

Nanosatellite Attitude Estimation using Kalman-Type Filters with Non-Gaussian Noise

Abstract— In order to control the orientation of a satellite, it is important to estimate the attitude accurately. Time series estimation is especially important in micro and nanosatellites, whose sensors are usually low-cost and have higher noise levels than high end sensors. Also, the algorithms should be able to run on systems with very restricted computer power. In this work, we evaluate five Kalman-type filtering algorithms for attitude estimation with 3-axis magnetometer and sun sensor measurements. The Kalman-type filters are selected so that each of them is designed to mitigate one error source for the unscented Kalman filter that is used as baseline. We investigate the distribution of the magnetometer noises and show that the Student's t-distribution is a better model for them than the Gaussian distribution. We consider filter responses in four operation modes: steady state, recovery from incorrect initial state, short-term sensor noise increment, and long-term increment. We find that a Kalman-type filter designed for Student's t sensor noises has the best combination of accuracy and computational speed for these problems, which leads to a conclusion that one can achieve more improvements in estimation accuracy by using a filter that can work with heavy tailed noise than by using a nonlinearity minimizing filter that assumes Gaussian noise.

Index Terms— attitude estimation; fault, nonlinear filter; nanosatellite; non-Gaussian noise.

1. INTRODUCTION

Satellites need to be oriented in space and there may be very strict requirements for the attitude estimation accuracy depending on the mission. Nanosatellites are required to maintain the attitude with attitude sensors that are typically cheap, simple, and light, and having poor accuracy. To mitigate the limitations of cheap and inaccurate sensors, better attitude estimation algorithms need to be used, and because of limited computational resources, the algorithms should have low computational complexity.

The conventional approach to attitude estimation of a satellite is to use the extended Kalman filter (EKF) [1] or its derivative-free version, the unscented Kalman filter (UKF) [2]. These algorithms are approximate solutions of the Bayesian filtering problem based on the nonlinear models of satellite motion and measurements, with Gaussian noises.

Another approach to attitude estimation, the single-frame method, was introduced in [3]–[6]. In this method, measurements at each time step are preprocessed to produce a set of linear measurements and associated covariances that are fed to the Kalman filter. The preprocessing involves the minimization of Wahba's loss function [7]. In [8], a comparison of minimization methods concludes that the singular value decomposition (SVD) and q methods are the most robust for single-frame attitude estimation methods.

Several algorithms have been proposed to improve the ability of Kalman type filters to cope with highly nonlinear situations. The filter introduced in [9] can automatically process multiple measurements in an optimized order so that the errors caused by nonlinearities are minimized.

Normal distribution based statistical estimation is vulnerable to outliers. In [10], the authors compare different geomagnetic orbit determination filters under different measurement noise distributions: normal, Student's t, and uniform in the simulations. They concluded that an unscented particle filter, which can make use of non-Gaussian noise models, is more accurate than EKF and UKF, which use Gaussian models. However, the computational complexity for particle filters is typically much greater than for Kalman-type filters.

In [11], a generalization and extension of a linear Student's t filter is proposed. In the paper, simulation results show that the proposed methods provide better accuracy than the existing methods in an application with heavy tailed process and measurement noises. In [12], an outlier-robust Kalman-type filter for nonlinear dynamical models is proposed based on Student's t-distribution in the measurement model. In the paper, the filter is compared to alternative filters in a computer simulation and is found to provide a good trade-off between accuracy and computational efficiency.

A nonlinear Kalman-type filter called maximum correntropy unscented Kalman filter (MCUKF) is proposed in [13] for spacecraft relative state estimation. Heavy-tailed non-Gaussian measurement noises are considered in the paper. The proposed MCUKF uses a non-linear regression model combined with maximum correntropy to update the measurement information. A practical test of relative motion of two spacecrafts is performed and it is found that the proposed filter gives better performance than the other filters such as EKF and UKF.

A noise-covariance adapting EKF algorithm is proposed in [14] to cope with sensor faults in the attitude estimation of a small satellite equipped with only a three-axis magnetometer. Similarly, a covariance adapting UKF with multiple measurement noise scaling factor is presented in [15] for nanosatellite attitude estimation, and is found to be more accurate than EKF, UKF, and the covariance-adapting UKF with single scaling factor. In [16], it is shown that the covariance-adapting filters are much more accurate than UKF and EKF in the faulty period and have faster recovery after the end of a fault. Furthermore, according to that study, the covariance-adapting UKF outperforms all other considered filters.

In this work, we present a computer simulation study of the performance of state-of-the-art filters for attitude estimation, taking

into consideration both accuracy and computational complexity. We show in section 3 that the magnetometer errors can be modelled with Student's t-distribution better than with a normal distribution. We consider four attitude estimation operational modes:

- attitude recovery after de-tumbling phase: this mode is simulated by initializing the filter with an incorrect initial state
- steady state, after the attitude recovery phase error has settled
- short-term noise increment in one magnetometer channel, which simulates a sensor fault due to transient disturbance
- long-term noise increment, simulating the onset of a persistent disturbance

The main purpose in this paper is to give attention on the fact that measurement noise of the magnetometers is not normal distributed as it is assumed in the most of the studies and the filtering methods should be designed considering this fact. To the best of our knowledge, there are only a few studies taking the magnetometer distributions as non-Gaussian for attitude estimation purposes and they did not use onboard magnetometer measurements to examine the data but only simulated the distributions.

The rest of the paper is organized as follows. In section 2, the filters used in this study are presented. The mathematical models of the satellite's rotational motion and of the measurements are given in section 3. In section 4, the results of the simulations performed are presented and discussed. Finally, in section 5, we summarize our conclusions and discuss possible further work.

2. FILTERING ALGORITHMS

In this section, five filtering algorithms considered for nanosatellite attitude estimation are presented. For all filters, the attitude estimation problem is formulated using the standard discrete-time nonlinear state space model

$$x_k = f(x_{k-1}) + w_k, \quad (1a)$$

$$y_k = h_k(x_k) + \varepsilon_k, \quad (1b)$$

where $f(\cdot)$ and $h(\cdot)$ are nonlinear dynamic and measurement functions respectively, which will be explained in detail in section 3, x_k is a n -vector of states at time t_k , w_k is a zero-mean noise with covariance Q , y_k is a d -vector of measurements, and ε_k is a zero-mean noise with covariance R_k . The initial state x_0 is assumed to be multivariate normal with mean μ_0 and covariance P_0 ; its probability density function (PDF) is denoted $p(x_0) = N(x_0 | \mu_0, P_0)$. The random vectors x_0, w_k, ε_k ($k = 1, 2, \dots$) are assumed to be independent. The estimation problem is to determine, at each time t_k , the conditional distribution of the state x_k given the values of the measurements obtained so far, y_1, y_2, \dots, y_k . Because this problem does not have a closed form solution, approximations are used.

We have chosen 5 different Kalman type filtering algorithms to be tested in the nanosatellite attitude estimation problem:

- UKF is the baseline solution.
- Outlier-Robust Kalman Filter (ORKF) takes the non-Gaussianity into account and updates the measurement noise covariance.
- Kullback-Leibler Partitioned Update Kalman Filter (KLPUKF) minimizes the nonlinearity.
- Single-Frame Methods based Kalman Filter (SFMKF) updates measurement noise covariance using the SFM.
- R-Adaptive Kalman Filter (RUKF) updates measurement noise covariance using the innovation sequence.

There are also other algorithms that have the same strengths e.g. Posterior Linearization Filter [17], [18] also copes well with the nonlinear problems and MCUKF [13] works well with the heavy tailed noise. We have chosen different types of filter to our tests to see what kind of filter provides the largest improvements in the estimation accuracy.

