

Accuracy improvement by implementing sequential measurement update in robust Kalman filter

Guobin Chang^{1,2} · Yunjia Wang¹ · Qianxin Wang^{1,2}

Received: 17 June 2015 / Accepted: 6 July 2015 / Published online: 17 July 2015
© Akadémiai Kiadó 2015

Abstract A robust Kalman filter based on Chi square test with sequential measurement update is proposed. This approach can not only handle outliers in part or even individual measurement channel, but can also further improve the accuracy especially when a novel ordering strategy in processing the measurement elements is adopted. The accuracy improvement can be attributed to the higher statistical efficiency, i.e., an increased probability of correctly resisting the outlying measurement elements and retaining the good ones. The accuracy improvement of the proposed method is illustrated by a simulating example.

Keywords Robust Kalman filter · Chi square test · Sequential measurement update · Statistical efficiency

1 Introduction

Kalman filter (KF) is widely used in processing kinematic geodetic measurements (Bogatin and Kogoj 2008). For linear systems with Gaussian noises, KF is optimal in almost every conceivable sense (Simon 2006). Gaussianity is often adopted due to its tractability and its good asymptotic performances, but exact Gaussianity seems to be an idealistic assumption in many practical cases. In fact Gaussianity is only approximate for cases with small sample data, especially when outliers/biases are not negligible. As a 2-norm-of-error minimizer, KF is rather sensible to deviations from the assumed Gaussianity which is well known as the lack-of-robustness or lack-of-reliability. While robustness/reliability can be

✉ Guobin Chang
guobinchang@hotmail.com

¹ School of Environmental Science and Spatial Informatics, China University of Mining and Technology, Xuzhou, China

² Xi'an Research Institute of Surveying and Mapping, Xi'an, China

rather general concepts, only the robustness against uncertainties in noise probability distribution is considered, i.e., the distributional robustness to be more specific (Huber and Ronchetti 2009).

In the geodetic literature, there are two kinds of approaches to address the non-Gaussianity and/or outliers/biases (Lehmann 2013a), one is the test-based outlier detection methods and the other is robust statistics based method.

For the test-based outlier detection method, in his pioneer work, (Baarda 1968) proposed to use hypothesis test to detect outliers in geodetic measurement. While the a priori variance is used in (Baarda 1968), the test using estimated variance is introduced in (Pope 1976). Trying to extend the theory of Baarda (1968) and Pope (1976), Teunissen developed the recursive detection, identification, and adaptation (DIA) theory (Teunissen 1990a, b; Teunissen and Salzmann 1989). In DIA, the detection process aims to decide by using a overall test whether some kind of bias or outlier is present; the identification process serves to decide in which channel of the measurement and/or process vector and at which epoch the bias/outlier occurred; in the adaptation process, the detected bias/outlier are corrected or discarded. The DIA methods have found widely applications in the geodetic community, e.g., to detect GNSS pseudorange outliers, phase cycle slips or ionospheric disturbances (Teunissen 1998; Teunissen and De Bakker 2013), station coordinate discontinuities (Perfetti 2006), etc. Also, by inverting the power function of the test, a concept called minimal detectable bias/outlier, some kind of reliability measure, can be derived which can be used to evaluate the strength of the measurement model (De Jong 2000; Koch 2015; Teunissen 1998), and hence further to conduct the design of the measurement model (Salzmann 1991). Some possible difficulties in using DIA include the following. First, appropriate alternative hypotheses should be carefully chosen, this may be the most non-trivial task in DIA (Teunissen 1990a). The alternative hypotheses construction in the identification process depends heavily on the specific problem to be solved and determines directly how to adapt the detected unusual measurement in the adaptation process (Lehmann 2013b). Second, the critical values of the test statistics in the detection and/or identification process are often hard to determine because the complex distribution pattern of the test statistics, sometimes, some kind of numerical method, e.g., the Monte Carlo method, can be used instead of the analytical methods (Lehmann 2012). Third, the application of DIA in the multiple-outlier case needs further investigation.

For the robust statistics based method, the well-developed discipline called robust statistics, aiming to robustify the conventional statistical inference methods such as estimation (Huber and Ronchetti 2009), has been successfully used in geodesy (Guo 2013; Hekimoğlu et al. 2011; Třasák and Štroner 2014). The starting point, which make it different from the test-based method, is to only make the method insensitive to outliers or other statistical uncertainties. In other words, it does not try to detect/identify/correct the uncertainties but only to resist them. It seems also natural to borrow concepts and methods from robust statistics to analyze and improve the robustness of KF. It is indeed so, e.g., the Bayesian estimator, of which KF can be seen a special case, was robustified using the celebrated M-estimator in (Yang 1991). Robust KF for rank-deficient measurement models was studied in (Koch and Yang 1998) focusing on getting the initial estimates at the start of the filtering. Rank-deficient model together with process model uncertainties was addressed recently in (Chang and Liu 2015) focusing on getting more suitable initial estimate in iteratively solving the M-estimation problem, also in (Chang and Liu 2015), the influence function of KF is introduced and derived to evaluate the robustness of a KF-based approach. An adaptively robust KF is proposed in (Yang et al. 2001) to address

uncertainties in both process and measurement models. M-estimator based robust KF is also developed in the framework of nonlinear KF, e.g., the unscented KF (Karlgaard 2015).

