widely employed as the GPS navigation processor. The fact that EKF highly depends on a predefined dynamics model forms a major drawback. To solve the performance degradation problem with non-Gaussian errors or heavy-tailed non-Gaussian noises, some robust Kalman filters have been developed by using non-minimum MSE criterion as the optimality criterion.

As a unified probabilistic measure of uncertainty quantification, entropy [4–6] has been widely used in information theory. The novel schemes using entropy principle based nonlinear filters are suitable as alternatives for GPS navigation processing. The robustness of algorithms has become a crucial issue when dealing with the practical GPS navigation application in non-Gaussian noise environments. In the cases where the additive noises in signal processing is supposed as Gaussian process, the MSE can be adopted for construction of the kernel adaptive filtering algorithms. The algorithm can suppress the effects of impulsive noise through kernel function in entropy, thus guarantees a good performance for non-Gaussian application. By introducing entropy/correntropy, kernel recursive algorithm based on minimum error entropy/maximum correntropy criterion can be employed to overcome the deteriorating performance where the LMS algorithm is involved for non-Gaussian signal. The minimum error entropy (MEE) criterion [7–9] and maximum correntropy criterion (MCC) [10–16] are information theoretic learning (ITL) approaches, which have been successfully applied in robust regression, classification, system identification and adaptive filtering. The algorithm updates equation recursively by minimizing the error entropy/maximizing the correntropy between output of the system and the desired signal. As compared with LMS, both the MEE and MCC algorithm possess better stability in non-Gaussian environments.

As a novel performance index, some of filters applied to non-Gaussian systems have been proposed. The MEE criterion is an important learning criterion in ITL, which has been successfully applied in robust regression, classification, system identification and adaptive filtering and has been widely adopted in non-Gaussian signal processing. The MEE scheme is designed by introducing an additional term, which and is tuned according to the higher order moment of the estimation error. The algorithm has a high accuracy in estimation because entropy can characterize all the randomness of the residual. The MEE adopted to minimize the error to obtain the maximum amount of information through measuring error information and ensures the local stability of the

error dynamic. The MEE is a method of information theory learning which has been successfully applied to Kalman filters to improve robustness against pulsed noise. Information theory learning has been successfully applied to robust regression, classification, system recognition and adaptive filtering. The MCC is another important learning criterion which has been successfully used to handle the heavy-tailed non-Gaussian noise. Maximizing the mutual information between a state and the estimate is equivalent to minimizing the entropy of the estimation error. Based on information theory, another entropy criterion is proposed. Many experiments have shown that although MEE achieves excellent performance, the computational complexity is slightly higher than MCC.

The kernel entropy based EKF is adopted for the GPS navigation processing. Performance evaluation will be conducted to investigate the performance based on the two alternative entropy-related criteria: MEE and MCC. Results will be given to demonstrate the superiority of the designs with appropriate kernel bandwidth. Adaptive algorithms under MEE and MCC show enhanced robustness in the presence of non-Gaussian disturbances or heavy-tailed interference errors, such as the multipath interference. The remainder of this paper is organized as follows. In Section 2, preliminary background on the EKF and AEKF is reviewed. Section 3 addresses the basic principles on entropy theory, includes the MEE and MCC. The MEE- and MCC-based EKF's are introduced in Section 4, where the MCC-based AEKF is also presented. In Section 5, numerical experiments are carried out to evaluate the performance using the proposed MCC-AEKF as compared to the other approaches. Conclusion is given in Section 6.

## 2 The Extended Kalman Filter and Covariance Scaling

Given a non-linear single model equation in discrete time

$$\mathbf{x}_{k+1} = \mathbf{f}(\mathbf{x}_k) + \mathbf{w}_k \quad (1)$$

$$\mathbf{z}_k = \mathbf{h}(\mathbf{x}_k) + \mathbf{v}_k \quad (2)$$

where the state vector  $\mathbf{x}_k \in \mathfrak{R}^n$ , process noise vector  $\mathbf{w}_k \in \mathfrak{R}^m$ , measurement vector  $\mathbf{z}_k \in \mathfrak{R}^m$ , measurement noise vector  $\mathbf{v}_k \in \mathfrak{R}^m$ ,  $\mathbf{Q}_k$  is the process noise covariance matrix and  $\mathbf{R}_k$  is the measurement noise covariance matrix.

The vectors  $\mathbf{w}_k$  and  $\mathbf{v}_k$  in Eqs. (1) and (2) are zero mean Gaussian white sequences has zero cross-correlation with each other:

$$E[\mathbf{w}_k \mathbf{w}_i^T] = \mathbf{Q}_k \delta_{ik}; \quad E[\mathbf{v}_k \mathbf{v}_i^T] = \mathbf{R}_k \delta_{ik}; \quad E[\mathbf{w}_k \mathbf{v}_i^T] = \mathbf{0} \quad \text{for all } k \text{ and } i \quad (3)$$

where  $E[\cdot]$  represents expectation, and superscript "T" denotes matrix transpose. The symbol  $\delta_{ik}$  stands for the Kronecker delta function:

$$\delta_{ik} = \begin{cases} 1, & i = k \\ 0, & i \neq k \end{cases}$$

The discrete-time adaptive extended Kalman filter algorithm is summarized as follow:

—Initialize state vector and state covariance matrix:  $\hat{\mathbf{x}}_0$  and  $\mathbf{P}_0$

Stage 1: correction steps/measurement update equations

(1) Compute Kalman gain matrix:

$$\mathbf{K}_k = \mathbf{P}_{k|k-1} \mathbf{H}_k^T \left[ \mathbf{H}_k \mathbf{P}_{k|k-1} \mathbf{H}_k^T + \mathbf{R}_k \right]^{-1} \quad (4)$$

(2) Update state vector:

$$\hat{\mathbf{x}}_k = \hat{\mathbf{x}}_{k|k-1} + \mathbf{K}_k [\mathbf{z}_k - \mathbf{h}(\hat{\mathbf{x}}_{k|k-1})] \quad (5)$$

(3) Update error covariance

The error covariance relationships for a discrete filter with the same structure as the Kalman filter, but with an arbitrary gain matrix are written as

$$\mathbf{P}_k = [\mathbf{I} - \mathbf{K}_k \mathbf{H}_k] \mathbf{P}_{k|k-1} [\mathbf{I} - \mathbf{K}_k \mathbf{H}_k]^T + \mathbf{K}_k \mathbf{R}_k \mathbf{K}_k^T \quad (6)$$

Stage 2: Prediction steps/time update equations

(4) Predict state vector

$$\hat{\mathbf{x}}_{k+1|k} = \mathbf{f}(\hat{\mathbf{x}}_k) \quad (7)$$

(5) Predict state covariance matrix

$$\mathbf{P}_{k+1|k} = \Phi_k \mathbf{P}_k \Phi_k^T + \mathbf{Q}_k \quad (8)$$

where the linear approximation equations for system and measurement matrices are obtained through the relations

$$\Phi_k \approx \left. \frac{\partial \mathbf{f}}{\partial \mathbf{x}} \right|_{\mathbf{x}=\hat{\mathbf{x}}_{k|k-1}} ; \quad \mathbf{H}_k \approx \left. \frac{\partial \mathbf{h}}{\partial \mathbf{x}} \right|_{\mathbf{x}=\hat{\mathbf{x}}_{k|k-1}}$$

The discrete-time adaptive extended Kalman filter (AEKF) algorithm is summarized as follow.

(1) Compute measurement residual:

From the incoming measurement  $\mathbf{z}_k$  and the optimal prediction  $\hat{\mathbf{x}}_{k|k-1}$  obtained in the previous step, the innovations sequence is defined as  $\mathbf{v}_k = \mathbf{z}_k - \mathbf{h}(\hat{\mathbf{x}}_{k|k-1})$ . The innovation reflects the discrepancy between the predicted measurement  $\mathbf{h}(\hat{\mathbf{x}}_{k|k-1})$  and the actual measurement  $\mathbf{z}_k$ . It represents the additional information available to the filter as a consequence of the new observation.

