Asteasuain Mariano (Orcid ID: 0000-0003-4599-7310)

Efficient and robust state estimation: Application to a copolymerization process

Jhovany Tupaz[1], Mariano Asteasuain[1,2], Mabel Sánchez[1,2]

[1]Departamento de Ingeniería Química, Universidad Nacional del Sur (UNS), Bahía Blanca, Argentina.
[2]Planta Piloto de Ingeniería Química (PLAPIQUI), Universidad Nacional del Sur (UNS) - CONICET, Bahía Blanca, Argentina.

Correspondence
Mabel Sánchez
Departamento de Ingeniería Química, Universidad Nacional del Sur (UNS), Av. Alem 1253, Bahía Blanca 8000, Argentina.

Planta Piloto de Ingeniería Química (PLAPIQUI), Universidad Nacional del Sur (UNS) - CONICET, Camino La Carrindanga km. 7, Bahía Blanca 8000, Argentina.

E-mail address: msanchez@plapiqui.edu.ar

Abstract

Polymerization processes are highly non-linear systems that require strict control of their dynamic operation to be competitive. The unscented Kalman filter is a filtering strategy that has shown a rewarding performance for non-linear state estimation. Besides, filters based on robust statistics have been proposed to deal with the presence of outliers. However, reported robust filters have employed only the Huber M-estimator as the loss function of the estimation problem. This work presents a new state-estimation procedure based on the unscented transformation and robust statistics concepts. When outliers are present, estimates are more accurate than when using the conventional filter. In contrast to previous research, our methodology is also efficient when there are no outliers. The performances of different loss functions for solving the estimation problem are presented. The results show that redescending M-estimators outperform the Huber function. The behaviour of the technique is analyzed for a copolymerization process.

KEYWORDS

Polymers, robust statistics, state estimation, unscented transformation

1 INTRODUCTION

Process control is important in polymer manufacturing. Polymers are a "product by process," ie, the microstructure of the material, and therefore its application properties are strongly dictated by process operating conditions. In consequence, it is necessary to monitor and control quality variables associated with the polymer product. However, this is a difficult task due to the high non-linearity of polymerization processes, the infeasibility to measure some variables, and the presence of uncertainty; the latter may enter through noise sources that continuously affect the process and the available measurements.

State estimators are widely used for the implementation of monitoring and control strategies.[1] Ali et al[2] provided a comprehensive survey of different types of estimators applied to chemical process systems, and classified them into six classes. Most of these estimators need an accurate knowledge of the process dynamics. Bayesian estimator is one of those six classes; this approach uses the available measurements of a system to obtain the probability distribution of state estimates.

The availability of first-principles mathematical models in the literature that describe the behaviour of polymerization processes, and the report of successful applications of state estimators, have motivated the development of non-linear state-estimation methods for polymerization processes.[3,4]

Filtering methods are estimators based on Bayesian approach, which estimate unmeasurable properties from available measurements. Commonly, these methods use a recursive algorithm that assumes uncertainty as an additive white noise with a Gaussian distribution.[5] The extended Kalman filter (EKF) is a widespread filtering method for non-linear systems. This technique works fairly well only if the non-linearity is not severe.[5] Some contributions regarding the application of EKF to polymerization processes can be found in the literature.[6,7]

A more recent filtering method suitable for non-linear systems is the unscented Kalman filter (UKF). This approach makes use of the unscented transformation (UT) proposed by Julier and Uhlmann,[8] which estimates the statistics of a random variable when it undergoes a non-linear transformation. This technique uses a deterministic set of sample points, called sigma points, to represent the stochastic distribution of the states. The UKF has provided better results than EKF for the monitoring of polymerization processes.[9,10] Moreover, control strategies that use the UKF for non-linear state estimation have shown very good performance for those processes.[11,12]

Recently, particle filtering (PF) has also been applied to polymerization processes. This is a Monte Carlo simulation-based method that, unlike the EKF and UKF, does not assume the presence of Gaussian noise. This approach recursively estimates the conditional probability density function (pdf) of the states given the measurements using a set of random samples called particles, which have associated weights. If the number of particles increases to infinity, the approximated conditional pdf converges to the true one. Several works analyzed the performance of the EKF, UKF, and PF for the state estimation in polymerization processes, showing that both PF and UKF attained more accurate estimates than the EKF.[13–16] Besides, the PF is less accurate than the UKF when uncertainty is present, such as plant-model mismatch in polymerization processes.[17]

The previous discussion indicates that strategies based on both non-linear UT filters and PF are attractive for state estimation in polymerization processes. However, certain drawbacks are remarked for the second approach in the aforementioned works: strong computational complexity and the necessity of good approximations for the initial pdf of the states.

Several UKF algorithms have been proposed that mainly differ in how the noise is injected into the system, how the set of sigma points is selected, and how the weights are calculated.[18] These algorithms usually model uncertainty as an additive and random noise. However, these two assumptions may not always be appropriate. A more rigorous approach than additive noise is to include noise into the non-linear functions. The augmented UKF (AUKF) is a filtering strategy that joins together the states and both process and measurement noises in a single augmented state vector, and then proceeds as the conventional UKF. Kolås et al[19] applied the AUKF in considering different noise models to represent the uncertainty in control inputs, auxiliary variables, and time-varying parameters, and showed improvements in the estimation performance.

With respect to the assumption of random noise, some samples of the process and measurements noises may not follow a random Gaussian distribution. In this context, atypical observations are called outliers, and even a single outlier can have a large distorting influence on the statistical method. Different types of M-estimators, which are generalizations of the maximum likelihood estimator, have been proposed for developing robust estimation techniques that avoid the detrimental effect of outliers.[20]

The advantages and drawbacks of the M-estimators have been analyzed in the literature.[21] However, applications of the robust conventional UKF are limited to the monotone Huber M-estimator.[22] Besides, the AUKF strategies that include noise into the non-linear functions do not use M-estimators to deal with the presence of outliers. Furthermore, it should be remarked

that the trade-off between efficiency and robustness of filtering methods based on robust statistics has not been previously addressed.