Note that there are some kind of overlapping of these two categories. First of all, both aim to robustify the methods, though through different approaches. In the robust statistics method, e.g., the M-estimator, some kind of detection and adaptation (down weighting) can be safely considered existing (Lehmann 2013a). Some kinds of combinations of the two are also possible, e.g., in (Lehmann 2013a), it is stated that “Robust estimation procedures can also be considered as preparatory tools for improved outlier testing”.

A robust KF using Chi square test to detect outliers in the measurement is studied in (Chang 2014), in this approach, the Mahalanobis distance of the measurement under assumed Gaussian distribution is constructed as the test statistic. This approach bears some features of both the above two methods. Outliers, raised due to statistical uncertainties, is detected using a Chi square test, this is the same to the detection process of the DIA in its local test form. However, for the detected outlier, the corresponding measurement is down weighted to make the estimate insensitive to it, this follows the lines of robust statistics, of course one can also consider it is following one kind of the adaptation of DIA. In (Chang 2014) Only one total Mahalanobis distance of all measurement elements is calculated and only one scaling factor is introduced to inflate the overall covariance matrix of the innovation vector, so in its original form, the approach cannot efficiently address uncertainties in only part of the measurement channels (Chang 2014). It was mentioned in (Chang 2014) that this problem can be fixed through implementing sequential measurement update, i.e., processing the vectorial measurement element by element. This idea is detailed and further explored in the current work. In addition to addressing uncertainties in part or even individual measurement channel, there are some by-products of this idea, e.g., superior numerical stability can be expected because no matrix inversion is involved. More importantly, accuracy can be further improved especially through elaborately choosing the order of processing the elements of the measurement vector. We attribute this improvement to the higher statistical efficiency gained. More specifically, after part of the elements are processed, better estimation (better than the prediction) can be obtained which serves as better reference to detect outliers in processing the remaining part of the elements. Better reference will increase the probability of correctly resisting the outlying ones and retaining the good ones in the measurement elements, which means a higher statistical efficiency.

The remaining part of the paper is organized as follows. The method is presented in Sect. 2. Accuracy improvement is illustrated with a simulating example in Sect. 3. Some concluding remarks is given in Sect. 4.

2 Method

In the first subsection, after presenting the basic formulae of the KF, the approach of detecting and resisting outliers in the measurement is introduced. In the second subsection, sequentially implementing the measurement update with the previously introduced outlier handling method is derived emphasizing a novel ordering strategy in processing the elements of the measurement vector.

2.1 Kalman filter and the resistance of outliers

The problem studied is represented as the following discrete-time state space model,

$$\mathbf{x}_k = \mathbf{F}\mathbf{x}_{k-1} + \mathbf{w}_{k-1} \tag{1}$$

$$\mathbf{y}_k = \mathbf{H}\mathbf{x}_k + \mathbf{v}_k \tag{2}$$

where \mathbf{x}_k and \mathbf{y}_k are n - and m -dimensional state and measurement vectors at the k th epoch, \mathbf{F} and \mathbf{H} are transition and design matrices with appropriate dimensions, \mathbf{w}_k and \mathbf{v}_k are process and measurement noises which are assumed zero-mean Gaussianly distributed with nominal covariance matrix \mathbf{Q} and \mathbf{R} respectively. Note that in this study the real distribution of \mathbf{v}_k can deviate from this assumption. Assume the initial estimate at 0 epoch is $\hat{\mathbf{x}}_{0|0}$ with associate covariance estimate being $\hat{\mathbf{P}}_{\mathbf{x}_{0|0}, \mathbf{x}_{0|0}}$. At any epoch, say k , we have the following KF formulae.

$$\hat{\mathbf{x}}_{k|k-1} = \mathbf{F}\hat{\mathbf{x}}_{k-1|k-1} \tag{3}$$

$$\hat{\mathbf{P}}_{\mathbf{x}_{k|k-1}, \mathbf{x}_{k|k-1}} = \mathbf{F}\hat{\mathbf{P}}_{\mathbf{x}_{k-1|k-1}, \mathbf{x}_{k-1|k-1}}\mathbf{F}^T + \mathbf{Q} \tag{4}$$

$$\hat{\mathbf{y}}_{k|k-1} = \mathbf{H}\hat{\mathbf{x}}_{k|k-1} \tag{5}$$

$$\hat{\mathbf{P}}_{\mathbf{y}_{k|k-1}, \mathbf{y}_{k|k-1}} = \mathbf{H}\hat{\mathbf{P}}_{\mathbf{x}_{k|k-1}, \mathbf{x}_{k|k-1}}\mathbf{H}^T + \mathbf{R} \tag{6}$$

$$\mathbf{K}_k = \hat{\mathbf{P}}_{\mathbf{x}_{k|k-1}, \mathbf{x}_{k|k-1}}\mathbf{H}^T \left(\hat{\mathbf{P}}_{\mathbf{y}_{k|k-1}, \mathbf{y}_{k|k-1}} \right)^{-1} \tag{7}$$