(2) The covariance of measurement residual matrix

By taking variances on both sides, we have the theoretical covariance, the covariance matrix of the innovation sequence is given by

$$\mathbf{C}_{v_k} = E[\mathbf{v}_k \mathbf{v}_k^T] = \mathbf{H}_k \mathbf{P}_{k|k-1} \mathbf{H}_k^T + \mathbf{R}_k \quad (9)$$

(3) Estimate the innovation covariance

Defining  $\hat{\mathbf{C}}_{v_k}$  as the statistical sample variance estimate of  $\mathbf{C}_{v_k}$ , matrix  $\hat{\mathbf{C}}_{v_k}$  can be computed through averaging inside a moving estimation window of size  $N$

$$\hat{\mathbf{C}}_{v_k} = \frac{1}{N} \sum_{j=j_0}^k \mathbf{v}_j \mathbf{v}_j^T \quad (10)$$

where  $N$  is the number of samples (usually called the window size);  $j_0 = k - N + 1$  is the first sample inside the estimation window. The window size  $N$  is chosen empirically (a good size for the moving window may vary from 10 to 30, and  $N = 15$  has been used in this paper) to give some statistical smoothing.

(4) Compute the forgetting factor

$$\lambda_k = \max \left\{ 1, \frac{\text{trace} [\hat{\mathbf{C}}_{v_k}]}{\text{trace} [\mathbf{C}_{v_k}]} \right\} \quad (11)$$

One of the other approaches for adaptive processing is on the incorporation of fading factors. The idea of fading memory is to apply a factor matrix to the predicted covariance matrix to deliberately increase the variance of the predicted state vector:

$$\mathbf{P}_{k+1|k} = \lambda_k \Phi_k \mathbf{P}_k \Phi_k^T + \mathbf{Q}_k \quad (12)$$

where  $\lambda_k = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_m)$ . The main difference between different fading memory algorithms is on the calculation of scale factor matrix  $\lambda_k$ . One approach is to assign the scale factors as constants. When  $\lambda_i \leq 1$  ( $i = 1, 2, \dots, m$ ), the filtering is in a steady state processing while  $\lambda_i > 1$ , the filtering may tend to be unstable. For the case  $\lambda_i = 1$ , it deteriorates to the standard Kalman filter.

### 3 Entropy

From the definition of the information theoretic and kernel methods, entropy is a measure of the uncertainty associated with random variables. ITL is a framework to non-parametrically adapt systems based on entropy and divergence. Correntropy denotes a generalized similarity measure between two random variables.

#### 3.1 Minimum Error Entropy

Originally presented by Shannon in 1948, many definitions of entropy have been introduced for various purposes, such as Shannon entropy and Renyi's entropy. Renyi's entropy, named after Alfred Renyi, is usually used for quantifying the diversity, uncertainty or randomness of a random variable. The quadratic Renyi's entropy, which has the form

$$H_2(\mathbf{X}) = -\log \int f_{\mathbf{X}}^2(\mathbf{x}) d\mathbf{x} \quad (13)$$

There are numerous methods to estimate the probability density. The kernel density estimation (KDE) has wide applicability and is closely related to the Renyi's entropy. Kernel density estimation, also called Parzen window method, is a nonparametric method to estimate the probability density function of a random process. One can estimate the quadratic information potential of error entropy using a sample mean estimator as follow as

$$\hat{V}_2(e) = \frac{1}{N} \sum_{i=1}^N G_{\sigma}(e_i) = \frac{1}{N^2} \sum_{i=1}^N \sum_{j=1}^N \kappa_{\sigma}(e_i - e_j) \quad (14)$$

where

$$k_{\sigma}(x, y) = G_{\sigma}(e_i) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{e_i^2}{2\sigma^2}\right) \quad (15)$$

denotes the Gaussian kernel, which is the most popular kernel function and is also adopted in this paper. Due to the negative logarithmic function monotonically decreasing function, it can be seen that minimizing the error entropy  $H_2(e)$  means maximizing the information potential  $\hat{V}_2(e)$ .

### 3.2 Maximum Correntropy Criterion

In recent years, the maximum correntropy criterion has been successfully applied in many areas of signal processing, pattern recognition and machine learning with the existence of non-Gaussian noise, especially the large outliers.

Correntropy between two scalar variables measures the second-order information as well as higher-order statistical information in the joint probability density function. The correntropy of two random scalar variables  $X$  and  $Y$  is defined as

$$V_{\sigma}(X, Y) = E[\kappa_{\sigma}(X, Y)] = \int \kappa_{\sigma}(x, y) dF_{X, Y}(x, y) = \int \kappa_{\sigma}(x, y) f_{XY}(x, y) dx dy \quad (16)$$

where  $E[\cdot]$  represents the expectation operator,  $F_{X, Y}(x, y)$  is the joint distribution,  $f_{XY}(x, y)$  is the joint probability density function, and  $\kappa_{\sigma}(x, y)$  is a positive definite kernel function that satisfies the Mercer theory. In most practical situations, only limited amount of data is available and the joint distribution is usually unknown. In these cases, one can estimate the correntropy using a sample mean estimator as follow as

$$\hat{V}_{\sigma}(X, Y) = \frac{1}{N} \sum_{i=1}^N G_{\sigma}(e_i) = \frac{1}{N} \sum_{i=1}^N \kappa_{\sigma}(x_i - y_i) \quad (17)$$

where  $e_i = x_i - y_i$ ,  $\sigma > 0$  denotes the kernel bandwidth. In most practical applications, the number of data we could access is scant and the joint distribution  $F_{X, Y}(x, y)$  is usually unavailable. For solving this problem, one can use the sample mean square to estimate the correntropy. Taking the Taylor series expansion of the Gaussian kernel, we have

$$V_{\sigma}(X, Y) = \sum_{n=0}^{\infty} \frac{(-1)^n}{2^n \sigma^{2n} n!} E[(X - Y)^{2n}] \quad (18)$$

It can be seen that correntropy represents a weighted sum of all even order moments of the two random variables  $X$  and  $Y$ . The kernel bandwidth appears as a parameter weighting the second order and higher order moments. With a very large (compared to the dynamic range of the data), the correntropy will be dominated by the second order moment, and then the maximum correntropy criterion will be approximately equal to the minimum mean square error criterion.

#### 4 The Kernel Entropy Based Extended Kalman Filter

Consider an augmented model given by state prediction error with the measurement equation as

$$\begin{bmatrix} \hat{\mathbf{x}}_{k|k-1} \\ \mathbf{z}_k \end{bmatrix} = \begin{bmatrix} \mathbf{x}_k \\ \mathbf{h}(\mathbf{x}_k) \end{bmatrix} + \begin{bmatrix} \delta \mathbf{x}_k \\ \mathbf{v}_k \end{bmatrix} \quad (19)$$

where  $\delta \mathbf{x}_k$  is the state prediction error as  $\delta \mathbf{x}_k = \hat{\mathbf{x}}_{k|k-1} - \mathbf{x}_k$  and we have

$$\boldsymbol{\varphi}_k = \begin{bmatrix} \hat{\mathbf{x}}_{k|k-1} - \mathbf{x}_k \\ \mathbf{v}_k \end{bmatrix} \quad (20)$$

The covariance matrix for  $\boldsymbol{\varphi}_k$  can be represented by the joint matrix composed of  $\mathbf{P}_{k|k-1}$  and  $\mathbf{R}_k$  matrices, and  $\mathbf{B}_k$  can be obtained by Cholesky decomposition of the joint matrix.

$$\begin{aligned} E[\boldsymbol{\varphi}_k \boldsymbol{\varphi}_k^T] &= \mathbf{B}_k \mathbf{B}_k^T = \begin{bmatrix} \mathbf{P}_{k|k-1} & 0 \\ 0 & \mathbf{R}_k \end{bmatrix} = \begin{bmatrix} \mathbf{B}_{p,k|k-1} \mathbf{B}_{p,k|k-1}^T & 0 \\ 0 & \mathbf{B}_{r,k} \mathbf{B}_{r,k}^T \end{bmatrix} \\ &= \begin{bmatrix} \mathbf{B}_{p,k|k-1} & 0 \\ 0 & \mathbf{B}_{r,k} \end{bmatrix} \begin{bmatrix} \mathbf{B}_{p,k|k-1} & 0 \\ 0 & \mathbf{B}_{r,k} \end{bmatrix}^T \end{aligned} \quad (21)$$