2.1. Gaussian Filter (GF) and Unscented Kalman Filter (UKF)

Many Kalman type approximations can be represented using the Gaussian filter (GF) formulation [19]. In this formulation, the filtering distribution is approximated by $p(x_k | y_{1:k}) = N(x_k | \mu_k, P_k)$, where the parameters μ_k, P_k are computed recursively in two stages (see Fig.1a):

Prediction Stage:

$$\text{Predicted mean: } \mu_k^- = \int f(x_{k-1}) N(x_{k-1} | \mu_{k-1}, P_{k-1}) dx_{k-1} \quad (2)$$

$$\text{Predicted covariance: } P_k^- = \int (f(x_{k-1}) - \mu_{k-1})(f(x_{k-1}) - \mu_{k-1})^T N(x_{k-1} | \mu_{k-1}, P_{k-1}) dx_{k-1} + Q \quad (3)$$

Update Stage:

$$\hat{y}_k = \int h_k(x_k) N(x_k | \mu_k^-, P_k^-) dx_k \quad (4)$$

$$\Psi_k = \int (x_k - \mu_k^-)(h_k(x_k) - \hat{y}_k)^T N(x_k | \mu_k^-, P_k^-) dx_k \quad (5)$$

$$\Phi_k = \int (h_k(x_k) - \hat{y}_k)(h_k(x_k) - \hat{y}_k)^T N(x_k | \mu_k^-, P_k^-) dx_k \quad (6)$$

$$\text{Innovation: } e_k = y_k - \hat{y}_k \quad (7)$$

$$\text{Innovation covariance: } S_k = \Phi_k + R_k \quad (8)$$

$$\text{Kalman gain: } K_k = \Psi_k S_k^{-1} \quad (9)$$

$$\text{Posterior mean: } \mu_k = \mu_k^- + K_k e_k \quad (10)$$

$$\text{Posterior covariance: } P_k = P_k^- - K_k S_k K_k^T \quad (11)$$

Different Kalman-type filters correspond to different ways of approximating the integrals in (2-6). The extended Kalman filter

(EKF) uses derivatives of the system and measurement functions, $f(\cdot)$ and $h(\cdot)$. The unscented Kalman filter (UKF) uses values of $f(\cdot)$ and $h(\cdot)$ evaluated on a set of points in state space ([20], [21]). In this work, we use UKF.

2.2. R-Adaptive Kalman Filter (RUKF)

Various algorithms have been proposed for adapting the Kalman filter starting with the earlier publications by Mehra [22]. One of these ideas is to scale the covariance of the filter. The parameters that are used for scaling can be calculated using different means [23], [24]. This idea is also applicable for the UKF adaptation. In [25], the state estimation performance of the UKF was improved by proper tuning of both the unscented transform parameters and the process and measurement noise covariance matrices of the dynamic system model. The R-adaptive Kalman filter was introduced with the goal of making the estimate insensitive to measurement faults [16]. The filter makes use of the empirical covariance of the windowed innovation sequence

$$\tilde{S}_k = \frac{1}{M} \sum_{j=k-M+1}^k e_j e_j^T, \quad (12)$$

where M is the width of the moving window. RUKF uses the UKF recursion (2-11) except that the measurement covariance matrix

R_k in (8) is replaced by a scaling matrix $V_k^* R_k$ where the scaling matrix is $V_k^* = \text{diag}(v_1^*, v_2^*, \dots, v_d^*)$,

$v_i^* = \max \{1, V(i,i)\}$ $i=1,\dots,d$, and $V = (\tilde{S}_k - \Phi_k) R_k^{-1}$. When a fault occurs in the i^{th} measurement the corresponding term of

$\tilde{S}_k(i,i)$ will become larger and the faulty measurement influence will decrease because of having a lower gain.

2.3. Kullback-Leibler Partitioned Update Kalman Filter (KLPUKF)

In general, the partitioned update Kalman filter performs the update stage of a multivariate measurement in several steps [9]:

1. Evaluate the nonlinearity of the measurements.
2. Minimize the nonlinearity of part of the measurement by applying a linear transformation.
3. Update the state using part of the measurement whose nonlinearity is small.
4. If the whole measurement has not yet been applied, use the partially updated state as the prior and the unused part of the measurements as a new measurement, and return to step 1.

In KLPUKF, the Kullback-Leibler divergence (KLD) of the error of approximation is used as the measure of nonlinearity. Denoting the approximated joint density of the state x_k and the measurement y_k in GF as $q(x_k, y_k)$ and the exact joint density as $p(x_k, y_k)$, the KLD of the error of the approximation is [26]

$$\eta_k = \text{KLD}(p(x_k, y_k) \| q(x_k, y_k)) = \iint p(x_k, y_k) \log \left(\frac{p(x_k, y_k)}{q(x_k, y_k)} \right) dx_k dy_k, \quad (13)$$

which can be shown [27] to be

$$\eta_k = \frac{1}{2} \log |I + R_k^{-1} \Upsilon_k|, \quad (14)$$

where

$$\Upsilon_k = \Phi_k - \Psi_k^T (P_k^{-1})^{-1} \Psi_k. \quad (15)$$

Here, it should be noted that some of the KF extensions such as EKF make linearization such that Υ_k always become zero; therefore, KLD cannot be found by using those extensions.

In [9] it is shown that the linear transformation in step 2 is the matrix

$$D_k = U_k^T \sqrt{R_k}^{-1}, \quad (16)$$

where $\sqrt{R_k}$ is the lower-triangular Cholesky factor of R_k , $(\sqrt{R_k} \sqrt{R_k}^T = R_k)$, and U_k is the orthogonal matrix in the eigenvalue decomposition

$$U_k \Lambda_k U_k^T = \sqrt{R_k}^{-1} \Upsilon_k \sqrt{R_k}^{-T}, \quad (17)$$

where diagonal matrix Λ_k has the eigenvalues in the diagonal in ascending order. The transformed measurement noise covariance is

$$\tilde{R}_k = U_k^T \sqrt{R_k}^{-1} R_k \sqrt{R_k}^{-T} U_k = I. \quad (18)$$

Here, the nonlinearity of the first measurement element is minimized by this transformation. In our implementation, we update one measurement at a time, i.e. steps 1-4 are carried out d times at each update. The algorithm scheme for KLPUKF can be seen in Fig.1b.

2.4. Outlier-Robust Kalman Filter (ORKF)

The variational Bayes based Student's t Kalman filter, ORKF [12] assumes that the measurement noises are distributed according to Student's t-distribution, which has "longer tails" than a Gaussian distribution. The Student's t distribution can be expressed using the auxiliary random variable λ_k as

$$p(y_k | x_k) = \int p(y_k | x_k, \lambda_k) p(\lambda_k) d\lambda_k, \quad (19)$$

where $y_k | x_k, \lambda_k \sim N(h(x_k), \frac{1}{\lambda_k} R_k)$ and $\lambda_k \sim \text{Gamma}\left(\frac{\nu}{2}, \frac{\nu}{2}\right)$. Here, $\nu \geq 1$ is the degrees of freedom parameter which determines

the distribution's kurtosis. If $\nu \rightarrow \infty$ the distribution converges to the normal distribution $N(h(x_k), R_k)$.

In the prediction step, $p(x_k | y_{1:k-1})$ is approximated as having a Gaussian density $N(x_k | \mu_k^-, P_k^-)$ as in (2-3). The update step with non-Gaussian noise is computationally intractable so the ORKF uses a Variational-Bayes approximation as follows. The KLD between the true posterior and the product approximation $p(x_k, \lambda_k | y_{1:k}) \approx q(x_k)q(\lambda_k)$ is

$$\text{KLD}\left(q(x_k)q(\lambda_k) \| p(x_k, \lambda_k | y_{1:k-1})\right) = \int q(x_k)q(\lambda_k) \log \left(\frac{q(x_k)q(\lambda_k)}{p(x_k, \lambda_k | y_{1:k})} \right) dx_k d\lambda_k. \quad (20)$$

The KLD is minimized with respect to $q(\lambda_k)$ while keeping $q(x_k)$ fixed by setting

$$\begin{aligned} \log q(\lambda_k) &= E_x (\log p(x_k, \lambda_k, y_{1:k})) + \text{constant} \\ &= -\frac{1}{2} \lambda_k \bar{\gamma}_k + \left(\frac{v+d}{2} - 1 \right) \log \lambda_k - \frac{v \lambda_k}{2} + \text{constant}, \end{aligned} \quad (21)$$

where

$$\begin{aligned} \bar{\gamma}_k &= E_x \left((y_k - h(x_k))^T R_k^{-1} (y_k - h(x_k)) \right) \\ &= \text{tr} \left\{ E_x \left((y_k - h(x_k)) (y_k - h(x_k))^T \right) R_k^{-1} \right\}, \end{aligned} \quad (22)$$

where $E(\cdot)$ is the expectation operator. Similarly, the KLD is minimized with respect to $q(x_k)$ while keeping $q(\lambda_k)$ fixed by setting

$$\begin{aligned} \log q(x_k) &= E_\lambda (\log p(x_k, \lambda_k, y_{1:k})) + \text{constant} \\ &= -\frac{1}{2} \bar{\lambda}_k (y_k - h(x_k))^T R_k^{-1} (y_k - h(x_k)) - (x_k - \mu_k^-)^T (P_k^-)^{-1} (x_k - \mu_k^-) + \text{constant}, \end{aligned} \quad (23)$$

where

$$\bar{\lambda}_k = E_\lambda (\lambda_k) = \int \lambda_k q(\lambda_k) d\lambda_k = \frac{v+d}{v+\bar{\gamma}_k}. \quad (24)$$

Then, (22) and (24) should be iterated to convergence or a fixed number of times. In the update step of the GF (4-11) is carried out

with $\frac{1}{\bar{\lambda}_k} R_k$ measurement noise covariance which is updated in every iteration and every time step (see Fig.1c).