$$\hat{\mathbf{x}}_{k|k} = \hat{\mathbf{x}}_{k|k-1} + \mathbf{K}_k \left(\tilde{\mathbf{y}}_k - \hat{\mathbf{y}}_{k|k-1} \right) \tag{8}$$

$$\hat{\mathbf{P}}_{\mathbf{x}_{k|k}, \mathbf{x}_{k|k}} = \hat{\mathbf{P}}_{\mathbf{x}_{k|k-1}, \mathbf{x}_{k|k-1}} - \mathbf{K}_k \hat{\mathbf{P}}_{\mathbf{y}_{k|k-1}, \mathbf{y}_{k|k-1}} \mathbf{K}_k^T \tag{9}$$

where $\hat{\mathbf{x}}_{i|j}$ represents an estimate of the state vector at the i th epoch using measurements up to the j th epoch, specifically, $\hat{\mathbf{x}}_{k|k-1}$ and $\hat{\mathbf{x}}_{k|k}$ are also called the a priori and the a posteriori estimates; $\mathbf{P}_{a,b}$ represents the (cross) covariance matrix between \mathbf{a} and \mathbf{b} ; $\tilde{\mathbf{y}}_k$, a non-random constant, is the actual measurement, or a realization of \mathbf{y}_k ; \mathbf{K}_k is gain matrix which combines $\hat{\mathbf{x}}_{k|k-1}$ and $\tilde{\mathbf{y}}_k - \hat{\mathbf{y}}_{k|k-1}$ linearly to get $\hat{\mathbf{x}}_{k|k}$. Note that $\mathbf{e}_k = \tilde{\mathbf{y}}_k - \hat{\mathbf{y}}_{k|k-1}$ is often called the innovation vector (Kailath et al. 2000) whose covariance is equal to $\hat{\mathbf{P}}_{\mathbf{y}_{k|k-1}, \mathbf{y}_{k|k-1}}$ because $\tilde{\mathbf{y}}_k$ is non-random.

Under the Gaussian assumption, \mathbf{y}_k should be Gaussian with mean $\hat{\mathbf{y}}_{k|k-1}$ and covariance $\hat{\mathbf{P}}_{\mathbf{y}_{k|k-1}, \mathbf{y}_{k|k-1}}$, and the squared Mahalanobis distance of \mathbf{y}_k should be Chi square distributed with m freedoms, i.e.,

$$\gamma(\mathbf{y}_k) = d^2 = \left(\mathbf{y}_k - \hat{\mathbf{y}}_{k|k-1} \right)^T \left(\hat{\mathbf{P}}_{\mathbf{y}_{k|k-1}, \mathbf{y}_{k|k-1}} \right)^{-1} \left(\mathbf{y}_k - \hat{\mathbf{y}}_{k|k-1} \right) \sim \chi_m^2 \tag{10}$$

So we can do a Chi square test to judge whether an actual measurement is a realization of \mathbf{y}_k under the Gaussian assumption. Let the null hypothesis be that \mathbf{v}_k is Gaussianly distributed. For a given significance level, say $1 - \alpha$, and the corresponding upper α -quantile $\chi_{m,\alpha}^2$, we have

$$\Pr \left[\gamma(\mathbf{y}_k) > \chi_{m,\alpha}^2 \right] < \alpha \tag{11}$$

if the null hypothesis holds, where $\text{Pr}[\cdot]$ denotes the probability of an event. Given α a rather small value, if $\gamma(\tilde{\mathbf{y}}_k) > \chi_{m,\alpha}^2$, we can say with a rather high probability, i.e., $1 - \alpha$, that the null hypothesis should be rejected, and in this case, $\tilde{\mathbf{y}}_k$ is deemed to be outlier. This shows how the outliers are detected. For a detected outlier, its contribution in the measurement update should be decreased, and this is achieved by inflating the covariance of the innovation vector. The inflating factor can be calculated as

$$\kappa = \frac{\gamma(\tilde{\mathbf{y}}_k)}{\chi_{m,\alpha}^2} \quad (12)$$

This shows how the detected outlier is resisted.

As in Eq. (10), only one statistic is calculated to judge whether the overall measurement vector is an outlier, and as in Eq. (12), only one scaling factor is introduced to inflate the whole covariance matrix. If only part or even a single element of the measurement is outlying, the overall measurement vector may be deemed as being outlying. This means that some good measurement elements may be mistaken as outliers. Also an outlying element may fail to be detected because it may be masked by other good elements. This means that some outlying measurement elements may be mistaken as good ones. To summarize, either mistaking good measurement elements as outlying ones or mistaking outlying ones as good ones will result in loss of statistical efficiency and hence in lower accuracy of the estimates.

2.2 Implementing sequential measurement update in robust Kalman filter

As mentioned in (Chang 2014), outliers in individual measurement channel can be efficiently addressed by implementing sequential measurement update of the KF, i.e., only a single measurement element is processed once in a measurement update and hence there will be m measurement updates in one recursion at any epoch. In every measurement update, the previous outlier detection and resistance will be performed.