Multiplying both sides on Eq. (19) by  $\mathbf{B}_k^{-1}$ , we have

$$\boldsymbol{\varepsilon}_k = \mathbf{D}_k - \mathbf{g}_k(\mathbf{x}_k) \quad (22)$$

where

$$\mathbf{D}_k = \mathbf{B}_k^{-1} \begin{bmatrix} \hat{\mathbf{x}}_{k|k-1} \\ \mathbf{z}_k \end{bmatrix}, \quad \mathbf{g}_k(\mathbf{x}_k) = \mathbf{B}_k^{-1} \begin{bmatrix} \mathbf{x}_k \\ \mathbf{h}(\mathbf{x}_k) \end{bmatrix}, \quad \boldsymbol{\varepsilon}_k = \mathbf{B}_k^{-1} \begin{bmatrix} \delta \mathbf{x}_k \\ \mathbf{v}_k \end{bmatrix} = \mathbf{B}_k^{-1} \boldsymbol{\varphi}_k \quad (23)$$

$$E[\boldsymbol{\varepsilon}_k \boldsymbol{\varepsilon}_k^T] = E[\mathbf{B}_k^{-1} \boldsymbol{\varphi}_k \boldsymbol{\varphi}_k^T (\mathbf{B}_k^{-1})^T] = E[\mathbf{I}] = \mathbf{I} \quad (24)$$

where  $\mathbf{I}$  is denoted the  $n \times n$  identity matrix, and the  $i$ -th row element in  $\boldsymbol{\varepsilon}_k$  is

$$\boldsymbol{\varepsilon}_{i,k} = \mathbf{d}_{i,k} - \mathbf{g}_{i,k}(\mathbf{x}_k) \quad (25)$$

##### 4.1 Minimum Error Entropy-Based Extended Kalman Filter

The idea for the MEE-based EKF is to optimize the following cost function  $J_{MEE}$

$$J_{MEE}(\mathbf{x}_k) = \frac{1}{L^2} \sum_{i=1}^L \sum_{j=1}^L G_\sigma(\boldsymbol{\varepsilon}_{j,k} - \boldsymbol{\varepsilon}_{i,k}) \quad (26)$$

Taking its derivative with respect to  $\mathbf{x}_k$  and setting it to zero

$$\frac{\partial J_{MEE}(\mathbf{x}_k)}{\partial \mathbf{x}_k} = 0 \quad (27)$$

we have

$$\hat{\mathbf{x}}_{k|k} = \arg \max_{\mathbf{x}_k} J_{MEE}(\mathbf{x}_k) = \arg \max_{\mathbf{x}_k} \sum_{i=1}^L \sum_{j=1}^L G_{\sigma}(\boldsymbol{\varepsilon}_{j,k} - \boldsymbol{\varepsilon}_{i,k}) \quad (28)$$

The solution cannot be obtained in closed form even for a simple linear regression problem, so one has to solve it using an iterative update algorithm such as the gradient based methods. The gradient based methods are simple and widely used. However, they depend on a free parameter step-size and usually converge to an optimal solution slowly. The fixed-point iterative algorithm is an alternative efficient way to solve the solution, which involves no step-size and may converge to the solution very fast. The computation procedures for the minimum error entropy based extended Kalman filter (MEE-EKF) are summarized as follows:

- (1) Choose a kernel bandwidth  $\sigma$  and the error tolerance  $\varepsilon$  with small positive value. Set the initial estimate and covariance  $\hat{\mathbf{x}}_0, \mathbf{P}_0$ ;
- (2) Perform Cholesky decomposition to obtain  $\mathbf{B}_{p,k|k-1}$ ;
- (3) Let  $\hat{\mathbf{x}}_{k,0} = \hat{\mathbf{x}}_{k|k-1}$ , where  $\hat{\mathbf{x}}_{k,t}$  represents the state estimate at the fixed-point iteration  $t$ ;
- (4) Iteration loop: Calculation of  $\hat{\mathbf{x}}_{k,t}$  using the following steps:

- 1)  $\tilde{\mathbf{e}}_{i,k} = \mathbf{d}_{i,k} - \mathbf{g}_{i,k}(\hat{\mathbf{x}}_{k,t})$

- 2)  $[\Phi_k]_{ij} = G_{\sigma}(\boldsymbol{\varepsilon}_{j,k} - \boldsymbol{\varepsilon}_{i,k})$

- 3)  $\Psi_k = \text{diag} \left\{ \sum_{i=1}^L G_{\sigma}(\boldsymbol{\varepsilon}_{1,k} - \boldsymbol{\varepsilon}_{i,k}), \dots, \sum_{i=1}^L G_{\sigma}(\boldsymbol{\varepsilon}_{L,k} - \boldsymbol{\varepsilon}_{i,k}) \right\}$

- 4)  $\tilde{\mathbf{C}}_k = \Phi_k - \Psi_k = \begin{bmatrix} \tilde{\mathbf{C}}_{x,k} & \tilde{\mathbf{C}}_{zx,k} \\ \tilde{\mathbf{C}}_{xz,k} & \tilde{\mathbf{C}}_{z,k} \end{bmatrix}$

- 5)  $\tilde{\mathbf{P}}_{xx,k|k-1} = (\mathbf{B}_{p,k|k-1}^{-1})^T \tilde{\mathbf{C}}_{x,k} \mathbf{B}_{p,k|k-1}^{-1}$

- 6)  $\tilde{\mathbf{P}}_{xz,k|k-1} = (\mathbf{B}_{p,k|k-1}^{-1})^T \tilde{\mathbf{C}}_{xz,k} \mathbf{B}_{r,k}^{-1}$

- 7)  $\tilde{\mathbf{P}}_{zx,k|k-1} = (\mathbf{B}_{r,k}^{-1})^T \tilde{\mathbf{C}}_{zx,k} \mathbf{B}_{p,k|k-1}^{-1}$

- 8)  $\tilde{\mathbf{R}}_k = (\mathbf{B}_{r,k}^{-1})^T \tilde{\mathbf{C}}_{z,k} \mathbf{B}_{r,k}^{-1}$

- 9)  $\tilde{\mathbf{K}}_k = (\tilde{\mathbf{P}}_{xx,k|k-1} + \mathbf{H}_k^T \tilde{\mathbf{P}}_{xz,k|k-1} + (\tilde{\mathbf{P}}_{zx,k|k-1} + \mathbf{H}_k^T \tilde{\mathbf{R}}_k) \mathbf{H}_k)^{-1} (\tilde{\mathbf{P}}_{zx,k|k-1} + \mathbf{H}_k^T \tilde{\mathbf{R}}_k)$

- 10)  $\hat{\mathbf{x}}_{k,t} = \hat{\mathbf{x}}_{k|k-1} + \tilde{\mathbf{K}}_k [\mathbf{z}_k - \mathbf{h}_k(\hat{\mathbf{x}}_{k|k-1})]$

- (5) Compare the estimation for the current steps with the previous steps for convergence check

$$\frac{\|\hat{\mathbf{x}}_{k,t} - \hat{\mathbf{x}}_{k,t-1}\|}{\|\hat{\mathbf{x}}_{k,t-1}\|} \leq \varepsilon$$

If the above condition holds, then set  $\hat{\mathbf{x}}_k = \hat{\mathbf{x}}_{k,t}$ , go to Step 6, otherwise, go back to Step 4

- (6) Calculation of update covariance matrix:

$$\mathbf{P}_k = (\mathbf{I} - \tilde{\mathbf{K}}_k \mathbf{H}_k) \mathbf{P}_{k|k-1} (\mathbf{I} - \tilde{\mathbf{K}}_k \mathbf{H}_k)^T + \tilde{\mathbf{K}}_k \mathbf{R}_k \tilde{\mathbf{K}}_k^T$$



(7) Predict  $\hat{\mathbf{x}}_{k+1|k}$  and  $\mathbf{P}_{k+1|k}$ . Set  $k+1 \rightarrow k$ , and go back to Step 2

$$\hat{\mathbf{x}}_{k+1|k} = \mathbf{f}(\hat{\mathbf{x}}_k)$$

$$\mathbf{P}_{k+1|k} = \Phi_k \mathbf{P}_k \Phi_k^T + \mathbf{Q}_k$$

#### 4.2 Maximum Correntropy Criterion-Based Extended Kalman Filter

The idea to optimize the following cost function  $J_{MCC}$

$$J_{MCC}(\mathbf{x}_k) = \sum_{i=1}^L G_\sigma(\mathbf{d}_{i,k} - \mathbf{g}_{i,k}(\mathbf{x}_k)) \tag{29}$$

can be obtained by solving

$$\frac{\partial J_{MCC}(\mathbf{x}_k)}{\partial \mathbf{x}_k} = 0 \tag{30}$$

and we have

$$\mathbf{x}_{k|k} = \arg \max_{\mathbf{x}_k} J_{MCC}(\mathbf{x}_k) = \arg \max_{\mathbf{x}_k} \sum_{i=1}^L G_\sigma(\mathbf{d}_{i,k} - \mathbf{g}_{i,k}(\mathbf{x}_k)) = \arg \max_{\mathbf{x}_k} \sum_{i=1}^L G_\sigma(\mathbf{e}_{i,k}) \tag{31}$$