This work presents a new efficient and robust state-estimation procedure based on the AUKF, called efficient and robust augmented UKF (ERAUKF), which deals with the presence of atypical measurements using the concepts of robust statistics. Unlike previously reported robust-statistics based filters, the procedure proposed in this work improves the efficiency using a critical value of the weight function to switch between a conventional AUKF and the ERAUKF approach. In addition, modifications of the filter update step are introduced to enhance estimates accuracy. Process noises are related to uncertainties in auxiliary variables and time-varying parameters, which enter into the non-linear process model. The performance analysis of different M-estimators allows selecting the most appropriate estimator function. The behaviour of the estimation technique is analyzed for the copolymerization of methyl methacrylate (MMA) and vinyl acetate (VA) in a continuous stirred-tank reactor (CSTR) with a recycle loop.[23]

The paper is structured as follows. In section 2, the features of the M-estimators are briefly revisited. In section 3, the proposed methodology is presented. Section 4 describes the application example. Results are presented and discussed in section 5. Finally, the conclusions are given in section 6.

2 M-ESTIMATORS

Robust statistics aims at providing reliable estimates, not only when data follow a given distribution exactly but also when this happens approximately. This involves a trade-off between efficiency and robustness.

Let us assume that $\hat{x}$ is a location M-estimate of x obtained by solving the following optimization problem:

$$\hat{x} = \arg\min_{x} \sum_{i=1}^{n_s} \rho\left(\frac{y_i - x}{\sigma}\right) = \arg\min_{x} \sum_{i=1}^{n_s} \rho(r_i) \tag{1}$$

where y_i is a measurement of variable x; σ is a known dispersion estimate; r_i is the *i*-th residue; n_s is the number of samples; and $\rho(.)$ represents the loss function (LF) of an M-estimator. It has been demonstrated that the distribution of $\hat{x}$ is approximately $N(x_o, v/n_s)$ when n_s increases. Then, $\hat{x}$ is asymptotically normal with asymptotic value x_o and asymptotic variance

v.[20] The asymptotic efficiency of $\hat{x}$ is defined as the ratio v_o/v, where v_o is the asymptotic variance of the maximum likelihood estimate, and indicates how near $\hat{x}$ is to the optimum.

To date, only the Huber (Hub) M-estimator has been used for all robust UKF algorithms reported in the literature. It is a convex function, sensitive to very large values of r; therefore, it presents low efficiency for heavy-tailed error distributions. This can be seen as follows:

$$\rho_{\text{Hub}}(r) = \begin{cases} r^2 & \text{if} \quad |r| \le \text{c}_{\text{Hub}} \\ 2\text{c}_{\text{Hub}}|r| - \text{c}_{\text{Hub}}^2 & \text{if} \quad |r| > \text{c}_{\text{Hub}} \end{cases} \quad (\text{c}_{\text{Hub}} = 1.4) \tag{2}$$

Another type of robust estimators is the so-called redescending M-estimator. These can be made efficient for heavy-tailed data, but require a good starting point to ensure the "good" solution. Redescending M-estimators can be of two types. The first one is unbounded, but the first derivative of the LF, $\psi(.) = \rho'(.)$, called influence function (IF), tends to zero at infinity. The Welsch (Wel) and correntropy (Cor) M-estimators have the following features:

$$\rho_{\text{Wel}}(r) = \text{c}_{\text{Wel}}^2\left(1 - exp\left[-\left(\frac{r}{\text{c}_{\text{Wel}}}\right)^2\right]\right) \quad (\text{c}_{\text{Wel}} = 2.98) \tag{3}$$

$$\rho_{\text{Cor}}(r) = \left(\frac{1}{\text{c}_{\text{Cor}}\sqrt{2\pi}}\right) exp\left[-\left(\frac{r^2}{2\text{c}_{\text{Cor}}^2}\right)\right] \quad (\text{c}_{\text{Cor}} = 2.05) \tag{4}$$

The second type of redescending estimators are bounded, but their IFs are strictly equal to zero for $|r|$ greater than a certain parameter value. The biweight (Biw) M-estimator belongs to the following kind:

$$\rho_{\text{Biw}}(r) = \begin{cases} 1 - \left[1 - \left(\frac{r}{\text{c}_{\text{Biw}}}\right)^2\right]^3 & \text{if} \quad |r| \le \text{c}_{\text{Biw}} \\ 1 & \text{if} \quad |r| > \text{c}_{\text{Biw}} \end{cases} \quad (\text{c}_{\text{Biw}} = 4.68) \tag{5}$$

To compare the estimation capabilities of different M-estimators, they should behave in the same way under the presence of random errors. That is, the asymptotic efficiencies of all M-estimators should be equal. It is a common practice to set them at 95% and then tune the parameter of each LF using the jackknife procedure to satisfy that asymptotic efficiency. The values of parameters c_{Hub}, c_{Wel}, c_{Cor}, and c_{Biw} in Equations (2)-(5) are obtained in this way.

The weight function ω is defined as follows:

$$\omega(r) = \begin{cases} \dfrac{\psi(r)}{r} & \text{if} \quad r \neq 0 \\ \psi'(0) & \text{if} \quad r = 0 \end{cases} \tag{6}$$

It is a non-increasing function of $|r|$; therefore, the weights of the residuals, $\omega(r)$, associated to atypical observations are smaller. This behaviour will be applied in the next section.

3 EFFICIENT AND ROBUST AUGMENTED UKF

Let us consider a dynamic system that is represented by the following non-linear and discrete state-space model:

$$x_k = f(x_{k-1}, u_{k-1}, w_{k-1}) \tag{7}$$

$$y_k = h(x_k, v_k) \tag{8}$$

where x_k and y_k are the *n*-dimensional state vector and the *m*-dimensional measurement vector, respectively, at the time step k (k = 1,…, K); $f(.)$ is the process model function; $h(.)$ is the measurement model function; u_{k-1} is the vector of deterministic control inputs; w_{k-1} is the process noise vector of dimension $q \leq n$; and v_k is the measurement noise vector of dimension m. It is assumed that $w_k \sim N(0, Q_k)$ and $v_k \sim N(0, R_k)$.

In this section, a new procedure to achieve robust state inferences under the presence of atypical measurements, denoted as ERAUKF, is presented. Initially, the methodology uses the AUKF to predict the states and their covariance matrix using the information gathered until time step (*k-1*).[18,24,25] Then, and in contrast to previous robust-statistics-based filters, the strategy checks the presence of atypical observations when y_k is obtained. If outliers are detected a robust update estimation technique is applied. If not, the AUKF is executed because it provides optimal estimates when only random errors are present. The presence of atypical observations is detected using a critical value of the weight function, ω_c, as described below.