2.5. Single-Frame Methods based Kalman Filter (SFMKF)

As a pre-processing step before the update stage, the linear attitude measurements are obtained from single-frame method (SFM) by minimizing Wahba's loss function [7]

$$L(A_k) = \frac{1}{2} \sum_j a_{j_k} |b_{j_k} - A_k r_{j_k}|^2, \quad (25)$$

where A is the orthogonal transformation matrix from reference coordinates to body coordinates fixed on the sensor body (Section

3.1), a_j is the inverse variance of sensor j , b_j is the unit vector of the 3-axis sensor measurements in the body coordinates, and r_j is the unit vector of the measurement model in the reference coordinates (Section 3.2). Those linear attitude angle measurements can be used in the Gaussian filter instead of the nonlinear measurements. The loss function can be minimized using the SVD method, which has the best robustness compared to other minimization methods such as QUEST, q, FOAM [4], [28]–[30]. Therefore, SVD method is used as an SFM in this paper.

The attitude measurement vector from SFM is,

$$\tilde{y}_k = Hx_k + \varepsilon_k, \quad (26)$$

where the measurement matrix is taken as $H = [I_{3 \times 3} \quad 0_{3 \times 3}]$ in this study. However, it should be noted that if another SFM is used with attitude representation other than Euler angles, the determined attitude should be represented as Euler angles after necessary transformations. The GF with update based on this linear measurement model is called SFMKF (see Fig.2 for the algorithm scheme). In the update stage, measurements \tilde{y}_k are coming from SFM with measurement covariance R_k , which is updated in each step based on the single-frame method. In the GF, the innovation (9) is replaced by

$$e_k = \tilde{y}_k - \hat{y}_k \quad (27)$$

and Eq. (7) can be replaced by

$$S_k = HP_k^{-1}H^T + R_k. \quad (28)$$

It should also be noted that the attitude measurements from SFM can be used as the initial attitude in the SFMKF. Therefore, this method might overcome the problem of having grossly incorrect initial values which affect the transient behavior of the filter.

3. MATHEMATICAL MODELS FOR NANOSATELLITE'S ROTATIONAL MOTION

3.1. Equations of Rotational Motion

The state vector for the continuous-time dynamic model is composed of attitude angles and angular velocities as

$$x = [\varphi \quad \theta \quad \psi \quad \omega_x \quad \omega_y \quad \omega_z]^T, \quad (29)$$

where φ is roll (angle of rotation about x axis), θ is pitch (y axis), ψ is yaw (z axis) with respect to the orbital frame, ω_x , ω_y , ω_z are the angular velocities in the body axis set with respect to the inertial coordinate system. The transformation matrix from reference (orbital) coordinates to body coordinates is [31, p. 764]

$$A(x) = A = \begin{bmatrix} c(\theta)c(\varphi) & c(\theta)s(\varphi) & -s(\theta) \\ -c(\psi)s(\varphi)+s(\psi)s(\theta)c(\varphi) & c(\psi)c(\varphi)+s(\psi)s(\theta)s(\varphi) & s(\psi)c(\theta) \\ s(\psi)s(\varphi)+c(\psi)s(\theta)c(\varphi) & -s(\psi)c(\varphi)+c(\psi)s(\theta)s(\varphi) & c(\psi)c(\theta) \end{bmatrix}, \quad (30)$$

where $c(\cdot)$ and $s(\cdot)$ are cosine and sine functions. The dynamic equations, obtained by the principle of conservation of angular momentum, are [31, p. 522]

$$J_x \frac{d\omega_x}{dt} = N_x + (J_y - J_z)\omega_y\omega_z, \quad (31.a)$$

$$J_y \frac{d\omega_y}{dt} = N_y + (J_z - J_x)\omega_z\omega_x, \quad (31.b)$$

$$J_z \frac{d\omega_z}{dt} = N_z + (J_x - J_y)\omega_x\omega_y, \quad (31.c)$$

where (J_x, J_y, J_z) are the principal moments of inertia, and (N_x, N_y, N_z) are the external disturbances. If only ideal spherical body gravity of Earth is taken into consideration, the external disturbances are [32, p. 238]

$$\begin{bmatrix} N_x \\ N_y \\ N_z \end{bmatrix} = -3 \frac{\mu^*}{r_0^3} \begin{bmatrix} (J_y - J_z)A_{23}A_{33} \\ (J_z - J_x)A_{13}A_{33} \\ (J_x - J_y)A_{13}A_{23} \end{bmatrix}, \quad (32)$$

where μ^* is the product of the universal gravitational constant and the mass of Earth, and r_0 is the distance between the satellite and Earth's center of mass. The angular velocity vector in the body axes, with respect to the reference (orbital) coordinate axes, is

$$\begin{bmatrix} p \\ q \\ r \end{bmatrix} = \begin{bmatrix} \omega_x \\ \omega_y \\ \omega_z \end{bmatrix} - A \begin{bmatrix} 0 \\ -\omega_o \\ 0 \end{bmatrix}, \quad (33)$$

where ω_o is the orbital angular velocity with respect to inertial reference, $\omega_o = \left(\frac{\mu^*}{r_0^3} \right)^{1/2}$ for a circular orbit. The satellite's

kinematics equation of motion derived using the Euler angle attitude representation is [31, p. 765],

$$\begin{bmatrix} \dot{\phi} \\ \dot{\theta} \\ \dot{\psi} \end{bmatrix} = \begin{bmatrix} 0 & s(\psi)/c(\theta) & c(\psi)/c(\theta) \\ 0 & c(\psi) & -s(\psi) \\ 1 & s(\psi)s(\theta)/c(\theta) & c(\psi)s(\theta)/c(\theta) \end{bmatrix} \begin{bmatrix} p \\ q \\ r \end{bmatrix}. \quad (34)$$

The system dynamics function f in (1a) is obtained by time-discretization of the differential equations (34) and (31) using Euler's method with constant time step, Δt :

$$\varphi_{k+1} = \varphi_k + \Delta t \left(\frac{s(\psi_k)}{c(\theta_k)} q_k + \frac{c(\psi_k)}{c(\theta_k)} r_k \right), \quad (35a)$$

$$\theta_{k+1} = \theta_k + \Delta t (c(\psi_k)q_k - s(\psi_k)r_k), \quad (35b)$$

$$\psi_{k+1} = \psi_k + \Delta t \left(p_k + s(\psi_k) \frac{s(\theta_k)}{c(\theta_k)} q_k + c(\psi_k) \frac{s(\theta_k)}{c(\theta_k)} r_k \right), \quad (35\text{c})$$

$$\omega_{x_{k+1}} = \omega_{x_k} + \Delta t \left(N_{x_k} + \omega_{y_k} \omega_{z_k} \right) (J_y - J_z) / J_x, \quad (35\text{d})$$

$$\omega_{y_{k+1}} = \omega_{y_k} + \Delta t \left(N_{y_k} + \omega_{z_k} \omega_{x_k} \right) (J_z - J_x) / J_y, \quad (35\text{e})$$

$$\omega_{z_{k+1}} = \omega_{z_k} + \Delta t \left(N_{z_k} + \omega_{x_k} \omega_{y_k} \right) (J_x - J_y) / J_z. \quad (35\text{f})$$

3.2. Measurement Models

In this study, a nanosatellite system is considered having as measurement sensors a sun sensor and a magnetometer. These sensors are commonly used for satellite missions in low Earth orbit.