The covariance matrices  $\mathbf{P}_{k|k-1}$  and  $\mathbf{R}_k$  can be further expressed as

$$\begin{aligned} \begin{bmatrix} \mathbf{P}_{k|k-1} & 0 \\ 0 & \mathbf{R}_k \end{bmatrix} &= \begin{bmatrix} \mathbf{B}_{p,k|k-1} & 0 \\ 0 & \mathbf{B}_{r,k} \end{bmatrix} \begin{bmatrix} \mathbf{C}_{x,k} & 0 \\ 0 & \mathbf{C}_{z,k} \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{B}_{p,k|k-1} & 0 \\ 0 & \mathbf{B}_{r,k} \end{bmatrix}^T \\ &= \begin{bmatrix} \mathbf{B}_{p,k|k-1} \mathbf{C}_{x,k}^{-1} \mathbf{B}_{p,k|k-1}^T & 0 \\ 0 & \mathbf{B}_{r,k} \mathbf{C}_{z,k}^{-1} \mathbf{B}_{r,k}^T \end{bmatrix} \end{aligned} \tag{32}$$

where we have

$$\mathbf{P}_{k|k-1} = \mathbf{B}_{p,k|k-1} \mathbf{C}_{x,k}^{-1} \mathbf{B}_{p,k|k-1}^T \tag{33}$$

$$\mathbf{R}_k = \mathbf{B}_{r,k} \mathbf{C}_{z,k}^{-1} \mathbf{B}_{r,k}^T \tag{34}$$

The computation procedures for the maximum correntropy criterion based extended Kalman filter (MCC-EKF) are summarized as follows:

- (1) Choose a kernel bandwidth  $\sigma$  and the error tolerance  $\varepsilon$  with small positive value. Set the initial estimate and covariance  $\hat{\mathbf{x}}_0, \mathbf{P}_0$ ;
- (2) Perform Cholesky decomposition to obtain  $\mathbf{B}_{p,k|k-1}$ ;
- (3) Let  $\hat{\mathbf{x}}_{k,0} = \hat{\mathbf{x}}_{k|k-1}$ , where  $\hat{\mathbf{x}}_{k,t}$  represents the state estimate at the fixed-point iteration  $t$ ;
- (4) Iteration loop: Calculation of  $\hat{\mathbf{x}}_{k,t}$  using the following steps:
  - 1)  $\tilde{e}_{i,k} = d_{i,k} - g_{i,k}(\hat{\mathbf{x}}_{k,t})$
  - 2)  $\tilde{\mathbf{C}}_{x,k} = \text{diag}(G_\sigma(\tilde{e}_{1,k}), \dots, G_\sigma(\tilde{e}_{n,k}))$

- 3)  $\tilde{\mathbf{C}}_{z,k} = \text{diag}(G_\sigma(\tilde{\epsilon}_{n+1,k}), \dots, G_\sigma(\tilde{\epsilon}_{n+m,k}))$
  - 4)  $\tilde{\mathbf{P}}_{k|k-1} = \mathbf{B}_{p,k|k-1} \tilde{\mathbf{C}}_{x,k}^{-1} \mathbf{B}_{p,k|k-1}^T$
  - 5)  $\tilde{\mathbf{R}}_k = \mathbf{B}_{r,k} \tilde{\mathbf{C}}_{z,k}^{-1} \mathbf{B}_{r,k}^T$
  - 6)  $\tilde{\mathbf{K}}_k = \tilde{\mathbf{P}}_{k|k-1} \mathbf{H}_k^T [\mathbf{H}_k \tilde{\mathbf{P}}_{k|k-1} \mathbf{H}_k^T + \tilde{\mathbf{R}}_k]^{-1}$
  - 7)  $\hat{\mathbf{x}}_{k,t} = \hat{\mathbf{x}}_{k|k-1} + \tilde{\mathbf{K}}_k [\mathbf{z}_k - \mathbf{h}_k(\hat{\mathbf{x}}_{k|k-1})]$
- (5) Compare the estimation for the current steps with the previous steps for convergence check

$$\frac{\|\hat{\mathbf{x}}_{k,t} - \hat{\mathbf{x}}_{k,t-1}\|}{\|\hat{\mathbf{x}}_{k,t-1}\|} \leq \varepsilon$$

If the above condition holds, then set  $\hat{\mathbf{x}}_k = \hat{\mathbf{x}}_{k,t}$ , go to Step 6, otherwise, go back to Step 4

- (6) Calculation of update covariance matrix:

$$\mathbf{P}_k = (\mathbf{I} - \tilde{\mathbf{K}}_k \mathbf{H}_k) \mathbf{P}_{k|k-1} (\mathbf{I} - \tilde{\mathbf{K}}_k \mathbf{H}_k)^T + \tilde{\mathbf{K}}_k \mathbf{R}_k \tilde{\mathbf{K}}_k^T$$

- (7) Predict  $\hat{\mathbf{x}}_{k+1|k}$  and  $\mathbf{P}_{k+1|k}$ . Set  $k+1 \rightarrow k$ , and go back to Step 2

$$\hat{\mathbf{x}}_{k+1|k} = \mathbf{f}(\hat{\mathbf{x}}_k)$$

$$\mathbf{P}_{k+1|k} = \Phi_k \mathbf{P}_k \Phi_k^T + \mathbf{Q}_k$$

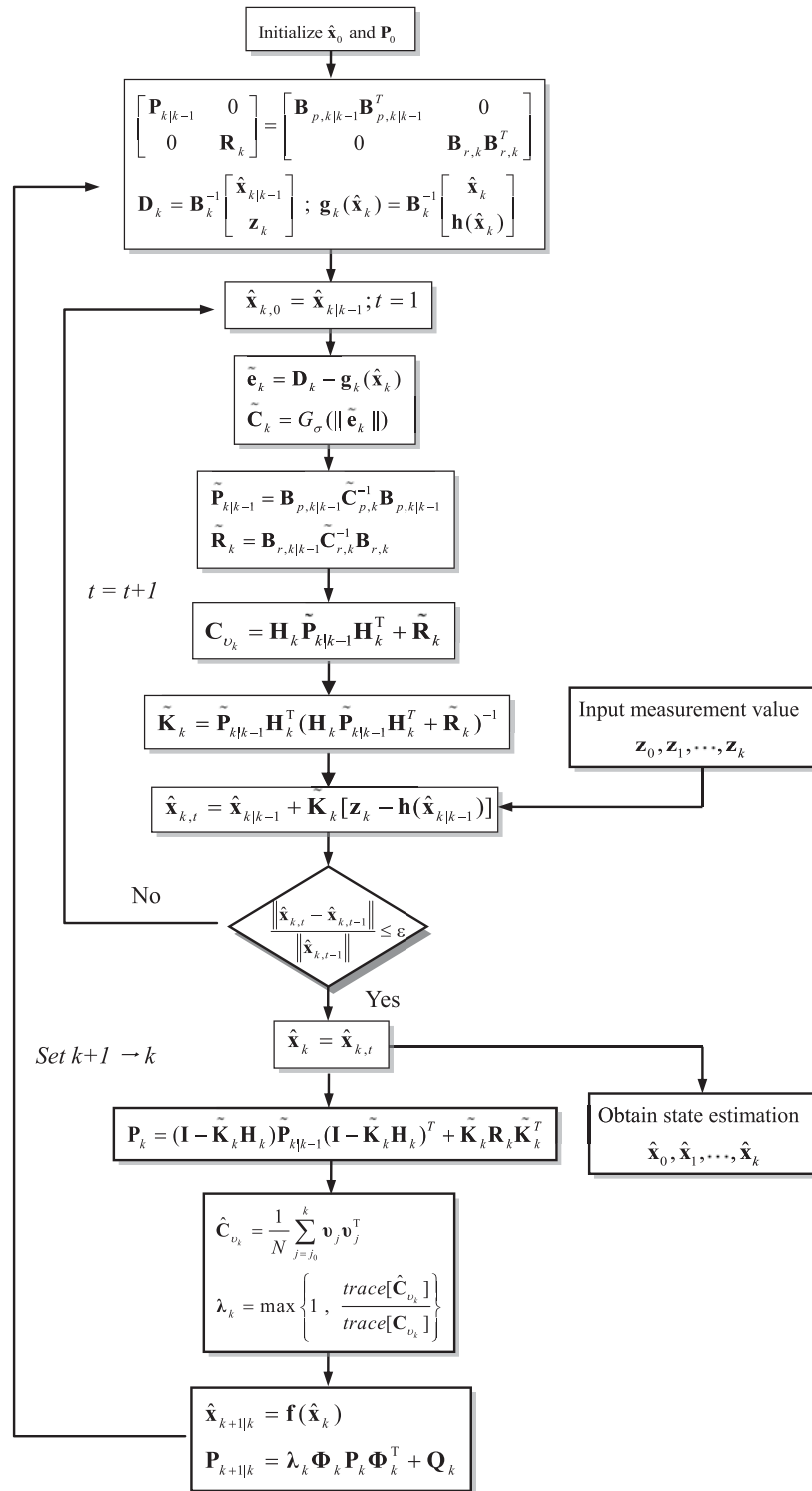
### 4.3 Maximum Correntropy Criterion-Based Adaptive Extended Kalman Filter

Utilization of the MCC-AEKF is a treatment for further performance enhancement. Fig. 1 provides the flow chart for one cycle of the maximum correntropy criterion-based adaptive extended Kalman filter (MCC-AEKF), which involves the computation procedure in both MCC and AEKF.