The ERAUKF uses an augmented state vector x_{k-1}^A of dimension L, which is defined by incorporating the noise vectors w_{k-1} and v_{k-1} to the state vector. The augmented state covariance matrix P_{k-1}^A is formulated in terms of the state covariance matrix P_{k-1} and the noises covariance matrices, denoted as Q_{k-1} and R_{k-1}, respectively. Then, the predicted state

vector $\hat{x}_{k|k-1}$ and its associated covariance $P_{k|k-1}$ are calculated in the same way as the AUKF algorithm.[5]

When y_k is obtained, the predicted estimations are updated using the new information. This work endeavours to detect the presence of outliers in vector y_k before the estimations update. Assuming $\hat{x}_{k|k-1}$ is a good approximation of x_k, an estimation $\hat{v}_k$ is obtained and then standardized as $r_{k,y}$ by the following:

$$r_{k,y} = R_k^{-1}\left(y_k - h\left(\hat{x}_{k|k-1}\right)\right) = R_k^{-1}\hat{v}_k \tag{9}$$

To determine if the *j*-th measurement of y_k ($j = 1,\ldots,m$) is atypical, ω_j is calculated by replacing the *j*-th element of $r_{k,y}$ in Equation (6). If $\omega_j < \omega_{cj}$, the *j*-th measurement is considered an outlier. Let us recall that $\omega(r)$ diminishes in the presence of atypical observations.

To set ω_c, 10 000 simulations of the prediction step and Equation (9) are run offline (for $k = 1,\ldots, K$) considering that measurements are subject only to random errors. The values of ω are calculated by Equation (6) using the LF of the selected M-estimator for all the generated measurements. Then the empirical cumulative distribution function of ω is obtained, and ω_c is selected as the weight value, for which the cumulative probability is α. This parameter is the probability of wrongly identifying an observation as atypical, and it is set equal to 0.05.

If no measurements of y_k are outliers, the update step of the classic AUKF is run. In contrast, if at least one measurement of y_k is identified as an outlier, the update of the predicted estimates is performed, taking advantage of the ω features. The procedure is explained below.

Assuming the true value of $x_k = \hat{x}_{k|k-1}$, let us define the error vector $\bar{e}_k$ as follows:

$$\bar{e}_k = \bar{z}_k - g(x_k) \tag{10}$$

and each term of Equation (10) can be seen as follows:

$$\bar{e}_k = \begin{bmatrix} e_{k|k-1} \\ v_k \end{bmatrix} = \begin{bmatrix} \hat{x}_{k|k-1} \\ y_k \end{bmatrix} - \begin{bmatrix} x_k \\ h(x_k) \end{bmatrix} \tag{11}$$

where $e_{k|k-1}$ is the error between the true state vector and its prediction.

The impact of events that do not correspond to the normal system behaviour can be reduced by generating a pseudo-measurement vector $\tilde{z}_k$ used as input of the update-estimation step. For

this purpose, first $\bar{e}_k$ is standardized. Let us consider M_k as the error covariance matrix of $\bar{e}_k$, as follows:

$$M_k = E\left\{\begin{bmatrix} e_{k|k-1} \\ v_k \end{bmatrix}\begin{bmatrix} e_{k|k-1} \\ v_k \end{bmatrix}^T\right\} = \begin{bmatrix} P_{k|k-1} & 0 \\ 0 & R_k \end{bmatrix} \tag{12}$$

and the Cholesky decomposition of M_k is represented by S_{M_k}. Multiplying Equation (10) by the inverse of S_{M_k}, $(S_{M_k}^{-1})$, the following residual vector r_k is obtained:

$$r_k = S_{M_k}^{-1} * \bar{e}_k = S_{M_k}^{-1} * \begin{bmatrix} e_{k|k-1} \\ v_k \end{bmatrix} \tag{13}$$

whose covariance matrix is the identity matrix as result of the error decorrelation performed using $S_{M_k}^{-1}$.

For the selected M-estimator, $\omega(r_k)$ is calculated by applying Equation (6). Vector $\tilde{e}_k$, which contains the weighted elements of $\bar{e}_k$, is obtained as follows:

$$\tilde{e}_k = \begin{bmatrix} \tilde{e}_{k|k-1} \\ \tilde{v}_k \end{bmatrix} = \text{diag}(\omega(r_k)) * \begin{bmatrix} e_{k|k-1} \\ v_k \end{bmatrix} \tag{14}$$

Finally, vector $\tilde{z}_k$ that involves the modified predicted states and measurement vectors is evaluated by the following:

$$\tilde{z}_k = g(x_k) + \tilde{e}_k \tag{15}$$

and each term of Equation (15) can be seen as follows:

$$\tilde{z}_k = \begin{bmatrix} \tilde{x}_{k|k-1} \\ \tilde{y}_k \end{bmatrix} = \begin{bmatrix} x_k \\ h(x_k) \end{bmatrix} + \begin{bmatrix} \tilde{e}_{k|k-1} \\ \tilde{v}_k \end{bmatrix} \tag{16}$$

The modified vectors, $\tilde{x}_{k|k-1}$ and $\tilde{y}_k$, are used instead of $\hat{x}_{k|k-1}$ and y_k, respectively, in the update step of the ERAUKF algorithm.

In contrast to previous works that consider $e_{k|k-1} = 0$, an iterative procedure is used until stopping criteria are satisfied regarding the accuracy of $\hat{x}_k$.[22] The ERAUKF algorithm is shown in detail in the Appendix.

The main advantages of the ERAUKF are: (i) the robust procedure is only run if outliers are detected; (ii) robustness is incorporated in the update step of the filtering technique, using information modified by the weight function of the M-estimator; and (iii) the solution is attained using few iterations because the solution procedure uses a good starting point.

4 CASE STUDY

Dynamic operation of polymerization reactors is a complex task. Some of the reasons for this complexity are the extreme sensitivity of the steady state to small changes in parameter values or operating conditions, and the highly interactive non-linear dynamic behaviour often exhibited by these reactors. The process selected as the case study for this work is the copolymerization of MMA and VA in a CSTR with a recycle loop. This process has been chosen in several works dealing with non-linear state-estimation and process-control applications.[9,23] Figure 1 shows a diagram of the copolymerization reactor.