If the measurement elements are cross correlated, we first de-correlate them through the Cholesky decomposition. One reviewer insisted on putting doubts on this decorrelation approach, as it would spread outliers over different observations. Unfortunately, the authors have no sounder solutions for now and will work in depth on one this issue in the future. Let

$$\mathbf{R} = \mathbf{L}\mathbf{L}^T \quad (13)$$

From Eq. (2), we have

$$\mathbf{L}^{-1}\mathbf{y}_k = \mathbf{L}^{-1}\mathbf{H}\mathbf{x}_k + \mathbf{L}^{-1}\mathbf{v}_k \quad (14)$$

Let $\bar{\mathbf{y}}_k = \mathbf{L}^{-1}\mathbf{y}_k$, $\bar{\mathbf{H}} = \mathbf{L}^{-1}\mathbf{H}$, and $\bar{\mathbf{v}}_k = \mathbf{L}^{-1}\mathbf{v}_k$, we have de-correlated measurement equation

$$\bar{\mathbf{y}}_k = \bar{\mathbf{H}}\mathbf{x}_k + \bar{\mathbf{v}}_k \quad (15)$$

$$\text{cov}[\bar{\mathbf{v}}_k] = \mathbf{L}^{-1}\mathbf{R}\mathbf{L}^{-T} = \mathbf{L}^{-1}\mathbf{L}\mathbf{L}^T\mathbf{L}^{-T} = \mathbf{I} \quad (16)$$

Note that for this measurement equation, the actual measurement should be $\tilde{\bar{\mathbf{y}}}_k = \mathbf{L}^{-1}\tilde{\mathbf{y}}_k$ accordingly.

Let

$$\bar{\mathbf{H}} = \begin{bmatrix} \bar{h}_1^T & \bar{h}_2^T & \cdots & \bar{h}_m^T \end{bmatrix} \quad (17)$$

In the j th measurement update at the k th epoch, we have

$$\hat{\bar{y}}_{k,j} = \bar{h}_j \hat{x}_{k,j-1} \tag{18}$$

$$\sigma_{\bar{y}_{k,j}}^2 = \bar{h}_j P_{\hat{x}_{k,j-1}, \hat{x}_{k,j-1}} \bar{h}_j^T + 1 \tag{19}$$

where $\hat{x}_{k,j-1}$ is the estimate after processing the $(j - 1)$ th measurement element and $\hat{x}_{k,0} = \hat{x}_{k|k-1}$, $\hat{\bar{y}}_{k,j}$ is the prediction of the j th element of \bar{y}_k . One-freedom Chi square test, is carried out to judge whether $\tilde{\bar{y}}_{k,j}$ is an outlier. The statistic is calculated as

$$\gamma(\tilde{\bar{y}}_{k,j}) = \frac{(\tilde{\bar{y}}_{k,j} - \hat{\bar{y}}_{k,j})^2}{\sigma_{\bar{y}_{k,j}}^2} \tag{20}$$

For a given significance level, say also $1 - \alpha$, and the corresponding upper α -quantile $\chi_{1,\alpha}^2$, we check if

$$\gamma(\tilde{\bar{y}}_{k,j}) > \chi_{1,\alpha}^2 \tag{21}$$

If Eq. (21) holds, $\tilde{\bar{y}}_{k,j}$ is deemed as an outlier, then the following scaling factor is calculated to inflate the variance in Eq. (19), i.e.,

$$\kappa = \frac{\gamma(\tilde{\bar{y}}_{k,j})}{\chi_{1,\alpha}^2} \tag{22}$$

$$\kappa \sigma_{\bar{y}_{k,j}}^2 \rightarrow \sigma_{\bar{y}_{k,j}}^2 \tag{23}$$

And then update the estimate as

$$K = \frac{P_{\hat{x}_{k,j-1}, \hat{x}_{k,j-1}} \bar{h}_j^T}{\sigma_{\bar{y}_{k,j}}^2} \tag{24}$$

$$\hat{x}_{k,j} = \hat{x}_{k,j-1} + K(\tilde{\bar{y}}_{k,j} - \hat{\bar{y}}_{k,j}) \tag{25}$$

$$P_{\hat{x}_{k,j}, \hat{x}_{k,j}} = P_{\hat{x}_{k,j-1}, \hat{x}_{k,j-1}} - \sigma_{\bar{y}_{k,j}}^2 K K^T \tag{26}$$

It is well known that $\hat{x}_{k,j}$ is a more accurate estimate than $\hat{x}_{k,j-1}$, because $P_{\hat{x}_{k,j}, \hat{x}_{k,j}} \leq P_{\hat{x}_{k,j-1}, \hat{x}_{k,j-1}}$, i.e., $P_{\hat{x}_{k,j-1}, \hat{x}_{k,j-1}} - P_{\hat{x}_{k,j}, \hat{x}_{k,j}}$ is positive semi-definite, which is explicitly shown in Eq. (26). So in checking the $(j + 1)$ th measurement elements, the reference information used, i.e., $\hat{x}_{k,j}$, is better than $\hat{x}_{k,j-1}$, of course even better than $\hat{x}_{k|k-1}$. Better reference information implies an increased probability of detecting outliers, so it is more probable to correctly resist the outlying elements and to retain the good ones.