To fulfil the requirement, an adaptive Kalman filter can be utilized as the noise-adaptive filter to estimate the noise covariance matrices and overcome the deficiency of Kalman filter. The benefit of the adaptive algorithm is that it keeps the covariance consistent with the real performance. The innovation sequences have been utilized by the correlation and covariance-matching techniques to estimate the noise covariances. The basic idea behind the covariance-matching approach is to make the actual value of the covariance of the residual consistent with its theoretical value.

## 5 Results and Discussion

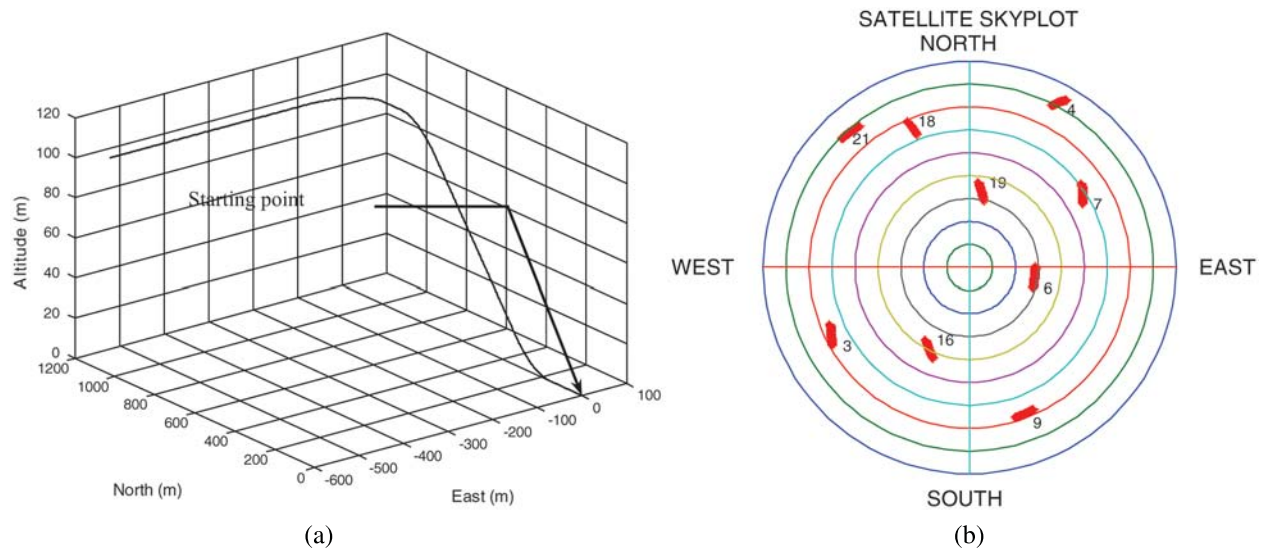
To validate the effectiveness of the proposed approaches, simulation experiments have been carried out to evaluate the performance of the proposed kernel entropy based approach in comparison with the other conventional methods for GPS navigation processing. The kernel entropy principle assisted EKF for GPS navigation processing is presented. Two scenarios dealing with two types of interferences are carried out, including pseudorange observable errors involving (1) time-varying variance in the measurement noise, and (2) outlier type of multipath interferences, during the vehicle moving.



**Figure 1:** Flow chart for one cycle of the maximum correntropy criterion-based adaptive extended Kalman filter (MCC-AEKF)

The computer codes were developed by the authors using the Matlab<sup>®</sup> software. The commercial software Satellite Navigation (SATNAV) Toolbox by GPSofSoft LLC [17] was employed for generation of the GPS satellite orbits/positions and thereafter, the satellite pseudoranges, carrier phase measurement, and constellation, required for simulation. The simulated pseudorange error sources corrupting GPS measurements include ionospheric delay, tropospheric delay, receiver noise and multipath. Assume that the differential GPS (DGPS) mode is available and therefore most of the receiver-independent common errors can be corrected, while the multipath and receiver thermal noise cannot be eliminated. The multipath interferences are added into the GPS pseudorange observation data during the vehicle moving. Since the research focus on the mitigation of multipath errors, the influence of measurement noise is relatively critical.

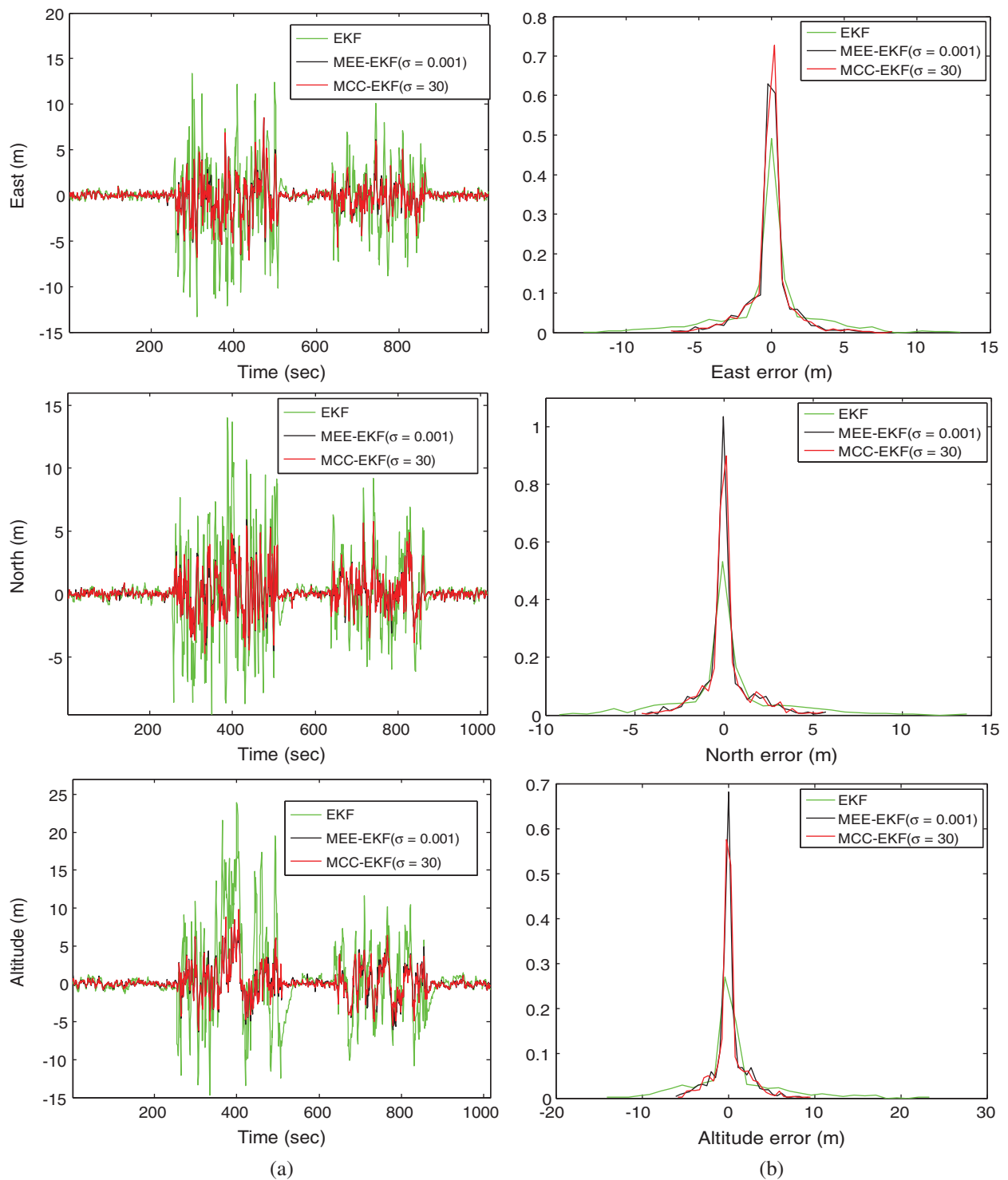
In the simulation, there are 9 GPS satellites available. The test trajectory for the simulated vehicle and the skyplot during the simulation time interval are shown as in Fig. 2. A vehicle is designed to perform the uniform accelerated motion to reduce the impact caused by unmodeling system dynamic errors. Performance comparison presented will cover two parts for each of the scenarios. Firstly, performance comparison for EKF, MEE-EKF and MCC-EKF is shown. Secondly, performance enhancement using covariance scaling is presented, where various types of approaches including EKF, AEKF MCC-EKF and MCC-AEKF are involved.