The fresh stream (F_1) contains the monomers MMA and VA, the initiator azoisobutyronitrile (AIBN), the solvent benzene (B), and the chain transfer agent (CTA), acetaldehyde. Besides, an inhibitor (INH) could be present. Stream F_1 is combined with a recycle stream (F_2) to form the reactor feed stream (F_3). The reactor outlet stream (F_4) goes to a separator. In this equipment, unreacted monomers and solvent continue to the purge point (F_8). After the purge, the remaining unreacted monomers and solvent in the recycle stream (F_9) are stored in a holding tank.

The mathematical model includes a comprehensive description of the generally accepted kinetic steps of this system. It is a generalization of models that have appeared in the literature and validated experimentally, which have proven to capture the essential dynamics of the system.[23,26–28] Details about the model equations, such as algebraic equations for stream connections, reaction rates, global concentration of radicals, and parameters can be found elsewhere.[23]

The mathematical model of this process comprises a set of differential algebraic equations for the mass and energy balances of the system. The state vector x involves the following variables:

$$x = \left[C_{MMA}, C_{VA}, C_{AIBN}, C_B, C_{CTA}, C_{inh}, T_r, \Gamma_{MMA}, \Gamma_{VA}, \gamma_0, \gamma_1, \gamma_2, C_{MMA,s}, C_{VA,s}, C_{AIBN,s}, C_{B,s}, C_{CTA,s}, C_{inh,s}, C_{MMA,h}, C_{VA,h}, C_{AIBN,h}, C_{B,h}, C_{CTA,h}, C_{inh,h}\right] \quad (17)$$

where C_j is the molar concentration of species j; T_r is the reactor temperature; Γ_j is the molar concentration of monomer j in the copolymer; γ_j is the j-th order moment of the copolymer molecular weight distribution (MWD); and subscripts s and h stand for separator and hold tank, respectively.

The general expressions of the reactor mass and moment balances by the following:

$$\frac{dC_j}{dt} = \frac{C_{j,f} - C_j}{\theta_r} + \phi_{C_j} \quad (j = \text{MMA, VA, AIBN, B, CTA, inh}) \tag{18}$$

$$\frac{d\Gamma_j}{dt} = \frac{\Gamma_{j,f} - \Gamma_j}{\theta_r} + \phi_{\Gamma_j} \quad (j = \text{MMA, VA}) \tag{19}$$

$$\frac{d\gamma_j}{dt} = \frac{\gamma_{j,f} - \gamma_j}{\theta_r} + \phi_{\gamma_j} \quad (j = 0, 1, 2) \tag{20}$$

In Equations (18)-(20), subscript f stands for feed stream; ϕ_{C_j}, ϕ_{Γ_j}, and ϕ_{γ_j} are reaction rates; and θ_r is the reactor residence time.

The separator and hold-tank balances are as follows:

$$\frac{dC_{j,s}}{dt} = \frac{C_{j,s,f} - C_{j,s}}{\theta_s} \quad (j = \text{MMA, VA, AIBN, B, CTA, inh}) \tag{21}$$

$$\frac{dC_{j,h}}{dt} = \frac{C_{j,h,f} - C_{j,h}}{\theta_h} \quad (j = \text{MMA, VA, AIBN, B, CTA, inh}) \tag{22}$$

where θ_s and θ_h are the residence time in the separator and in the hold tank, respectively.

The following is the reactor energy balance:

$$\begin{aligned}\frac{dT_r}{dt} = \frac{T_{r,f} - T_r}{\theta_r} + \frac{\left(-\Delta H_{paa}\right)k_{paa}C_{MMA}C^*_{MMA} + \left(-\Delta H_{pba}\right)k_{pba}C_{MMA}C^*_{VA}}{\rho_r C_{p,r}} \\ + \frac{\left(-\Delta H_{pab}\right)k_{pab}C_{VA}C^*_{MMA} + \left(-\Delta H_{pbb}\right)k_{pbb}C_{VA}C^*_{VA}}{\rho_r C_{p,r}} - \frac{U_r A_r\left(T_r - T_{jack}\right)}{V\rho_r C_{p,r}}\end{aligned} \tag{23}$$

where C^*_{MMA} and C^*_{VA} are the global concentration of radicals with a MMA and a VA terminal unit, respectively; ΔH_{paa}, ΔH_{pba}, ΔH_{pab}, and ΔH_{pbb} are reaction enthalpies; k_{paa}, k_{pba}, k_{pab}, and k_{pbb} are kinetic parameters of the propagation reactions; a = MMA and b = VA; ρ_r and $C_{p,r}$ represent the density and heat-capacity of the reaction mixture, respectively; U_r is the overall heat-transfer coefficient; A_r is the overall transfer area; V is the reactor volume; and T_{jack} is the temperature of the reactor jacket.

The measurement vector y considered for this work is as follows:

$$y = [Conv_{Total}, T_r, \bar{M}_W] \tag{24}$$

where $Conv_{Total}$ is the total conversion of the two monomers and $\bar{M}_W$ is the weight-average molecular weight. It is assumed that T_r, $Conv_{Total}$, and $\bar{M}_W$ are on-line measurements available from the process. Measurement T_r is acquired by a thermocouple; and data of $Conv_{Total}$ and $\bar{M}_W$ could be obtained from on-line sensors that use empirical correlations. For instance, on-line sensors can be directly plugged in the reacting mixture to get $Conv_{Total}$ through ultrasound propagation velocity measurements.[29] Besides, an on-line sensor based on the liquid viscosity measurements can be installed at the outlet of the reactor to estimate $\bar{M}_W$.[30] On-line viscosity measurement is a typical approach used in polymer control to estimate the weight-average molecular weight of the polymer,[31,32] with errors normally lying within the error bounds of direct experimental measurements of $\bar{M}_W$.[33] The on-line sensor provides estimates of $\bar{M}_W$ immediately after receiving the viscosity measurements, which can be obtained with a frequency up to 1 Hz for typical vibrating rod viscosimeters.[30]

The expressions of $Conv_{Total}$ and $\bar{M}_W$ in terms of the state variables are the following:

$$Conv_{Total} = \frac{\gamma_1}{(\gamma_1 + M_{MMA}C_{MMA} + M_{VA}C_{VA})} \tag{25}$$

$$\bar{M}_W = \gamma_2/\gamma_1 \tag{26}$$

where M_{MMA} and M_{VA} are the molecular weights of MMA and VA monomers, respectively.