IGRF (International Geomagnetic Reference Field) is used as the geomagnetic field model; its inputs are date and position of the satellite [31, pp. 779–782], [33]:

$$B_o(t, \bar{r}, \text{colat}, \text{lon}) = -\nabla \left\{ a \sum_{n=1}^N \sum_{m=0}^n \left(\frac{a}{\bar{r}} \right)^{n+1} \times \left[g_n^m(t) \cos(m \text{ lon}(t)) + h_n^m(t) \sin(m \text{ lon}(t)) P_n^m(\cos(\text{colat}(t))) \right] \right\}, \quad (36)$$

where B_o is the predicted magnetic field in nanoTesla (nT), \bar{r} is the distance between the mass centers of the satellite and Earth, $\text{colat}(t)$ is the co-latitude, $\text{lon}(t)$ is the longitude, $P_n^m(\cos(\text{colat}(t)))$ are the Schmidt quasi-normalized associated Legendre polynomials of degree n and order m , $a = 6371.2$ km is the geomagnetic conventional Earth's mean reference spherical radius, g_n^m and h_n^m are the Gaussian coefficients given in units of nT [33]. The noiseless magnetometer measurements are obtained by transforming the geomagnetic field to the body frame:

$$B_m(t, x) = A(x) B_o(t, \bar{r}(t), \text{colat}(t), \text{lon}(t)). \quad (37)$$

As the distance between the satellite and Earth's centers is negligible considering the distance between the Sun and Earth's centers, Sun direction with respect to the Earth-centered inertial (ECI) coordinates only depends on the Julian Day (T_{TDB}) and not on the satellite's position. The Julian Day can be derived from the satellite's reference epoch and the onboard clock time, t . The ecliptic longitude of the Sun ($\lambda_{\text{ecliptic}}$) and its linear model (ε) can be computed as [34, pp. 181–184].

$$M_{\text{Sun}} = 357.5277233^\circ + 35999.05034 T_{TDB}, \quad (38.\text{a})$$

$$\lambda_{\text{ecliptic}} = \lambda_{M_{\text{Sun}}} + 1.914666471^\circ \sin(M_{\text{Sun}}) + 0.019994643 \sin(2M_{\text{Sun}}), \quad (38.\text{b})$$

$$\lambda_{M_{\text{Sun}}} = 280.4606184^\circ + 36000.77005361 T_{TDB}, \quad (38.\text{c})$$

$$\varepsilon = 23.439291^\circ - 0.0130042T_{TDB}, \quad (38.d)$$

where the mean anomaly and the mean longitude of the Sun are given by (38a), (38c), and sind is degree-argument sine function.

From these relations, Sun direction vector (S_{ECI}) in ECI coordinates can be found from

$$S_{ECI}(t) = \begin{bmatrix} \cosd(\lambda_{\text{ecliptic}}) \\ \sind(\lambda_{\text{ecliptic}})\cosd(\varepsilon) \\ \sind(\lambda_{\text{ecliptic}})\sind(\varepsilon) \end{bmatrix}. \quad (39)$$

The noiseless sun sensor measurements S_m are obtained by transforming the Sun direction to the body frame:

$$S_m(t, x) = A(x)L(t)S_{ECI}(t), \quad (40)$$

where L is the transformation matrix from ECI to orbit coordinates. We compute the satellite's orbital elements (L) and position on the orbit $(\bar{r}, \text{colat}, \text{lon})$ using the SGP4 orbital propagation model [35]^{*}. For consistency, sun direction and magnetic field measurements are used as unit vectors, and the state space model's (1) measurement function is then

$$h_k(x) = \begin{bmatrix} B_m(t_k, x) \\ \|B_m(t_k, x)\| \\ S_m(t_k, x) \end{bmatrix}. \quad (41)$$

3.3. Magnetometer Noise Distribution Analysis

It is investigated which distribution is best fitting for the magnetometer noises. As representative magnetometer measurements we use 11 nonconsecutive days' SWARM-A data from the ESA website (<ftp://swarm-diss.eo.esa.int/>) and we use the magnetic field model IGRF as a reference. The differences between the unit vectors of real magnetometer measurements of SWARM-A and the reference magnetic field model IGRF in 13th order having the date and position inputs from the satellite are used in order to construct the noise density. Three symmetric distributions are fitted to these errors using standard maximum likelihood methods: normal, t, and extreme value distribution. The PDF of the t-distribution shows the best agreement with the density histogram (Fig.3). This is also the distribution that gives the smallest Akaike information criterion (Table I), which is defined as [36]

$$AIC_i = \ln(MSE_i) + \frac{2\bar{k}_i}{N_{mag}}, \quad (42)$$

where MSE_i is mean square error of the i^{th} candidate distribution, \bar{k}_i is the number of fitting parameters, N_{mag} is the number of the observations. Therefore, t-distributed noise on magnetometer measurements are used for the simulations (Section 4). In the analysis, the non-normalized magnetic field noises are in the order of 100-200 nT.

* The code is available at <http://www.centerforspace.com/downloads/>.

4. ANALYSIS AND RESULTS

In this section, we describe the numerical simulation tests and the results obtained. We begin by presenting the parameters common to all tests. We consider a nanosatellite with principal moments of inertia $J = \text{diag} [2.1 \times 10^{-3} \quad 2.0 \times 10^{-3} \quad 1.9 \times 10^{-3}] \text{ kg m}^2$. The orbit is almost circular with inclination $i = 87.4^\circ$, eccentricity $e = 0.0009$, and average altitude 500 km. Sun sensors and magnetometers have 3-axis measurements and both sensors have 1-Hz frequency; estimation time step is also 1 sec. The sensor noises are characterized using standard deviations $\sigma_B = 0.008$ for each measurement of the magnetometers and $\sigma_S = 0.002$ for sun sensors which is based on the angular deviation of sun vector with respect to the actual direction of the Sun. The noises for magnetometer measurements are sampled from Student's t distribution and sun sensors

are sampled from normal. The dynamic system's noise covariance is selected as $Q = \begin{bmatrix} 10^{-4} (\text{rad})^2 I_{3 \times 3} & 0 \\ 0 & 10^{-6} (\text{rad/s})^2 I_{3 \times 3} \end{bmatrix}$. The

non-zero Q matrix is introduced to avoid numerical issues with the approximate nonlinear covariance propagation. We consider attitude estimation over a single orbit (6000 s) starting at position $\text{colat}_0 = 24^\circ$, $\text{lon}_0 = 173^\circ$, $\bar{r}_0 = 6878 \text{ km}$ and time 00:00:00 UT at January 1, 2014. The true initial state is $x_0^{true} = [0.015 \quad 0.01 \quad 0.005 \quad 0.0005 \quad 0.00075 \quad 0.0005]^T$ with the units of the first three elements in radians and the last in radians per seconds, and the initial state of filters is

$$x_0 = 2x_0^{true} \quad (43)$$

(except SFMKF, which uses SFM). If process noise is applied to the true states in the simulation environment, the outputs will be different in different simulations. We wanted to concentrate only on the effect of the measurement noise and, thus, the true track is generated using (35) without added noise.

Filter algorithm parameters:

UKF: unscented transformation parameters from [20], $\alpha = 10^{-3}$, $\beta = 2$, $\kappa = 0$,

RUKF: sliding window width, $M = 20$,

ORKF: number of iterations, N_{iter} , and degrees of freedom parameter, $v = 4$.

4.1. Normal Operation Mode

Error formulation used for the evaluation of the filters' attitude angle estimation accuracies in Fig.4 is normalized root mean square of the norm of estimation errors $\sqrt{\sum_{k=1}^n \|e_k\|^2} / \sqrt{\sum_{k=1}^n \|X_k^{true}\|^2}$ where $e_k = X_k^{true} - X_k^{est}$ is the absolute error vector between the true and the estimated data during the whole simulation time. Simulations are performed on a computer with Intel® Core™ i7

@2.93 GHz CPU and 8 GB RAM. Fig.4 shows the work-accuracy diagram for the 5 filter algorithms with one of them having 3 variants. The x-axis shows the average NRMSE from 100 simulations with different random number generator (RNG) seeds for the sensor noises. There is not much change in the results when the parameters are changed in the filters other than ORKF. Therefore, only ORKF is applied using $N_{iter} = \{2, 5, 10\}$ and the iteration number is shown at the filter label's subscript. KLPUKF, ORKF_s and ORKF₁₀ are the most accurate algorithms, with KLPUKF and ORKF variations computationally the heaviest. The other three have smaller, roughly equal, computational load, but SFMKF and RUKF's accuracy, which is roughly equal, is much better than UKF's. For ORKF, ORKF_s was selected and named without the subscript for the rest of the paper in order to have accurate but computationally light estimations. It should be noted that the non-Gaussianity seems to be the dominating factor compared to non-linearity that cause error to the UKF.