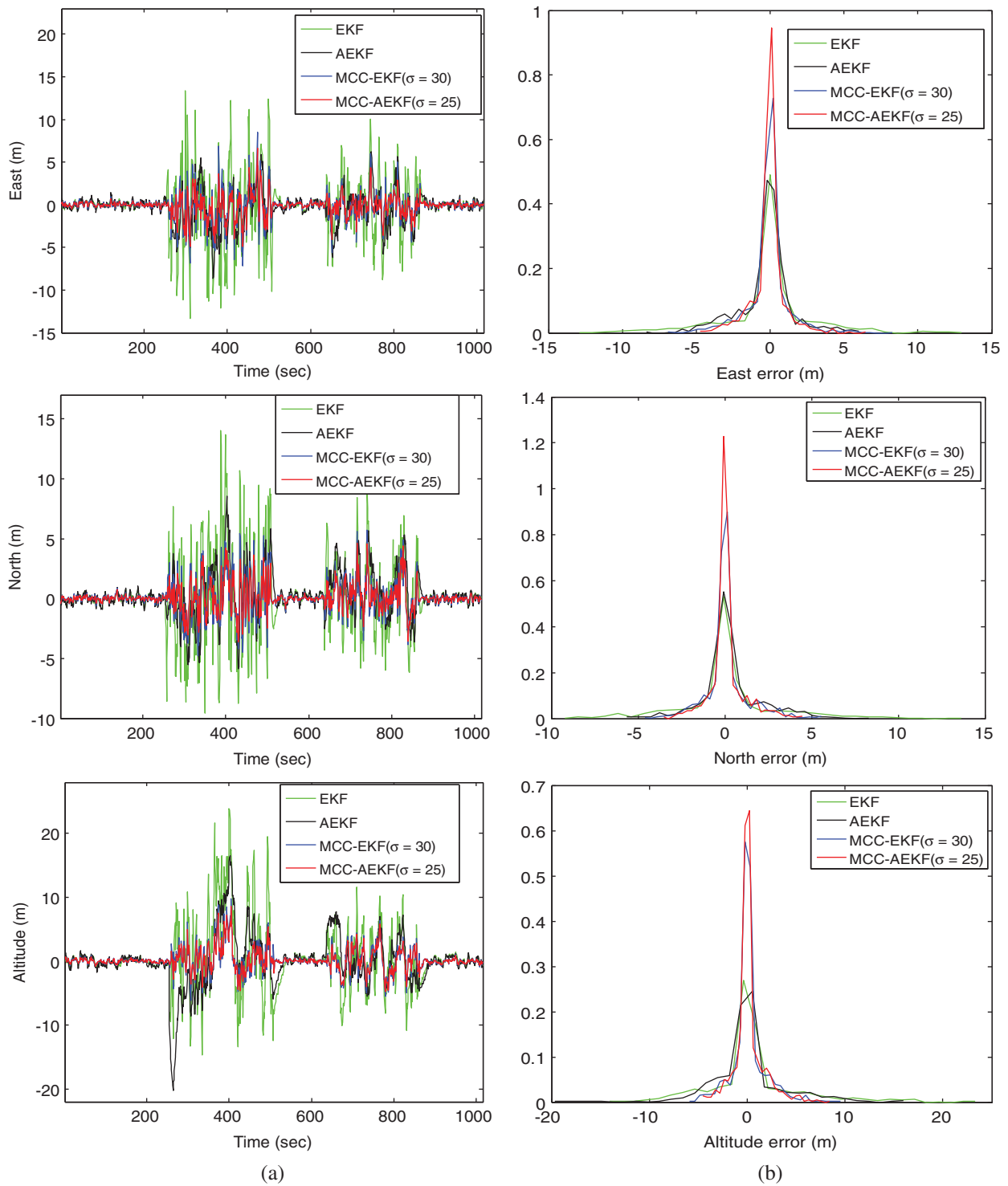
**Figure 2:** (a) Test trajectory for the simulated vehicle; and (b) the skyplot during the simulation

**Table 1:** Description of the time-varying noise strength in the five time intervals for Scenario 1

Segment	Time interval (s)	Noise variance (m <sup>2</sup> )
1	[0–254]	$N(0, 1)$
2	[255–508]	$N(0, 16)$
3	[509–635]	$N(0, 1)$
4	[636–863]	$N(0, 10)$
5	[864–1016]	$N(0, 1)$



**Figure 3:** (a) Comparison of positioning accuracy and (b) the corresponding error probability density functions (pdf's) for EKF, MEE-EKF and MCC-EKF—Scenario 1. (a) Position errors, (b) probability density functions



**Figure 4:** (a) Comparison of positioning accuracy and (b) the corresponding error pdf's for EKF, AEKF MCC-EKF and MCC-AEKF—Scenario 1. (a) Position errors, (b) probability density functions

### 5.1 Scenario 1: Environment Involving Time-Varying Variance in Measurement Noise

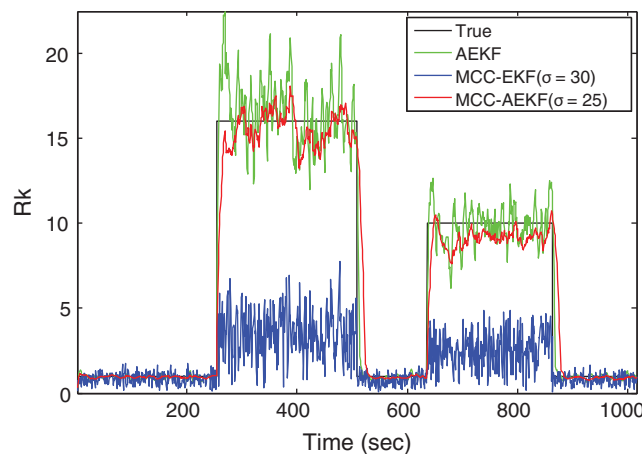
Scenario 1 is designed for investigating the performance comparison when dealing with the time-varying measurement noise statistics. Description of time varying measurement variances in the five time intervals is shown in Tab. 1. The time-varying measurement noise variances  $r_{\rho_i}$  in the five time intervals for this scenario are assumed to follows the variation:  $r_{\rho_i} = 1 \rightarrow 16 \rightarrow 1 \rightarrow 10 \rightarrow 1$  (in unit of  $m^2$ ), where the ‘arrows ( $\rightarrow$ )’ is employed for indicating the time-varying trajectory of variances. The set of unknown time-varying statistical parameters of noise needs to be estimated with the system state and the error covariance.

#### 5.1.1 Performance Comparison for EKF, MEE-EKF and MCC-EKF

Comparison of GPS navigation accuracy for the EKF, MEE-EKF and MCC-EKF is shown in Fig. 3 where the positioning accuracy comparison and the corresponding error pdf’s, respectively, are shown. The results show that both MEE and MCC based EKF can effectively improve the positioning performance. As can be seen, both the MEE and MCC can be adopted to assist the EKF to improve GPS navigation accuracy in time-varying Gaussian noise environment where the filtering performance based on the two optimization criterion lead to equivalent results with no noticeable distinction.