5 RESULTS AND DISCUSSION

The analysis considers that the process is initially at steady state and that on-line measurements are continuously available, with a sampling interval of 5 min. At time t = 1 hour, a perturbation occurs in the reactor feed stream leading to a transition to a new steady state. Table 1 shows the values of the relevant variables for this scenario.

Three case studies are considered for analyzing the performance of different filtering techniques:

Case 1. There is uncertainty in all process states, which are regarded as additive noises.

Case 2. There is uncertainty in the reactions rates $\phi_{C_{MMA}}$, $\phi_{C_{VA}}$, $\phi_{C_{AIBN}}$, ϕ_{Y_0}, ϕ_{Y_1}, and ϕ_{Y_2}. This uncertainty is regarded as process noise that enters the non-linear model.

Case 3. Process noise enters, as in Case 2, but affects the kinetic parameters of the aforementioned reactions rates, which are time-varying parameters. In addition, these parameters are unknown and estimated.

The variables and parameters suffering from uncertainty in Cases 2 and 3, respectively, were selected after performing a sensitivity analysis of the process model to choose the most sensitive ones. Uncertainty in the process and in the measurements was assumed as white noise, mean-zero, and normal distribution, ie, $w_k = N(0, Q_k)$ and $v_k = N(0, R_k)$, respectively.

For each case study, the performances of three different filtering techniques are compared with and without outliers. These strategies are the AUKF, robust AUKF (RAUKF), and ERAUKF. The RAUKF is a version of the ERAUKF that does not check the presence of atypical observations; therefore, it always runs a robust update procedure. In addition, different types of M-estimators (Hub, Wel, Cor, Biw) are considered for RAUKF and ERAUKF. The filtering algorithms of these strategies were implemented in MATLAB, noises were generated using the MATLAB randn function, and the Equations (18)-(23) are a set of differential-algebraic equations that were integrated numerically using a solver function ode15s. This solver allows the computation of both dynamic and steady state values for the state vector in Equation (17).

The mean square error (MSE) was used as performance measure. It was calculated as the average of the square differences between the true values of the state variables and measurements and the estimated ones for 10 000 simulation trials of the dynamic system. When outliers are included, they are randomly generated in the range from 10 to 15 times the standard deviation for measurements $Conv_{Total}$, T_r, and $\bar{M}_W$.

The critical values of ω_{cj} (j=1,…, m) were calculated using the empirical cumulative distribution function (ecdf) in MATLAB. For each ω_j, the sample is formed by $K \times 10^4$ values

obtained running RAUKF for different M-estimators without including outliers. The ecdf returns the value of ω_j, for which the cumulative distribution is 5%, and this is set as ω_{cj}.

Comparison results are displayed in Tables 2-9. For measurement $\bar{M}_W$, Tables 2 and 3 show the MSE reduction achieved using RAUKF (expressed in percentage) in comparison to the method proposed by Chang et al,[22] when only random errors are considered and when outliers are present, respectively.

Results in Table 2 and 3 indicate the accuracy of the estimations increases (the MSE diminishes) when using an iterative robust technique in comparison to the method that sets $e_{k|k-1} = 0$ and does not perform an iterative solution scheme. It should be remarked that the proposed procedure is fast; only three to five iterations are sufficient to achieve convergence given that the predicted estimates provide a good initial point.

Tables 4 and 5 display the MSE of $\bar{M}_W$ attained using AUKF, RAUKF, and ERAUKF, with random errors and atypical measurements, respectively. Table 6 shows the MSE reduction percentages of RAUKF and ERAUKF with respect to AUKF in the presence of outliers. In a similar manner, Table 7 shows the MSE reduction percentages of the ERAUKF regarding RAUKF.

When outliers are not present, AUKF provides the best estimates in comparison with robust methodologies for the three cases (Table 4), because AUKF uses the least square estimator that satisfies a set of optimality conditions when there are only random errors. However, an improvement of the trade-off between efficiency and robustness is noticed for ERAUKF-Biw since it attains MSEs similar to the ones obtained using AUKF for the analyzed scenarios.

Under the presence of outliers, the robust techniques significantly outperform AUKF for the three cases (Table 5). Moreover, the performance of ERAUKF is better than the corresponding one to RAUKF. For the robust techniques, the MSEs attained using the redescending M-estimators are lower than the ones obtained using the Hub function because this is sensitive to outliers. Besides, the second type of redescending M-estimators, eg, the Biw function, gives better results than those provided by the first type, which is represented by the Cor and Wel functions in this work. This occurs because the IF of the Biw estimator is equal to zero when the residual is greater than c_{Biw}. Finally, RAUKF-Cor performs better than RAUKF-Wel because the IF of the Cor function approximates to zero faster than the Wel function does.

Regarding the RAUKF-Biw, which gives better results than RAUKF methodology for other M-estimators, the MSE reduction percentages with respect to AUKF is greater than 65% (Table

6) for the three cases, and the MSE reduction percentages of the ERAUKF-Biw regarding RAUKF-Biw is approximately 20% (Table 7).

The MSE reduction percentages of $\bar{M}_W$ achieved considering non-additive noise (Cases 2 and 3) with respect to additive noise (Case 1) were: (a) 24.32% (Case 2) and 15.70% (Case 3) for random error; and (b) 24.00% (Case 2) and 10.16% (Case 3) in the presence of outliers. These reported values correspond to the MSE results of the ERAUKF-Biw technique.

The MSEs obtained when uncertainty enters into the non-linear model (Cases 2 and 3) are better than those achieved for additive noise (Case 1), even when kinetic parameters are estimated (Case 3). Besides, the performance of Case 2 is better than that of Case 3, because in the latter the kinetic parameters are unknown and estimated as state variables. Furthermore, the MSE reduction percentages of the non-additive noise (Cases 2) with respect to additive noise (Case 1) is approximately 25%, both in the absence and presence of outliers.

Figures 2-4 depict the real and estimated values of measurement $\bar{M}_W$ for each case when atypical measurements are present and different state-estimation techniques are applied. The estimates obtained using AUKF are biased for the three cases, and show a greater detrimental effect for additive noise (Case 1). Regarding the robust techniques, the ERAUKF-Hub provides estimates that tend to the real values but outliers are still affecting the results. In contrast, the attained estimates are close to the real values when ERAUKF-Biw is applied.

A similar behaviour of the proposed strategy is observed for the other measurements, $Conv_{Total}$ and T_r. Results are not included here for the sake of space.