A time series of estimation errors from a single simulation (Fig. 5) shows that all 5 filters behave similarly. Most of the error arises in the initial transient during which the filters recover from the incorrect initial state. Even SFMKF, which uses an initial state from SFM instead of (42), has a significant transient error. In order to evaluate the steady state where errors arise mainly from sensor errors, we consider also the errors after 1500 s. Box plots (Fig. 6) and RMSE (Table II) indicate that all filters perform reasonably well in the steady state regime, with ORKF having the smallest attitude RMSE.

4.2. Operation with Noise-Increment Magnetometer Fault

Magnetometers on satellite can easily be disturbed by other subsystems which are very close because of the small size of the nanosatellite. There can also be external disturbances. Those effects result in faults on the measurements. Generally, two types of error models have been widely adopted to describe frequently encountered faults: the continuous bias model and the noise increment error model. For disturbances having continuous bias on the measurements, sensors should be calibrated first. Here, noise increment type of magnetometer measurement fault is considered. Here, it is considered that a sensor-related fault occurred on the magnetometer's y channel, making the measurements noisy on that channel. The y magnetometer component $\varepsilon_{B_k}^y$ of the measurement noise term ε_k in (1b) is replaced by $\delta_k \varepsilon_{B_k}^y$, where δ_k is a scaling sequence. Two fault cases are considered: long-term fault

$$\delta_k = \begin{cases} 10 & \text{if } t > 4000 \\ 1 & \text{otherwise} \end{cases}, \quad (44)$$

and transient (short-term) fault

$$\delta_k = \begin{cases} 10 & \text{if } 3000 < t < 3400 \\ 1 & \text{otherwise} \end{cases}. \quad (45)$$

Fig.7 show the time series of the y-magnetometer component of the RUKF scaling matrix derived from the windowed innovation sequence (12). These show that the RUKF scale factor is adapting correctly: it changes from ~1 to ~100 during the fault. The scale factors for normal and Student's t distribution cases without considering any faults are presented in Fig.8 in order to see how RUKF copes with the heavy tails. We see that RUKF compensates the heavy tails by increasing the scale factor. It should be noted that the x and z axes are similar to the y axis outputs. Fig.9 and Table III summarize the estimation errors for different filters. From these we see

- UKF and KLPKF are generally the least accurate methods, especially for long-term fault. That's because these filters assume Gaussian noise and are not adaptive. Also, UKF is generally very slow in recovering after the end of the short-term fault.
- The remaining filters cope fairly well with both kinds of fault. ORKF is the most accurate, RUKF is almost as accurate, and SFMKF is the least accurate of the three. These three filters are adaptive in measurement noise covariance. So, the filter copes with faulty measurement by adapting the scaling factors and having a lower gain.

5. CONCLUSIONS

From this study, we find:

1. Real magnetometer noises are better modelled by Student's t distribution than normal distribution.
2. SFMKF, UKF, RUKF are the computationally lightest filters, KLPKF and ORKF are 2-4 times slower.
3. In normal operation mode, ORKF and KLPKF are the most accurate; SFMKF and RUKF errors are about twice as large. UKF has much larger errors because it does not cope well with t-distributed noises.
4. ORKF and RUKF cope best with transient and long-term noise-increment faults in magnetometer. SFMKF copes reasonably well.

In summary ORKF is the most accurate filter both for the normal operation mode and for the operation with noise-increment magnetometer fault. We identified that the outliers caused the difference on the estimation accuracy. We suggest that it is important to check the noise distribution of the magnetometers. If they are non-Gaussian as with the tested magnetometers the non-Gaussianity is a more significant source for estimation errors than nonlinearity and it is more important to use the limited computational resources in nanosatellites to compensate for non-Gaussianity than using more complex nonlinear filters.

For further work, algorithms can be tested for sensor faults of continuous bias type. This would require a calibration in which the state vector has more elements to be estimated. Other disturbances such as extraterrestrial magnetic field can be included in the models and results using different reduced order IGRF models can be compared. In addition to these, different Student's t filters and the filter proposed in [37], which is a Kalman-type filter for measurements having non-Gaussian noises, could be implemented and compared with other considered filters for further comparative studies. Another interesting possibility would be to investigate

whether the KLPUKF and ORKF algorithms could be fused into an algorithm that minimizes the nonlinearity and can cope with the heavy tailed noises.

ACKNOWLEDGMENT

The data of SWARM-A were obtained from the website (<ftp://swarm-diss.eo.esa.int/>). We thank the data providers for providing access to the data.

REFERENCES

- [1] R. DaForno, F. Reali, S. Bristor, and S. Debei, “Autonomous Navigation of MegSat1: Attitude, Sensor Bias and Scale Factor Estimation by EKF and Magnetometer-Only Measurement,” in *22nd AIAA International Communications Satellite Systems Conference and Exhibit*, 2004.
- [2] P. Sekhavat, Q. Gong, and I. M. Ross, “NPSAT 1 Parameter Estimation using Unscented Kalman Filter,” in *Proceedings of the 2007 American Control Conference*, 2007, pp. 4445–4451.
- [3] C. Hajiiev and M. Bahar, “Attitude Determination and Control System Design of the ITU-UUBF LEO1 Satellite,” *Acta Astronaut.*, vol. 52, no. 2–6, pp. 493–499, 2003.
- [4] K. Vinther, K. F. Jensen, J. A. Larsen, and R. Wisniewski, “Inexpensive Cubesat Attitude Estimation Using Quaternions And Unscented Kalman Filtering,” *Autom. Control Aerosp.*, vol. 4, no. 1, 2011.
- [5] R. Zanetti, T. Ainscough, J. Christian, and P. D. Spanos, “Q Method Extended Kalman Filter,” *NASA Technical Reports*, 2012. [Online]. Available: <http://ntrs.nasa.gov/archive/nasa/casi.ntrs.nasa.gov/20120017927.pdf>.
- [6] C. Hajiiev, D. Cilden, and Y. Somov, “Gyro-Free Attitude and Rate Estimation for a Small Satellite using SVD and EKF,” *Aerosp. Sci. Technol.*, vol. 55, pp. 324–331, 2016.
- [7] G. Wahba, “Problem 65-1: A Least Squares Estimate of Satellite Attitude,” *Soc. Ind. Appl. Math. Rev.*, vol. 7, no. 3, p. 409, 1965.
- [8] Y. Oshman and F. L. Markley, “Sequential gyroless attitude and attitude-rate estimation from vector observations,” *Acta Astronaut.*, vol. 46, no. 7, pp. 449–463, Apr. 2000.
- [9] M. Raitoharju, Á. García-Fernández, and R. Piche, “Kullback-Leibler Divergence Approach to Partitioned Update Kalman Filter,” *Signal Processing*, 2016.
- [10] X. Ning, X. Ma, C. Peng, W. Quan, and J. Fang, “Analysis of filtering methods for satellite autonomous orbit determination using celestial and geomagnetic measurement,” *Math. Probl. Eng.*, vol. 2012, pp. 1–16, Feb. 2012.
- [11] Y. Huang, Y. Zhang, N. Li, and J. Chambers, “Robust student’s t based nonlinear filter and smoother,” *IEEE Trans. Aerosp.*

Electron. Syst., vol. 52, no. 5, pp. 2586–2596, Oct. 2016.