#### 5.1.2 Performance Enhancement Using Covariance Scaling

Comparison of positioning accuracy for the four algorithms: EKF, AEKF, MCC-EKF and MCC-AEKF is shown in Fig. 4, where both the positioning accuracy and the corresponding error pdf’s are presented. Fig. 5 shows the variation and adaptation capability of the standard deviation for the time-varying statistics in the measurement model. The MCC-EKF did not catch the variation of noise strength very well. With the assistance of AEKF, the MCC-AEKF can further improve the performance. From the other view point, the adaptation capability of noise variance for the AEKF has been improved with the assistance of the MCC mechanism. Tab. 2 provides the performance comparison for various algorithms. As compared MEE, the MCC based approach provides similar positioning accuracy with better computation efficiency. Of the various approaches, the MCC-AEKF provides the best positioning accuracy with only a little more execution time as compared to MCC-EKF.



**Figure 5:** Variation and adaptation results of the variance for the time-varying measurement noise

**Table 2:** Performance comparison for various algorithms—Scenario 1

Filter used	RMSE (m)			Execution time (s)
	East	North	Altitude	
EKF	3.0629	2.8335	4.8644	1.3278
AEKF	2.3560	2.2041	4.9047	1.3478
MEE-EKF ( $\sigma = 0.001$ )	1.4832	1.3099	1.8564	2.7931
MCC-EKF ( $\sigma = 30$ )	1.4829	1.3025	1.8000	1.7364
MCC-AEKF ( $\sigma = 25$ )	1.1137	1.0318	1.5763	1.7525

### 5.2 Scenario 2: Pseudorange Observable Involving Outlier Type of Multipath Errors

In Scenario 2, mitigation of the pseudorange observable involving outlier type of multipath interferences is discussed. There are totally five time durations where additional randomly generated errors are intentionally injected into the GPS pseudorange observation data during the vehicle moving. [Tab. 3](#) shows the information of the outliers, including the numbers of outliers and their strengths.

**Table 3:** Information for the outliers

Time interval	Number of outliers	Standard deviation of the outliers (m)
[0–254]	5	15, 100, 70, 80, 100
[255–508]	6	55, 70, 10, 22, 15, 15
[509–635]	6	55, 20, 13, 65, 34, 100
[636–863]	10	66, 90, 105, 65, 21, 90, 105, 65, 50, 80
[864–1016]	6	16, 20, 45, 15, 20, 30

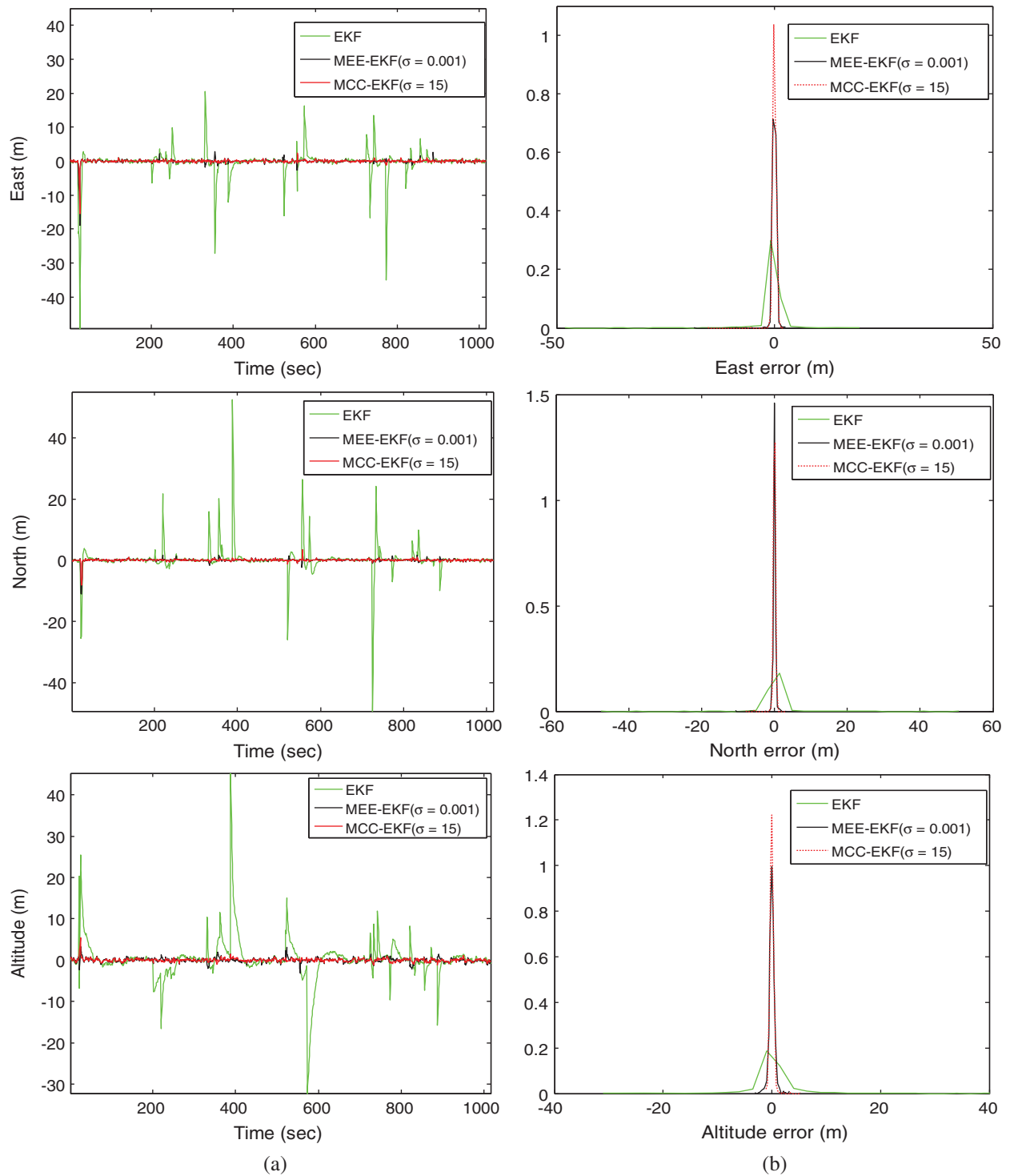
#### 5.2.1 Performance Comparison for EKF, MEE-EKF and MCC-EKF

Comparison of GPS navigation accuracy for the three schemes: EKF, MEE-EKF and MCC-EKF is shown in [Fig. 6](#) where both the comparison of positioning accuracy and the error corresponding pdf's are presented. The results show that the both the MEE and MCC can assist EKF to effectively deal with the outliers in the pseudorange observables such as multipath interferences.

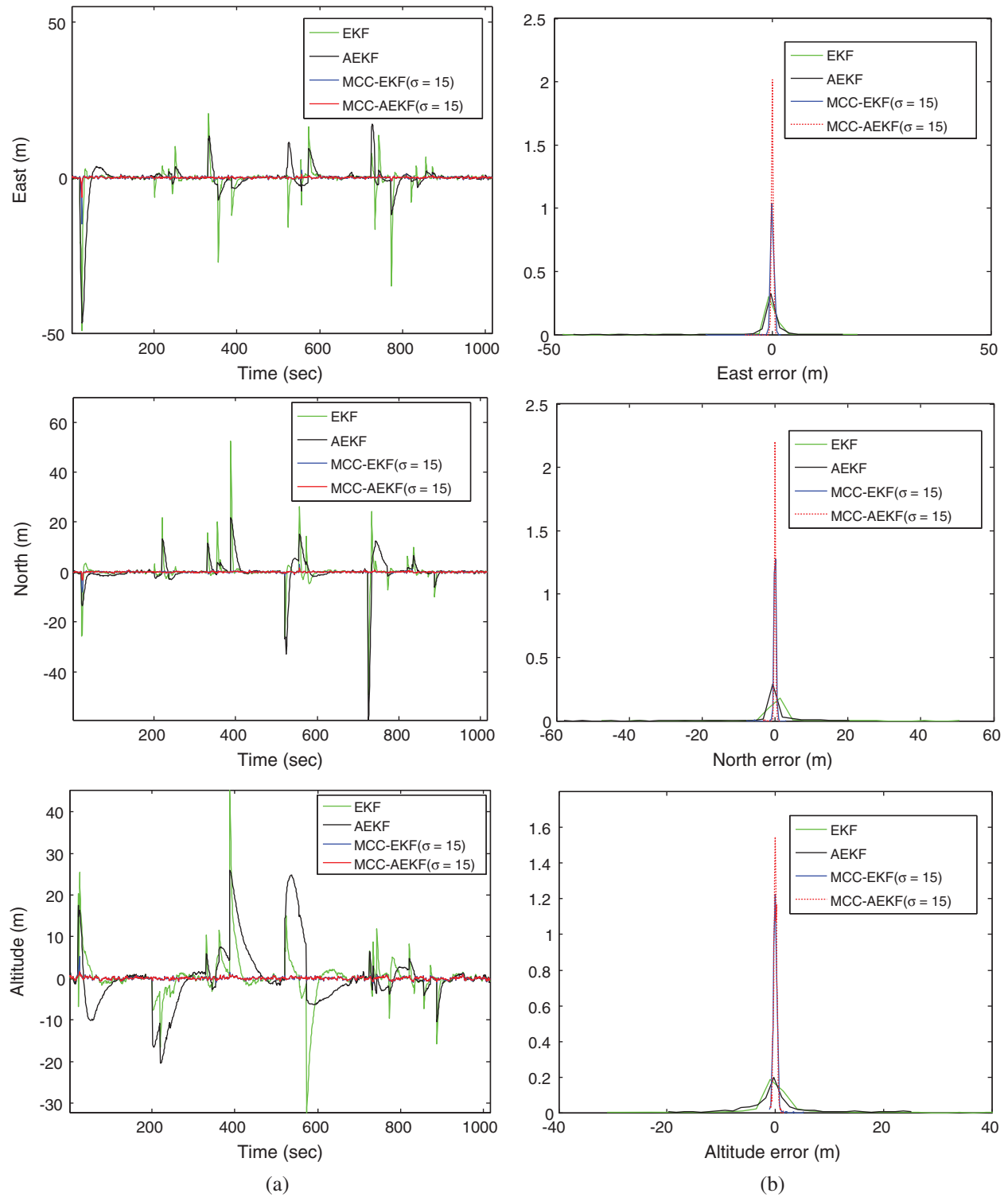
#### 5.2.2 Performance Enhancement Using Covariance Scaling