In implementing the above robust sequential measurement update, it should be apparent that the more reliable a measurement element is, the more accurate the estimate will become. So the more reliable elements should be firstly processed in order to get a even better reference information to check the remaining dubious elements. The Mahalanobis distances or their squares of individual elements can represent to some extent the relative qualities of these elements, i.e., the smaller one's Mahalanobis distance is, the more reliable the elements should be. So for given reference information, Mahalanobis distances

of all the remaining elements are calculated, and the element with the smallest Mahalanobis distance is processed in the next measurement update. Of course this sorting process will increase the computation, but the accuracy improvement may reward in some cases.

The algorithm is depicted in Table 1 in the form of pseudo code.

3 An illustrating example

Assume an object is moving forward without side slip in the horizon, the forward distance and velocity are of interest and the north and east position is measured.

A constant velocity model is assumed, so we have

$$\begin{bmatrix} \dot{p} \\ \dot{v} \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} p \\ v \end{bmatrix} + \begin{bmatrix} 0 \\ a \end{bmatrix} \quad (27)$$

where, p and v are the forward position and velocity which should be estimated, a is the forward acceleration which is assumed white Gaussian noise with variance σ_a^2 . Given the integration interval τ , Eq. (27) is discretized to get the process equation

$$\begin{bmatrix} p_k \\ v_k \end{bmatrix} = \begin{bmatrix} 1 & \tau \\ 0 & 1 \end{bmatrix} \begin{bmatrix} p_{k-1} \\ v_{k-1} \end{bmatrix} + \mathbf{w}_{k-1} \quad (28)$$

with

$$\mathbf{Q} = \text{cov}[\mathbf{w}_{k-1}] = \sigma_a^2 \begin{bmatrix} \frac{\tau^3}{3} & \frac{\tau^2}{2} \\ \frac{\tau^2}{2} & \tau \end{bmatrix} \quad (29)$$

The measurement equation is

$$\begin{bmatrix} p_{N,k} \\ p_{E,k} \end{bmatrix} = \begin{bmatrix} \cos \theta & 0 \\ \sin \theta & 0 \end{bmatrix} \begin{bmatrix} p_k \\ v_k \end{bmatrix} + \begin{bmatrix} \varepsilon_k \\ \zeta_k \end{bmatrix} \quad (30)$$

where θ is the heading angle. The measurement noises ε_k and ζ_k are assumed to be distributed as gross error model, which is proposed in (Huber 1964) to formulate the full kind of the “neighborhood” of an assumed parametric model, the probability density function is

$$f = (1 - \mu)f_0 + \mu f_c \quad (31)$$

where f , f_0 , and f_c represent the real, the nominal, and the contaminating distribution respectively, μ is called the contaminating ratio. When $\mu = 0$, the nominal distribution represent the real distribution exactly. In this study both f_0 and f_c are assumed to be zero-mean Gaussian but with different variance, in this case, the gross error model in Eq. (31) is also called a Gauss mix model.

The following three cases are studied.

Case 1 NoUn: uncertainty exists in neither of the two channels, i.e., the nominal distributions of both ε_k and ζ_k represent their real distributions;

Table 1 Outlier detection and resistance in sequential measurement update

Start filter

Given F, H, Q, R, \tilde{y}_k , and α ; given $\hat{x}_{0|0}$ and $P_{\hat{x}_{0|0}, \hat{x}_{0|0}}$.

1. Prediction,

$$\hat{x}_{k|k-1} = F\hat{x}_{k-1|k-1}, P_{\hat{x}_{k|k-1}, \hat{x}_{k|k-1}} = FP_{\hat{x}_{k-1|k-1}, \hat{x}_{k-1|k-1}}F^T + Q$$

Let $\hat{x}_k = \hat{x}_{k|k-1}$, $P_{\hat{x}_k, \hat{x}_k} = P_{\hat{x}_{k|k-1}, \hat{x}_{k|k-1}}$, and $j=1$;

2. De-correlation,

$$R = LL^T, \bar{y}_k = L^{-1}y_k, \text{ and } \bar{H} = L^{-1}H = [\bar{h}_1^T \ \bar{h}_2^T \ \dots \ \bar{h}_m^T]^T;$$

3. Sequential update,

3.1 Choosing the most reliable one among the $(m-j+1)$ unprocessed elements,

$$\hat{y}_{k,i} = \bar{h}_i \hat{x}_{k,j-1}, \sigma_{\hat{y}_{k,i}}^2 = \bar{h}_i P_{\hat{x}_{k,j-1}, \hat{x}_{k,j-1}} \bar{h}_i^T + 1, \gamma(\tilde{y}_{k,i}) = \frac{(\tilde{y}_{k,i} - \hat{y}_{k,i})^2}{\sigma_{\hat{y}_{k,i}}^2}, \text{ for } i=i_1, i_2, \dots, i_{m-j+1}.$$

The element with smallest is selected, say the i 'th $\tilde{y}_{k,i}$.