- [12] R. Piche, S. Sarkka, and J. Hartikainen, “Recursive outlier-robust filtering and smoothing for nonlinear systems using the multivariate student-t distribution,” in *2012 IEEE International Workshop on Machine Learning for Signal Processing*, 2012, pp. 1–6.
- [13] X. Liu, H. Qu, J. Zhao, P. Yue, and M. Wang, “Maximum Correntropy Unscented Kalman Filter for Spacecraft Relative State Estimation.,” *Sensors*, vol. 16, no. 9, Sep. 2016.
- [14] Y. Xing, S. Zhang, J. Zhang, and X. Cao, “Robust-extended Kalman filter for small satellite attitude estimation in the presence of measurement uncertainties and faults,” *Proc. Inst. Mech. Eng. Part G J. Aerosp. Eng.*, vol. 226, no. 1, pp. 30–41, Jan. 2012.
- [15] D. Lee, G. Vukovich, and R. Lee, “Robust unscented Kalman filter for nanosat attitude estimation,” *Int. J. Control. Autom. Syst.*, vol. 15, no. 5, pp. 2161–2173, Oct. 2017.
- [16] H. E. Soken, C. Hajiye, and S. Sakai, “Robust Kalman Filtering for Small Satellite Attitude Estimation in the Presence of Measurement Faults,” *Eur. J. Control*, vol. 20, no. 2, pp. 64–72, 2014.
- [17] A. F. Garcia-Fernandez, L. Svensson, M. R. Morelande, and S. Sarkka, “Posterior Linearization Filter: Principles and Implementation Using Sigma Points,” *IEEE Trans. Signal Process.*, vol. 63, no. 20, pp. 5561–5573, Oct. 2015.
- [18] M. Raitoharju, L. Svensson, A. F. Garcia-Fernandez, and R. Piche, “Damped Posterior Linearization Filter,” *IEEE Signal Process. Lett.*, pp. 1–1, 2018.
- [19] K. Ito and K. Xiong, “Gaussian filters for nonlinear filtering problems,” *IEEE Trans. Automat. Contr.*, vol. 45, no. 5, pp. 910–927, May 2000.
- [20] E. A. Wan and R. Van Der Merwe, “The unscented Kalman filter for nonlinear estimation,” in *Proceedings of the IEEE 2000 Adaptive Systems for Signal Processing, Communications, and Control Symposium (Cat. No.00EX373)*, 2000, pp. 153–158.
- [21] S. J. Julier and J. K. Uhlmann, “New extension of the Kalman filter to nonlinear systems,” in *Proc. SPIE 3068, Signal Processing, Sensor Fusion, and Target Recognition VI*, 1997, vol. 3068.
- [22] R. K. Mehra, “On the Identification of Variances and Adaptive Filtering,” *IEEE Trans. Automat. Contr.*, vol. 15, no. 2, pp. 175–184, 1970.
- [23] C. Hajiye, “Fault tolerant integrated radar/inertial altimeter based on Nonlinear Robust Adaptive Kalman filter,” *Aerospace Sci. Technol.*, vol. 17, no. 1, pp. 40–49, Mar. 2012.
- [24] C. H. Kang, S. Y. Kim, and C. G. Park, “A GNSS Interference Identification and Tracking based on Adaptive Fading Kalman Filter,” *IFAC Proc. Vol.*, vol. 47, no. 3, pp. 3250–3255, Jan. 2014.

- [25] L. A. Scardua and J. J. da Cruz, “Complete Offline Tuning of the Unscented Kalman Filter,” *Automatica*, vol. 80, pp. 54–61, Jun. 2017.
- [26] S. Kullback and R. A. Leibler, “On Information and Sufficiency,” *Ann. Math. Stat.*, vol. 22, no. 1, pp. 79–86, Mar. 1951.
- [27] M. R. Morelande and A. F. Garcia-Fernandez, “Analysis of Kalman Filter Approximations for Nonlinear Measurements,” *IEEE Trans. Signal Process.*, vol. 61, no. 22, pp. 5477–5484, Nov. 2013.
- [28] F. L. Markley, “Attitude Determination Using Vector Observations and Singular Value Decomposition,” *J. Astronaut. Sci.*, vol. 36, no. 3, pp. 245–258, 1988.
- [29] F. L. Markley, “Attitude Determination Using Vector Observations: a Fast Optimal Matrix Algorithm,” *J. Astronaut. Sci.*, vol. 41, no. 2, pp. 261–280, 1993.
- [30] D. Cilden-Guler, E. S. Conguroglu, and C. Hajiiev, “Single-Frame Attitude Determination Methods for Nanosatellites,” *Metrologia*, vol. 54, no. 3, pp. 313–324, 2017.
- [31] J. R. Wertz, *Spacecraft Attitude Determination and Control*, no. 73. Dordrecht, Holland: Kluwer Academic Publishers, 1988.
- [32] P. C. Hughes, *Spacecraft Attitude Dynamics*. Dover Publications, 2004.
- [33] E. Thébault, C. C. Finlay, C. D. Beggan, P. Alken, and E. Al., “International Geomagnetic Reference Field: the 12th generation,” *Earth, Planets Sp.*, vol. 67:69, 2015.
- [34] D. A. Vallado, *Fundamentals of Astrodynamics and Applications (3rd Ed.)*, vol. 21. USA: Microcosm Press/Springer, 2007.
- [35] D. A. Vallado and P. Crawford, “SGP4 orbit determination,” in *AIAA/AAS Astrodynamics Specialist Conf.*, 2008.
- [36] M. D. Fraser, *Network models for control and processing*. Intellect, 2000.
- [37] M. Raitoharju, R. Piche, and H. Nurminen, “A Systematic Approach for Kalman-type Filtering with non-Gaussian Noises,” in *Proceedings of the 19th International Conference on Information Fusion (Fusion 2016)*, IEEE, 2016.

FIGURES AND TABLES

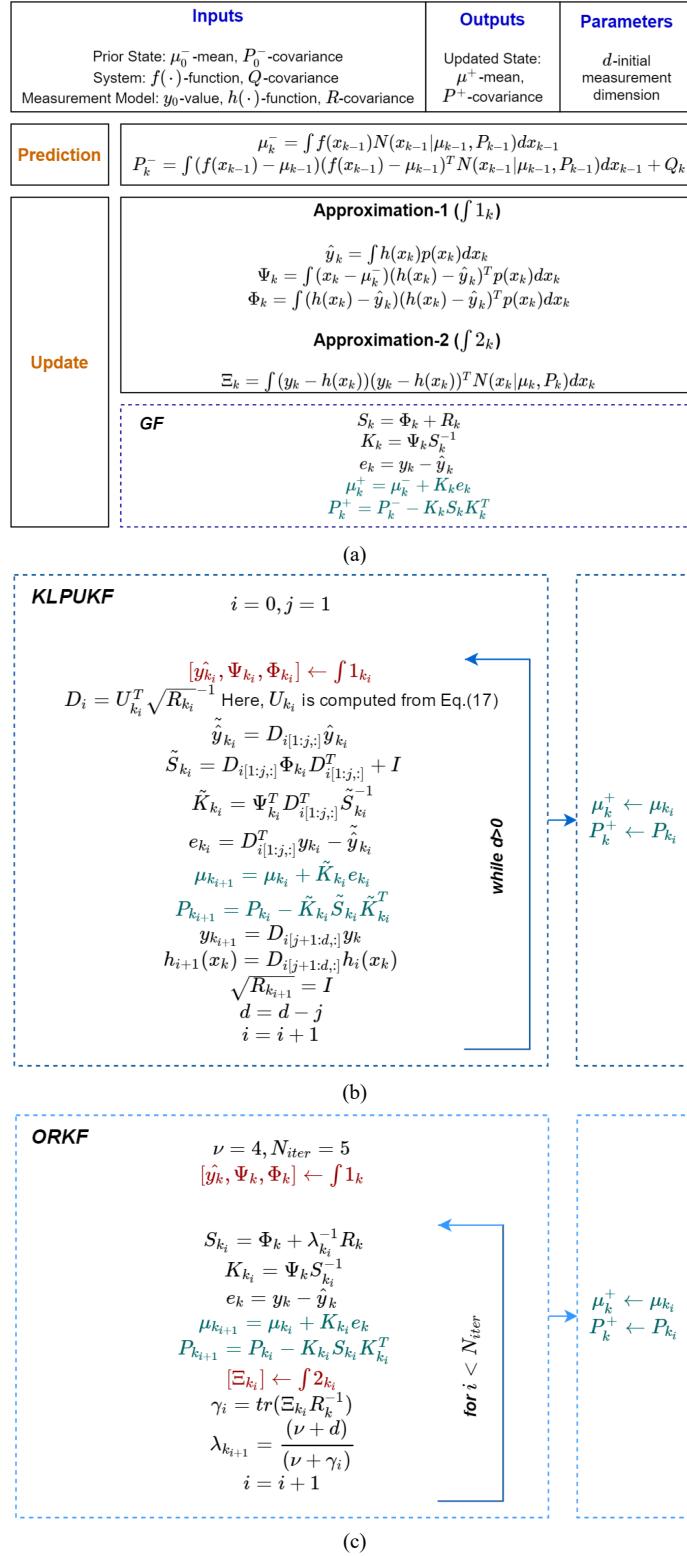


Fig. 1. Algorithm schemes of GF (a), KLPUKF (b), ORKF (c).