Utilization of the AEKF, referred to as the MCC-AEKF, is employed for further performance enhancement. [Fig. 7](#) illustrates the positioning accuracy and the corresponding pdf's for various algorithms: EKF, AEKF, MCC-EKF and MCC-AEKF. The AEKF does not possess sufficient capability to resolve the outlier type of interference, while the MCC-AEKF demonstrates substantial performance improvement in navigation accuracy with acceptable extra computational expense. [Tab. 4](#) summarizes the estimation performance and execution time for various algorithms.





**Figure 6:** (a) Comparison of positioning accuracy and (b) the corresponding error pdf's for EKF, MEE-EKF and MCC-EKF—Scenario 2. (a) Position errors, (b) probability density functions



**Figure 7:** (a) Positioning accuracy comparison and (b) the corresponding error pdf's for EKF, AEKF, MCC-EKF and MCC-AEKF—Scenario 2. (a) Position errors, (b) probability density functions

**Table 4:** Performance comparison for various algorithms—Scenario 2

Filter used	RMSE (m)			Execution time (s)
	East	North	Altitude	
EKF	3.9280	4.4652	5.0101	1.3412
AEKF	5.2129	5.7250	7.2116	1.4034
MEE-EKF ( $\sigma = 0.001$ )	0.6475	0.4500	0.4685	2.4639
MCC-EKF ( $\sigma = 15$ )	0.5773	0.4408	0.3996	1.7931
MCC-AEKF ( $\sigma = 15$ )	0.3040	0.2408	0.2954	1.8723

## 6 Conclusions

This paper investigates the kernel entropy principle based adaptive filtering for Global Positioning System (GPS) navigation processing. The algorithm is effective for dealing with non-Gaussian or heavy-tailed errors, such as the multipath interferences.

The standard EKF method is derived based on MSE criterion and is limited to the assumption of linearity and Gaussianity to be optimal. The robustness of nonlinear filter is improved using the optimization criterion based on entropy or correntropy. The GPS navigation algorithm based on kernel entropy related principles, including the MEE criterion and the MCC has been performed, which is especially useful for the heavy-tailed/impulsive types of interference errors. In addition, behavior of the innovation related parameters have been introduced, which are useful in designing the adaptive Kalman filter to form the MCC-AEKF for further performance improvement.

Simulation experiments for GPS navigation have been provided to illustrate the performance. Results show that the kernel entropy principle based adaptive filtering algorithm possesses noticeable improvement on navigation accuracy as compared to that of conventional methods and thus demonstrates good potential as the alternative as the GPS navigation processor, especially in the case of observables with non-Gaussian errors. Two scenarios, including (1) the environment involving time-varying measurement noise variance; and (2) the pseudorange observable involving outlier type of multipath errors, respectively, are presented for demonstration. Performance comparison for various approaches, including EKF, AEKF, MCC-EKF, MEE-EKF and MCC-AEKF have been carried out and the kernel entropy based EKF algorithm has demonstrated promising results in navigational accuracy improvement.

**Funding Statement:** This work has been partially supported by the Ministry of Science and Technology, Taiwan (Grant Number MOST 108-2221-E-019-013).

**Conflicts of Interest:** The authors declare that they have no conflicts of interest to report regarding the present study.

## References

- [1] E. D. Kaplan and C. J. Hegarty, *Understanding GPS: Principles and Applications*, Norwood, MA, USA: Artech House, Inc., 2006.
- [2] R. G. Brown and P. Y. C. Hwang, *Introduction to Random Signals and Applied Kalman Filtering*, New York, NY, USA: John Wiley & Sons, 1997.

- [3] A. H. Mohamed and K. P. Schwarz, "Adaptive kalman filtering for INS/GPS," *Journal of Geodesy*, vol. 73, pp. 193–203, 1999.
- [4] M. A. El-Sayed, A. A. Ali, M. E. Hussien and H. A. Sennary, "A multi-level threshold method for edge detection and segmentation based on entropy," *Computers, Materials & Continua*, vol. 63, no. 1, pp. 1–16, 2020.
- [5] W. Liu, J. C. Principe and S. Haykin, *Kernel Adaptive Filtering: A Comprehensive Introduction*, Hoboken, NJ, USA: John Wiley & Sons, 2010.
- [6] J. C. Principe, *Information Theoretic Learning, Renyi's Entropy and Kernel Perspectives*, New York, NY, USA: Springer, 2010.
- [7] B. Chen, L. Dang, Y. Gu, N. Zheng and J. C. Principe, "Minimum error entropy Kalman filter," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, pp. 1–11, 2019 (Early Access).
- [8] B. Chen, Z. J. Yuan, N. N. Zheng and J. C. Principe, "Kernel minimum error entropy algorithm," *Neurocomputing*, vol. 121, pp. 160–169, 2013.
- [9] M. Ren, J. Zhang, F. Fang, G. Hou and J. Xu, "Improved minimum entropy filtering for continuous nonlinear non-Gaussian systems using a generalized density evolution equation," *Entropy*, vol. 15, no. 7, pp. 2510–2523, 2013.
- [10] B. Chen, X. Liu, H. Zhao and J. C. Principe, "Maximum correntropy Kalman filter," *Automatica*, vol. 76, pp. 70–77, 2017.
- [11] B. Chen and J. C. Principe, "Maximum correntropy estimation is a smoothed MAP estimation," *IEEE Signal Processing Letters*, vol. 19, no. 8, pp. 491–494, 2012.
- [12] B. Chen, L. Xing, J. Liang, N. Zheng and J. C. Principe, "Steady-state mean-square error analysis for adaptive filtering under the maximum correntropy criterion," *IEEE Signal Processing Letters*, vol. 21, no. 7, pp. 880–884, 2014.
- [13] W. Liu, P. P. Pokharel and J. C. Principe, "Correntropy: Properties and applications in non-Gaussian signal processing," *IEEE Transactions on Signal Processing*, vol. 55, no. 11, pp. 5286–5298, 2007.
- [14] X. Liu, H. Qu, H. Zhao and B. Chen, "Extended Kalman filter under maximum correntropy criterion," in *Proc. IEEE Int. Joint Conf. on Neural Networks*, Vancouver, BC, Canada, pp. 1733–1737, 2016.
- [15] A. Singh and J. C. Principe, "Using correntropy as a cost function in linear adaptive filters," in *Proc. IEEE Proc. of the Int. Joint Conf. on Neural Networks*, Atlanta, GA, USA, pp. 2950–2955, 2009.
- [16] S. Zhao, B. Chen and J. C. Principe, "Kernel adaptive filtering with maximum correntropy criterion," in *Proc. IEEE Proc. of the Int. Joint Conf. on Neural Networks*, San Jose, CA, USA, pp. 2012–2017, 2011.
- [17] GPSSoft LLC., *Satellite Navigation Toolbox 3.0 User's Guide*, Athens, OH, USA, 2003.