3.2 Outlier detection and covariance inflation,

Calculating the scaling factor,

$$\kappa = \begin{cases} \frac{\gamma(\tilde{y}_{k,i'})}{\chi_{1,\alpha}^2}, & \text{if } \gamma(\tilde{y}_{k,i'}) > \chi_{1,\alpha}^2 \\ 1, & \text{otherwise} \end{cases}$$

$$\sigma_{\hat{y}_{k,i'}}^2 = \kappa \sigma_{\hat{y}_{k,i'}}^2$$

3.3 State estimate update,

$$K = \frac{P_{\hat{x}_k, \hat{x}_k} \bar{h}_j^T}{\sigma_{\hat{y}_{k,i'}}^2}, \hat{x}_k = \hat{x}_k + K(\tilde{y}_{k,i'} - \hat{y}_{k,i'}), P_{\hat{x}_k, \hat{x}_k} = P_{\hat{x}_k, \hat{x}_k} - \sigma_{\hat{y}_{k,i'}}^2 K K^T$$

3.4 Recursion

If $j < m$,

$j=j+1$, and go to 3.1,

Else,

$$\hat{x}_{k|k} = \hat{x}_k, P_{\hat{x}_{k|k}, \hat{x}_{k|k}} = P_{\hat{x}_k, \hat{x}_k}, k=k+1, \text{ and go to 1;}$$

End filter.

Case 2 UnOn: uncertainty exists in only one channel, say the north channel;

Case 3 UnBo: uncertainties exist in both channels.

Let $f_0 \sim N(0,1)$ and $f_c \sim N(0,10)$, if uncertainty exists, let $\mu = 0.1$, otherwise $\mu = 0$. Three approaches are checked in all three cases,

Approach 1 KF, the standard KF;

Approach 2 RKF, the robust KF proposed in (Chang 2014), also introduced in Sect. 2.1. In this approach let $\alpha = 0.05$, so $\chi_{2, 0.05}^2 = 5.99$.

Approach 3 RKF, the robust KF with sequential measurement update proposed in this work, illustrated in Table 1. In this approach let $\alpha = 0.05$, so $\chi_{1, 0.05}^2 = 3.84$.

Monte Carlo experiments with 10,000 independent runs are conducted. The mean squared (over different Monte Carlo runs) errors of the estimates at any epoch by all three approaches in all three cases are calculated and depicted in the figures. The overall mean squared errors (over both different Monte Carlo runs and different epochs) of the estimates by three approaches in three cases are also calculated and summarized in Table 2.

In Figs. 1 and 2, it is found that all three approaches perform well in the first case. As the KF is optimal in this case, the comparative performance of the two robust approaches validates their statistical efficiency in this case. More specifically, there exists inevitably probability that the good measurement may be mistaken as outliers, but this probability is rather low. The high efficiency is achieved by deliberately selecting a rather high significance level, i.e., $1 - \alpha$.

From Figs. 3 and 4, we see that the performance of the KF degrades significantly which clearly shows its lack of robustness. However, two robust approaches can still provide relatively good estimates in this case. The higher accuracy of the two robust approaches compared to that of the KF is due to the robustness of two. Note that in spite of the robustness, the root mean squared errors of the two robust approaches still increase compared to the first case, this has nothing to do with robustness, rather this is because the equivalent accuracy of the measurement decrease compared to the first case. In the first case, the standard deviation of the measurement noise (north channel) is 1 m, while the equivalent standard deviation (north channel) in the second case is

$$\bar{\sigma} = \sqrt{(1 - \mu)\sigma_{f_0}^2 + \mu\sigma_{f_c}^2} = \sqrt{0.9 \times 1 + 0.1 \times 100} \approx 3.3 > 1 \quad (32)$$

Table 2 Root mean squared errors in position and velocity estimations by three approaches, i.e., Kalman filter (KF), robust Kalman filter (RKF), and robust Kalman filter with sequential measurement update (RKF), in three cases, i.e., with no uncertainty (NoUn), with uncertainties in one channel (UnOn), and with uncertainties in both channels (UnBo)

Error terms	Approaches	NoUn	UnOn	UnBo
Position (m)	KF	0.3903	0.9192	1.2393
	RKF	0.5199	0.5497	0.5723
	RKFs	0.3933	0.4098	0.4300
Velocity (m/s)	KF	1.0809	1.8103	2.3350
	RKF	1.0829	1.1067	1.1407
	RKFs	1.0799	1.0927	1.1173

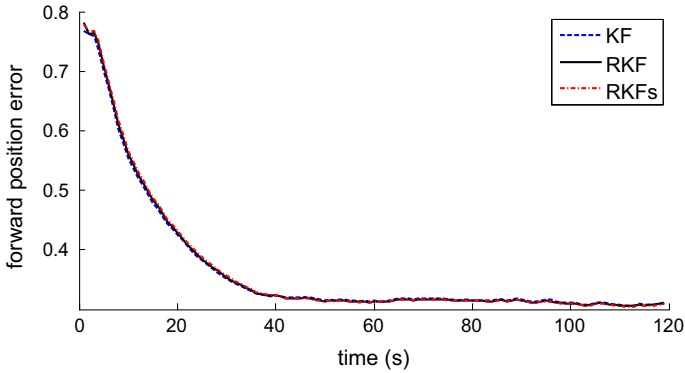


Fig. 1 Mean squared errors in the position estimates in Case 1 (NoUn)