6 CONCLUSIONS

In this work, a new efficient and robust state-estimation methodology based on the augmented unscented Kalman filter, called efficient robust augmented unscented Kalman filter, which deals with the presence of atypical measurements using the concepts of robust statistics, is presented. The behaviour of the estimation procedure is analyzed for the copolymerization of MMA and VA monomers.

The strategy shows a rewarding performance both when measurements are contaminated with outliers and when only random errors are present. This behaviour is achieved using the properties of the weigh function of the M-estimator. A critical value of this function allows switching between different estimate update methods, enhancing robustness and efficiency. If it is necessary to run a robust technique in the update step, an iterative solution scheme is applied that converges in few iterations, thanks to its good initial point. This iterative procedure increases the accuracy of estimates in comparison to previous existing techniques.

As regards the robustness on the estimation, the procedures that use the redescending M-estimators provide better results in comparison with those achieved using the monotone M-estimator, commonly used in the literature. Moreover, the second type of redescending M-estimators, which have bounded loss functions and influence functions equal to zero for residual values greater than their parameters (Biweight M-estimator), outperform all the other M-estimators.

The aforementioned conclusions are based on results obtained both for additive noise and for when uncertainty enters directly into the non-linear functions. As expected, better results were achieved for these case studies.

ACKNOWLEDGMENTS

The authors gratefully acknowledge the financial support of this work by Consejo Nacional de Investigaciones Científicas y Técnicas of Argentina (CONICET) through grant PIP 0653 and by Universidad Nacional del Sur (UNS) through grant PGI 24/M165.

APPENDIX

The ERAUKF algorithm

Table A1 shows the algorithm of the ERAUKF.

REFERENCES

[1] M. Soroush, *Comput. Chem. Eng.* **1998**, *23*, 229.

[2] J. M. Ali, N. H. Hoang, M. A. Hussain, D. Dochain, *Comput. Chem. Eng.* **2015**, *76*, 27.

[3] N. Arora, L. T. Biegler, *Comput. Chem. Eng.* **2001**, *25*, 1585.

[4] N. Sheibat-Othman, D. Peycelon, S. Othman, J. M. Suau, G. Févotte, *Chem. Eng. J.* **2008**, *140*, 529.

[5] D. Simon, *Optimal State Estimation: Kalman, H infinity, and Nonlinear Approaches*. Wiley-Interscience, Hoboken **2006**.

[6] F. V. Lima, R. M. Rajamani, T. A. Soderstrom, J. B. Rawlings, *IEEE Trans. Control Syst. Technol.* **2013**, *21*, 1249.

[7] I. D. Gil, J. C. Vargas, J. P. Corriou, *Ind. Eng. Chem. Res.* **2014**, *53*, 7397.

[8] S. J. Julier, J. K. Uhlmann, presented at Signal Processing, Sensor Fusion, and Target Recognition VI, Orlando, July **1997**.

[9] R. Galdeano, M. Asteasuain, M. Sánchez, *Macromol. React. Eng.* **2011**, *5*, 278.

[10] F. D. Rincón, M. Esposito, P. H. H. de Araujo, C. Sayer, G. A. C. Le Roux, *Macromol. React. Eng.* **2013**, *7*, 24.

[11] N. C. Jacob, R. Dhib, *J. Process Contr.* **2011**, *21*, 1332.

[12] D. Vasanthi, B. Pranavamoorthy, N. Pappa, *IFAC Proceedings Volumes* **2013**, *46*, 301.

[13] A. V. Shenoy, V. Prasad, S. L. Shah, *IFAC Proceedings Volumes* **2010**, *43*, 159.

[14] R. Hashemi, S. Engell, *IFAC PapersOnLine* **2016**, *49*, 365.

[15] Z. Zhang, J. Chen, *Comput. Chem. Eng.* **2014**, *69*, 66.

[16] Z. Zhu, Z. Meng, Z. Zhang, J. Chen, Y. Dai, *ISA Trans.* **2017**, *69*, 281.

[17] A. V. Shenoy, J. Prakash, V. Prasad, S. L. Shah, K. B. McAuley, *J. Process Contr*. **2013**, *23*, 120.

[18] S. Kolås, B. A. Foss, T. S. Schei, *Comput. Chem. Eng.* **2009**, *33*, 1386.

[19] S. Kolås, B. A. Foss, T. S. Schei, *J. Process Contr.* **2009**, *19*, 1111.

[20] R. A. Maronna, R. D. Martin, V. J. Yohai, M. Salibián-Barrera, *Robust Statistic: Theory and Methods (with R)*, 2nd ed., Wiley, Hoboken **2019**.

[21] C. E. Llanos, M. C. Sánchez, R. A. Maronna, *Ind. Eng. Chem. Res.* **2015**, *54*, 5096.

[22] L. Chang, B. Hu, G. Chang, A. Li, *IET Sci. Meas. Technol.* **2012**, *6*, 502.

[23] J. P. Congalidis, J. R. Richards, W. H. Ray, *AIChE J.* **1989**, *35*, 891.

[24] S. J. Julier, J. K. Uhlmann, *Proc. IEEE* **2004**, *92*, 401.

[25] R. van der Merwe, *PhD Thesis*, Oregon Health & Science University, Portland **2004**.

[26] J. W. Hamer, T. A. Akramov, W. H. Ray, *Chem. Eng. Sci.* **1981**, *36*, 1897.

[27] A. D. Schmidt, W. H. Ray, *Chem. Eng. Sci.* **1981**, *36*, 1401.

[28] A. D. Schmidt, A. B. Clinch, W. H. Ray, *Chem. Eng. Sci.* **1984**, *39*, 419.

[29] S. Canegallo, M. Apostolo, G. Storti, M. Morbidelli, *J. Appl. Polym. Sci.* **1995**, *57*, 1333.

[30] R. Hashemi, D. Kohlmann, S. Engell, *IFAC Proceedings Volumes* **2013**, *46*, 726.

[31] G. E. Fonseca, M. A. Dubé, A. Penlidis, *Macromol. React. Eng.* **2009**, *3*, 327.

[32] T. Meyer, J. Keurentjes, *Handbook of Polymer Reaction Engineering*, Wiley-VCH, Weinheim **2005**.

[33] M. P. Vega, E. L. Lima, J. C. Pinto, *Polymer* **2001**, *42*, 3909.