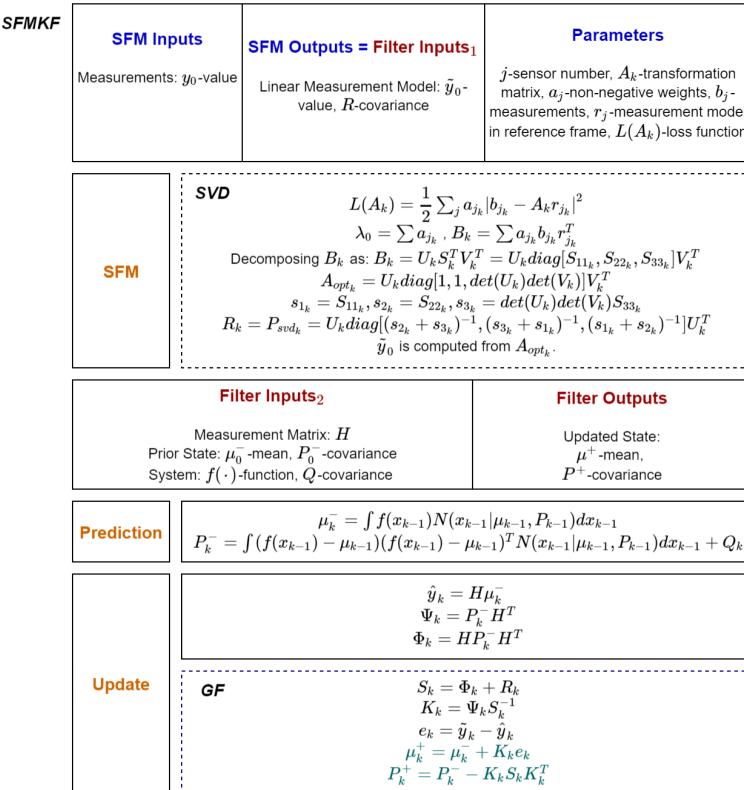


Fig. 2 SFMKF algorithm scheme.

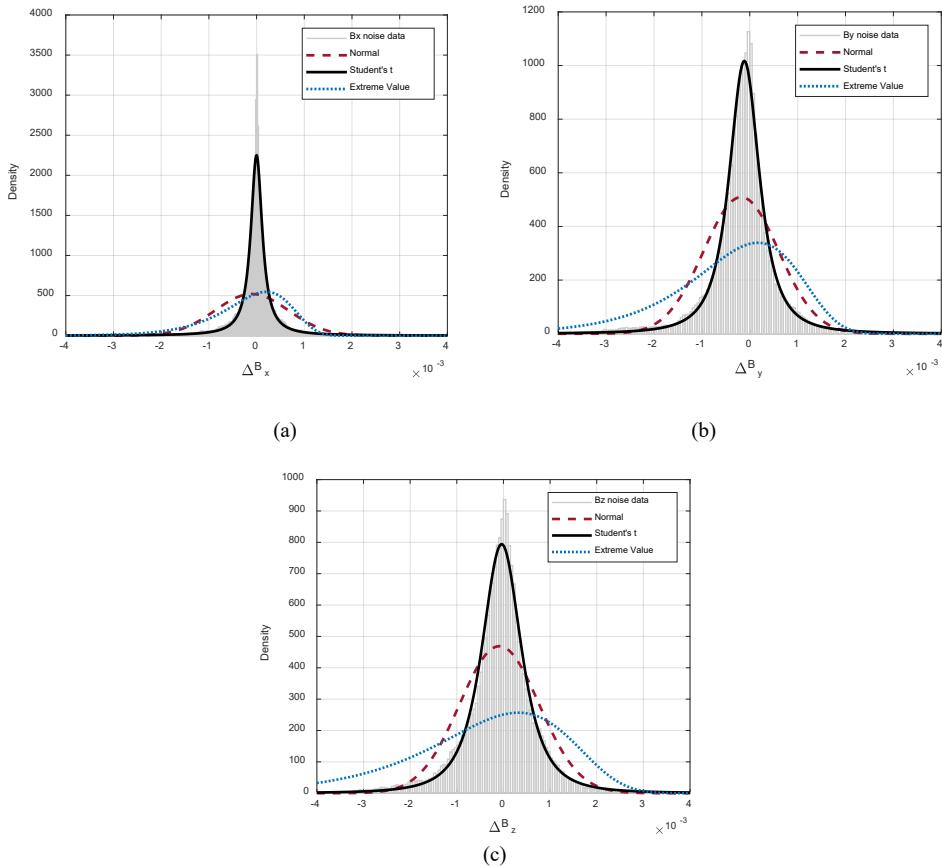


Fig. 3. Distribution fitting for magnetometer measurement noises on x, y, and z axes.

TABLE I AKAIKE'S INFORMATION CRITERION (AIC) ANALYSIS OF DISTRIBUTIONS FOR MAGNETOMETER NOISE.

Model	AIC		
	x	y	z
Normal	3.999	2.694	6.693
t	3.998	2.690	6.531
Extreme value	4.001	2.700	6.721

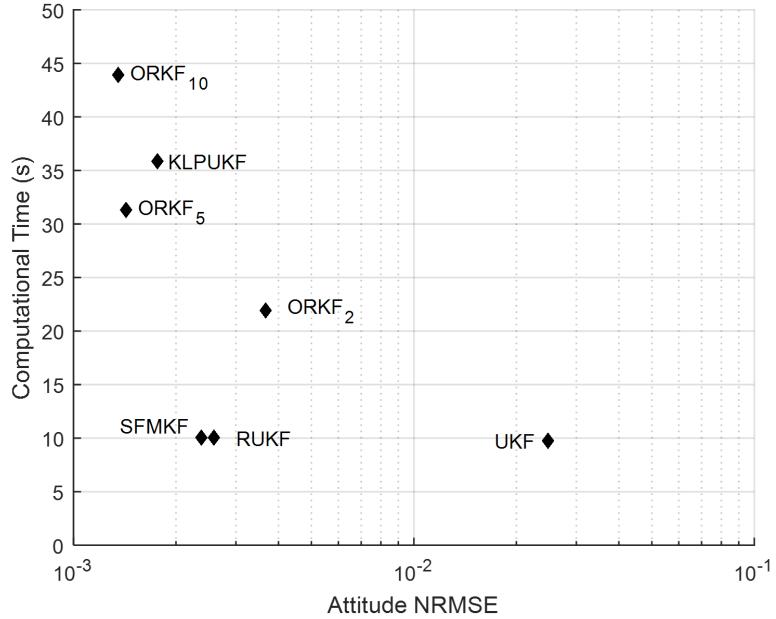


Fig. 4. Work vs Accuracy for 5 filters.

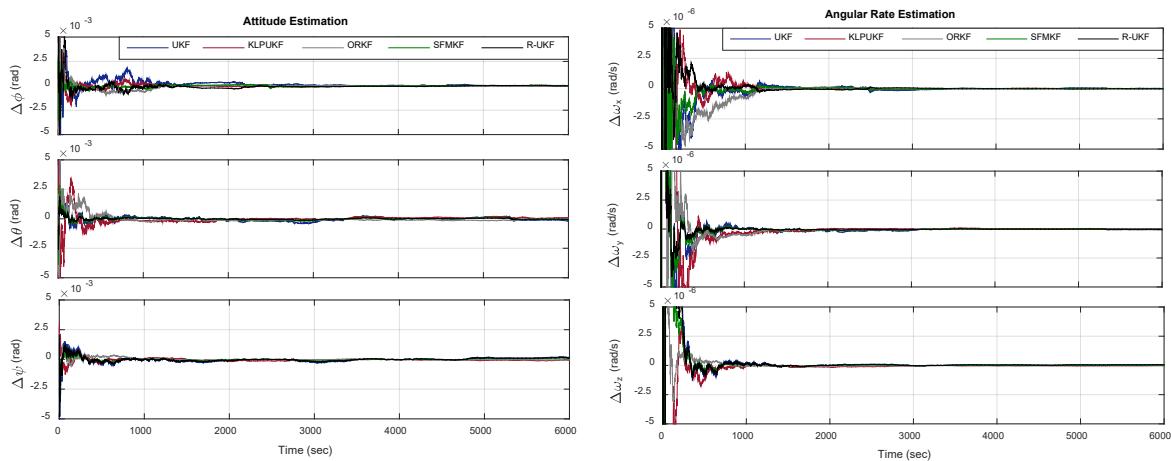


Fig. 5. Attitude (left) and angular velocity (right) estimation errors for normal operation mode, single simulation.