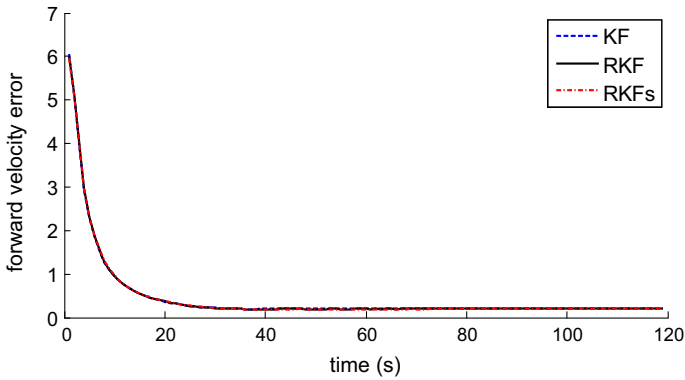


Fig. 2 Mean squared errors in the velocity estimates in Case 1 (NoUn)

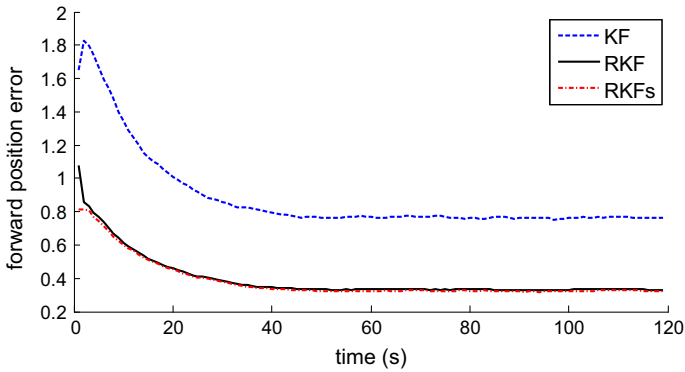


Fig. 3 Mean squared errors in the position estimates in Case 2 (UnOn)

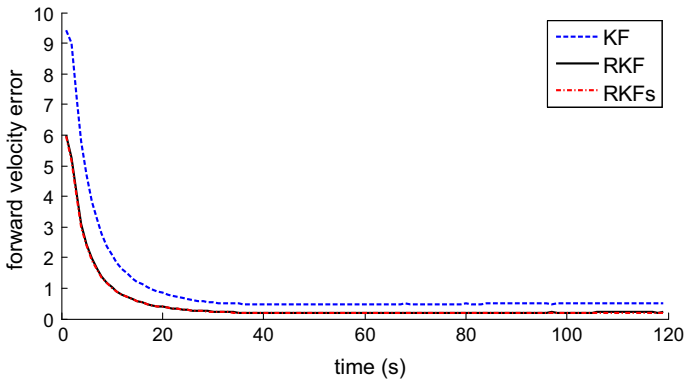


Fig. 4 Mean squared errors in the velocity estimates in Case 2 (UnOn)

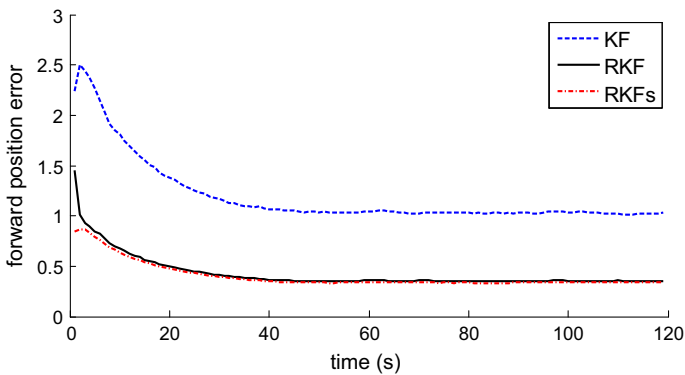


Fig. 5 Mean squared errors in the position estimates in Case 3 (UnBo)

From Figs. 3 and 4, we can also find that the RKF_s approach performs slightly better than the RKF approach. The superiority of RKF_s can be clearly seen in Table 2. As explained previously, this is due to the higher statistical efficiency of RKF_s than RKF.

From Figs. 5 and 6, we observe that again in this case, the two robust approaches outperform the standard KF which is due to the robustness. Again in this case, the performances of the two robust approaches decrease compared to the second case which is caused by the increased standard deviation of measurement noise in east channel. The RKF_s performs slightly better than RKF as is clearly demonstrated in Table 2, again, this is due to the relatively higher statistical efficiency of the former compared to the latter.