Figure Captions

FIGURE 1 Copolymerization reactor

FIGURE 2 Real and estimated values of measurement $\bar{M}_W$ for Case 1 with outliers

FIGURE 3 Real and estimated values of measurement $\bar{M}_W$ for Case 2 with outliers

FIGURE 4 Real and estimated values of measurement $\bar{M}_W$ for Case 3 with outliers

GRAPHICAL ABSTRACT Efficient and robust state estimation

Table Captions

TABLE 1 Initial conditions and process perturbations

TABLE 2 Mean square error reduction (%) of efficient and robust augmented unscented Kalman filter regarding the method by Chang et al[22] (random errors)

TABLE 3 Mean square error reduction (%) of robust augmented unscented Kalman filter regarding the method by Chang et al[22] (outliers)

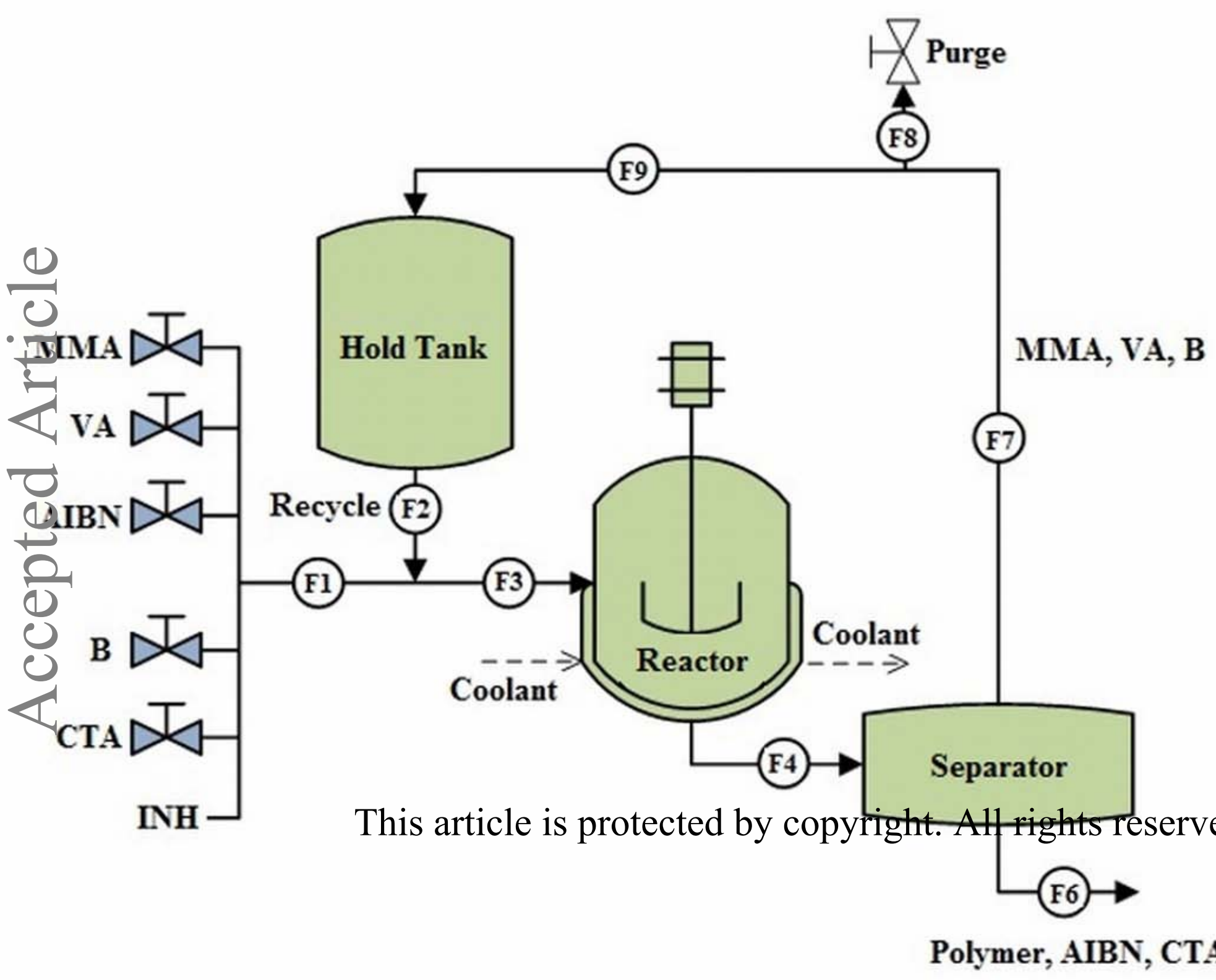
TABLE 4 Mean square error of $\bar{M}_W$ (random errors)