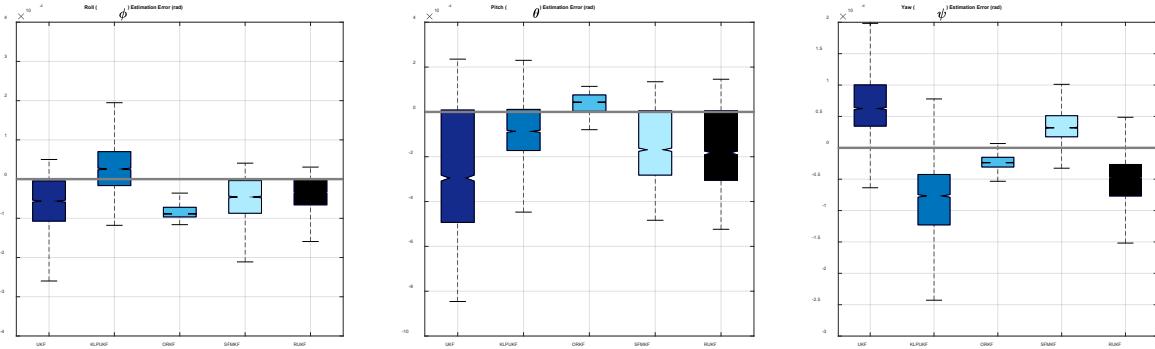
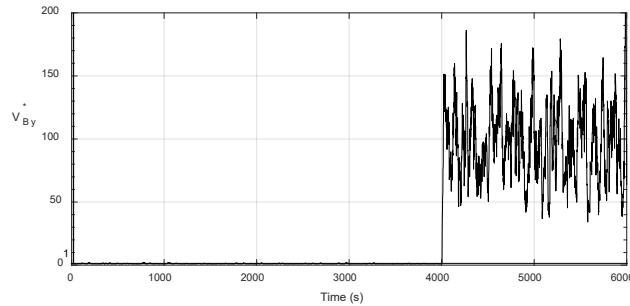


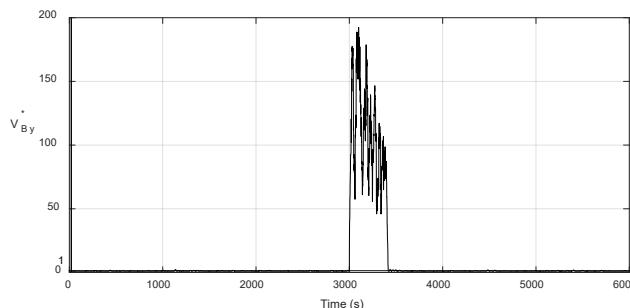
Fig. 6. Box plots of estimation errors during $t \geq 1500$ s, averaged over 100 simulations.

TABLE II RMSE OF ESTIMATIONS DURING $t \geq 1500$ s, AVERAGED OVER 100 SIMULATIONS.

RMSE	Roll ($\times 10^{-3}$ rad)	Pitch ($\times 10^{-3}$ rad)	Yaw ($\times 10^{-3}$ rad)	ω_x ($\times 10^{-6}$ rad/s)	ω_y ($\times 10^{-6}$ rad/s)	ω_z ($\times 10^{-6}$ rad/s)
UKF	0.659	0.664	0.065	0.2201	0.3562	0.9871
KLPUKF	0.171	0.183	0.078	0.0863	0.0591	0.0879
ORKF	0.075	0.133	0.041	0.0523	0.0163	0.0141
SFMKF	0.054	0.190	0.064	0.0883	0.2100	0.0419
RUKF	0.052	0.187	0.042	0.0801	0.2107	0.0378



(a)



(b)

Fig. 7. RUKF scaling factor for long-term (a) and short-term (b) magnetometer fault.

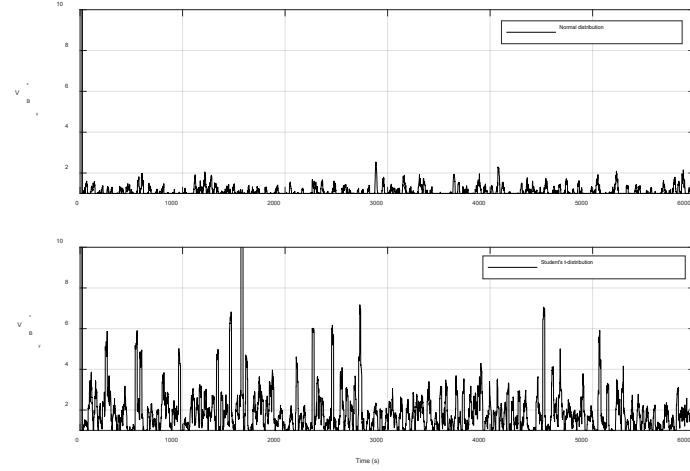
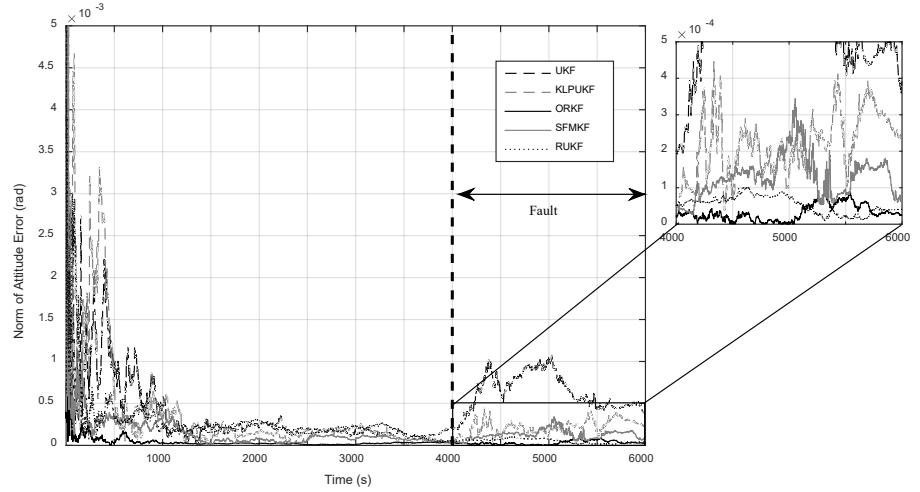
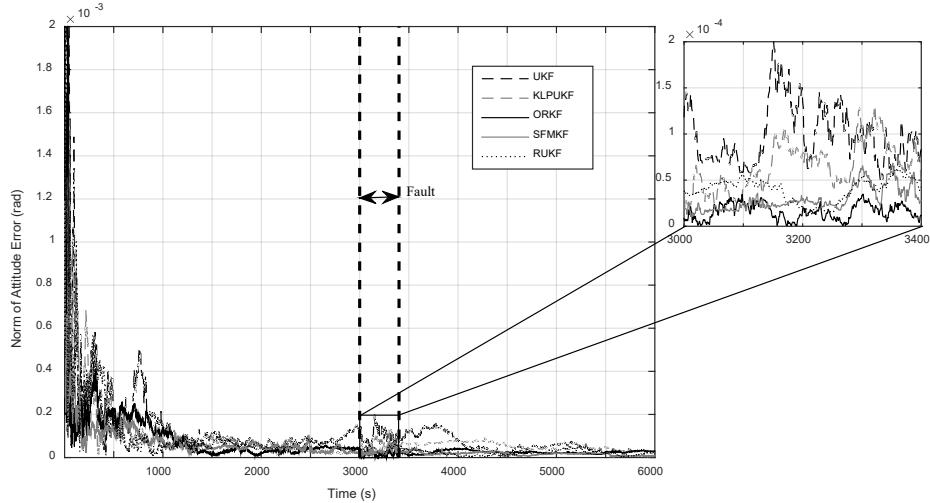


Fig. 8. RUKF scaling factor for normal and t-distributions without any magnetometer faults.



(a)



(b)

Fig. 9. Attitude errors for long-term (a) and short-term (b) magnetometer faults, single simulation.

TABLE III FILTER RMSE DURING FAULTY PERIOD, AVERAGED OVER 100 SIMULATIONS.

RMSE	Long-term			Short-term		
	Roll ($\times 10^{-3}$ rad)	Pitch ($\times 10^{-3}$ rad)	Yaw ($\times 10^{-3}$ rad)	Roll ($\times 10^{-3}$ rad)	Pitch ($\times 10^{-3}$ rad)	Yaw ($\times 10^{-3}$ rad)
UKF	0.991	1.843	2.001	0.705	0.771	0.816
KLPUKF	0.213	1.102	0.201	0.198	0.683	0.155
ORKF	0.146	0.264	0.091	0.076	0.120	0.047
SFMKF	0.170	0.375	0.102	0.091	0.168	0.071
RUKF	0.159	0.275	0.114	0.084	0.141	0.059