4 Concluding remarks

The distributional uncertainty in the measurement noise, or more specifically the non-Gaussianity of the measurement noise's distribution, is addressed through employing a robust KF based on Chi square test to detect outliers and innovation vector covariance inflation to resist the detected outliers. Through implementing the so called sequential

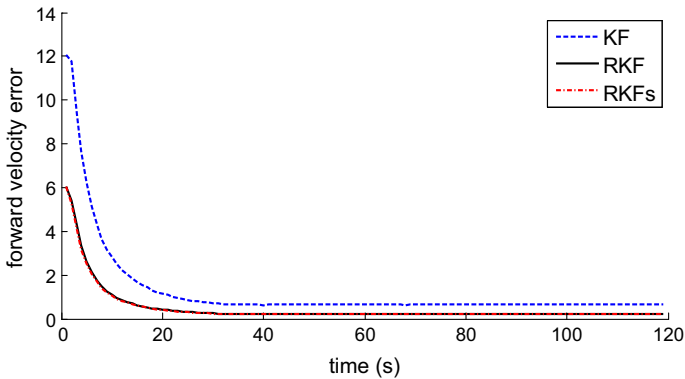


Fig. 6 Mean squared errors in the velocity estimates in Case 3 (UnBo)

measurement update of the robust KF, we achieved manifold merits: the ability to address outliers in part of or even individual measurement channel; higher numerical stability because matrix inverse is no longer needed; and most importantly higher accuracy because of a higher statistical efficiency in detecting outliers. The higher statistical efficiency is brought about by a higher probability of correctly detecting the outlying measurement elements and retaining the good ones.

It must be admitted that there is inevitably computation increase in the proposed method, mainly because hypothesis test should be done for every measurement element and a sorting process should be carried out to select the measurement element for the next processing, but we believe that in some case, the higher accuracy gained may reward the increased computation paid.

Acknowledgments The authors are grateful to two anonymous reviewers whose valuable comments improved the paper significantly. This work was supported by the National Natural Science Foundation of China (No. 41404001).

References

- Baarda W (1968) A testing procedure for use in geodetic networks. *Publ Geod New Ser* 2(5):1–97
- Bogatin S, Kogoj D (2008) Processing kinematic geodetic measurements using Kalman filtering. *Acta Geodaetica et Geophysica Hungarica* 43(1):53–74
- Chang G (2014) Robust Kalman filtering based on Mahalanobis distance as outlier judging criterion. *J Geod* 88(4):391–401
- Chang G, Liu M (2015) M-estimator based robust Kalman filter for systems with process modeling errors and rank deficient measurement models. *Nonlin Dyn* 80(3):1431–1449
- De Jong K (2000) Minimal detectable biases of cross-correlated GPS observations. *GPS Solut* 3(3):12–18
- Guo J (2013) The case-deletion and mean-shift outlier models: equivalence and beyond. *Acta Geodaetica et Geophysica* 48(2):191–197
- Hekimoğlu S, Erdogan B, Erenoglu R, Hosbas R (2011) Increasing the efficacy of the tests for outliers for geodetic networks. *Acta Geodaetica et Geophysica Hungarica* 46(3):291–308
- Huber PJ (1964) Robust estimation of a location parameter. *Ann Math Stat* 35(1):73–101
- Huber PJ, Ronchetti EM (2009) *Robust statistics*. Wiley, New Jersey
- Kailath T, Sayed AH, Hassibi B (2000) *Linear estimation*. Prentice Hall, New Jersey
- Karlgård CD (2015) Nonlinear regression huber-kalman filtering and fixed-interval smoothing. *J Guid Control Dyn* 38(2):322–330

- Koch KR (2015) Minimal detectable outliers as measures of reliability. *J Geod* 89(5):483–490
- Koch KR, Yang Y (1998) Robust Kalman filter for rank deficient observation models. *J Geod* 72(7):436–441
- Lehmann R (2012) Improved critical values for extreme normalized and studentized residuals in Gauss–Markov models. *J Geod* 86:1137–1146
- Lehmann R (2013a) 3σ -Rule for outlier detection from the viewpoint of geodetic adjustment. *J Surv Eng* 139(4):157–165
- Lehmann R (2013b) On the formulation of the alternative hypothesis for geodetic outlier detection. *J Geod* 87:373–386
- Perfetti N (2006) Detection of station coordinate discontinuities within the Italian GPS fiducial network. *J Geod* 80(7):381–396
- Pope A. (1976). The statistics of residuals and the detection of outliers. NOAA Technical Report, NOS 65(NGS 1): 1-133
- Salzmann M (1991) MDB: a design tool for integrated navigation systems. *Bull Géod* 65(2):109–115
- Simon D (2006) Optimal state estimation: Kalman, $H\infty$, and nonlinear approaches. Wiley, New Jersey
- Teunissen PJG. (1990a). An integrity and quality control procedure for use in multi sensor integration. In: Proceedings ION GPS, pp 19–21
- Teunissen PJG (1990b) Quality control in integrated navigation systems. *IEEE Aerosp Electron Syst Mag* 5(7):35–41
- Teunissen PJG (1998) Minimal detectable biases of GPS data. *J Geod* 72(4):236–244
- Teunissen PJG, De Bakker PF (2013) Single-receiver single-channel multi-frequency GNSS integrity: outliers, slips, and ionospheric disturbances. *J Geod* 87(2):161–177
- Teunissen PJG, Salzmann MA (1989) A recursive slippage test for use in state-space filtering. *Manuscripta geodaetica* 14(6):383–390
- Trásák P, Štroner M (2014) Outlier detection efficiency in the high precision geodetic network adjustment. *Acta Geodaetica et Geophysica* 49(2):161–175
- Yang Y (1991) Robust bayesian estimation. *Bull Géod* 65(3):145–150
- Yang Y, He H, Xu G (2001) Adaptively robust filtering for kinematic geodetic positioning. *J Geod* 75(2):109–116