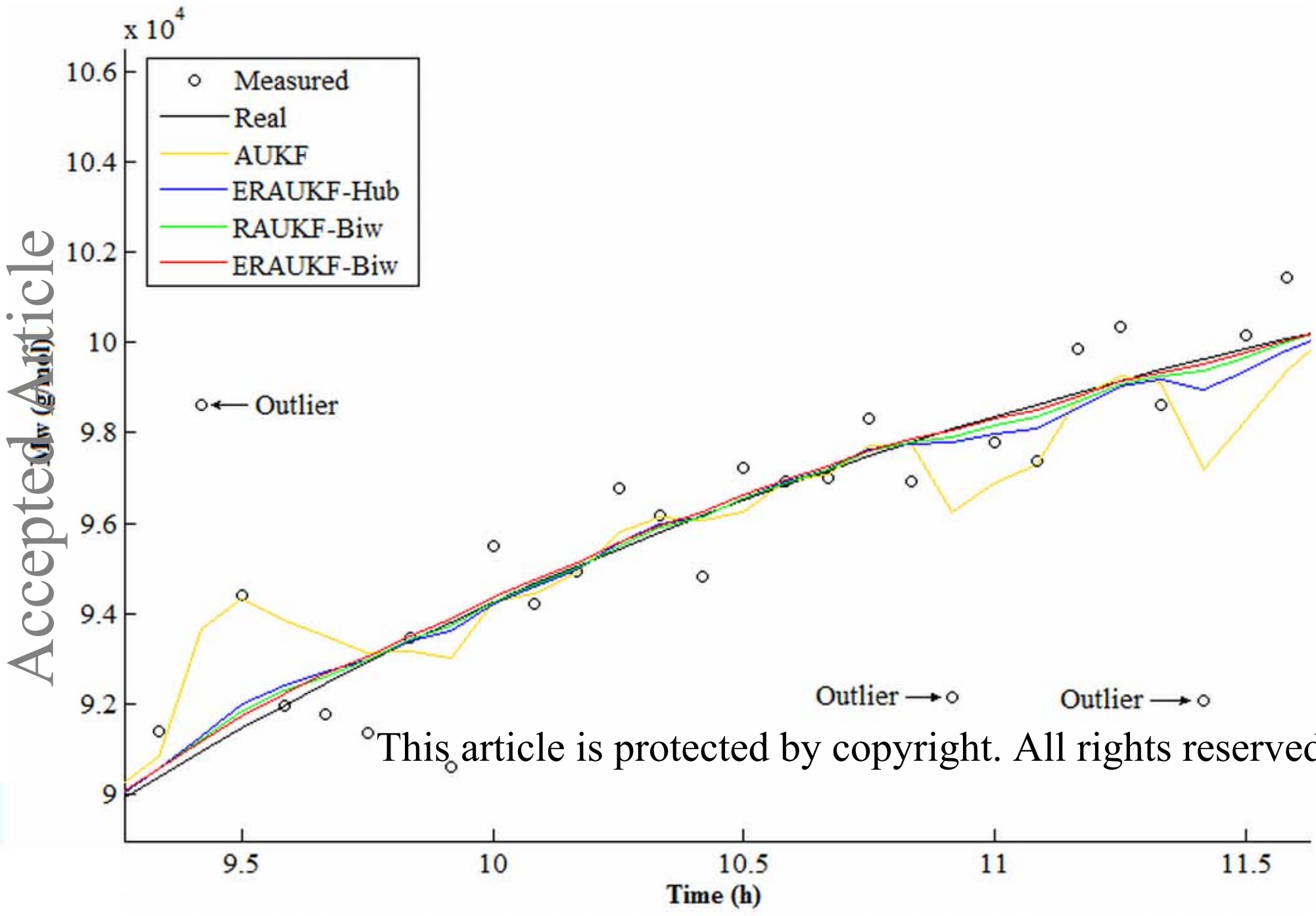
TABLE 5 Mean square error of $\bar{M}_W$ (outliers)

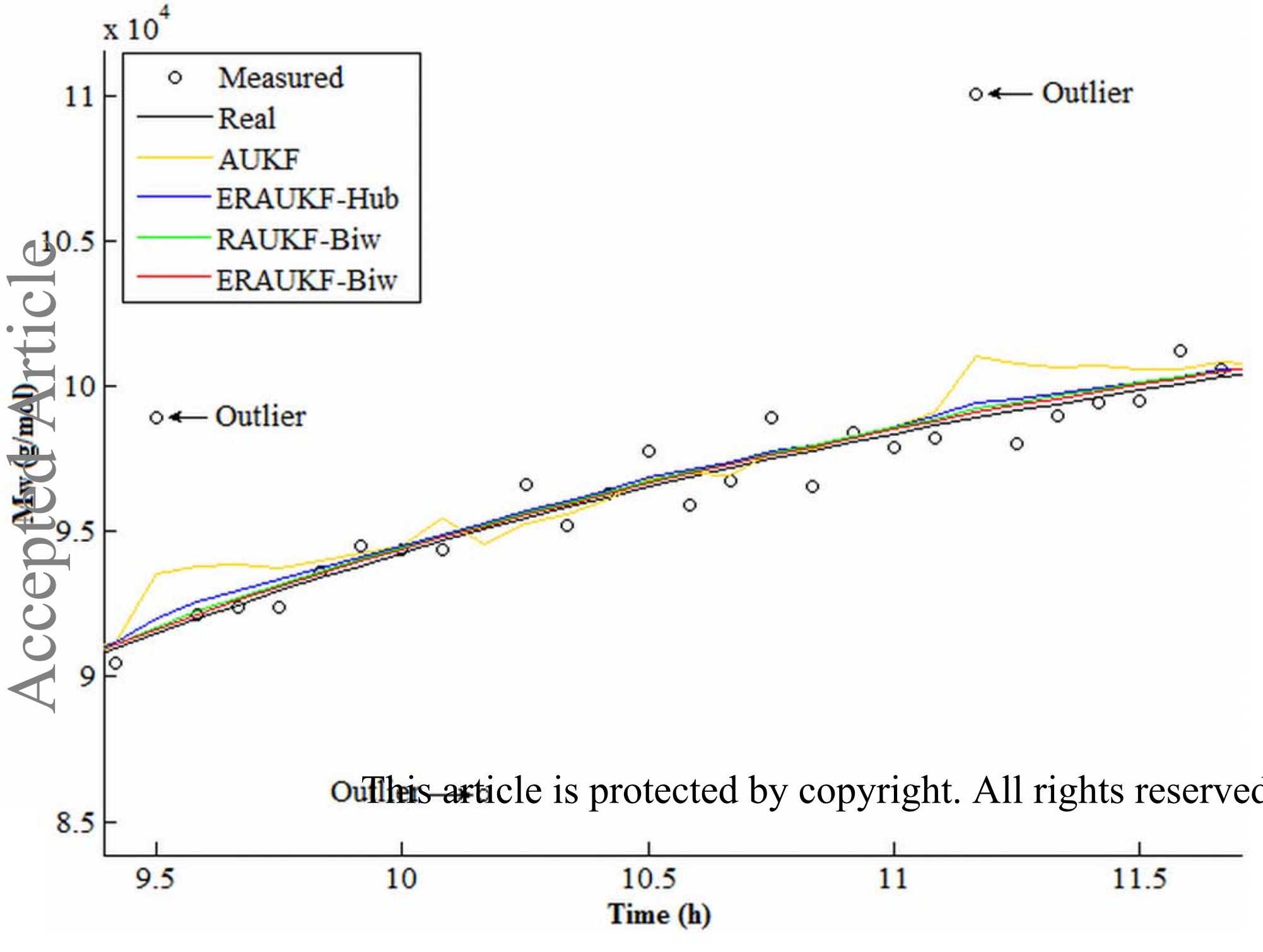
TABLE 6 Mean square error reduction (%) of robust techniques regarding augmented unscented Kalman filter (outliers)

TABLE 7 Mean square error reduction (%) of efficient and robust augmented unscented Kalman filter regarding robust augmented unscented Kalman filter (outliers)

TABLE A1 Efficient and robust augmented unscented Kalman filter algorithm







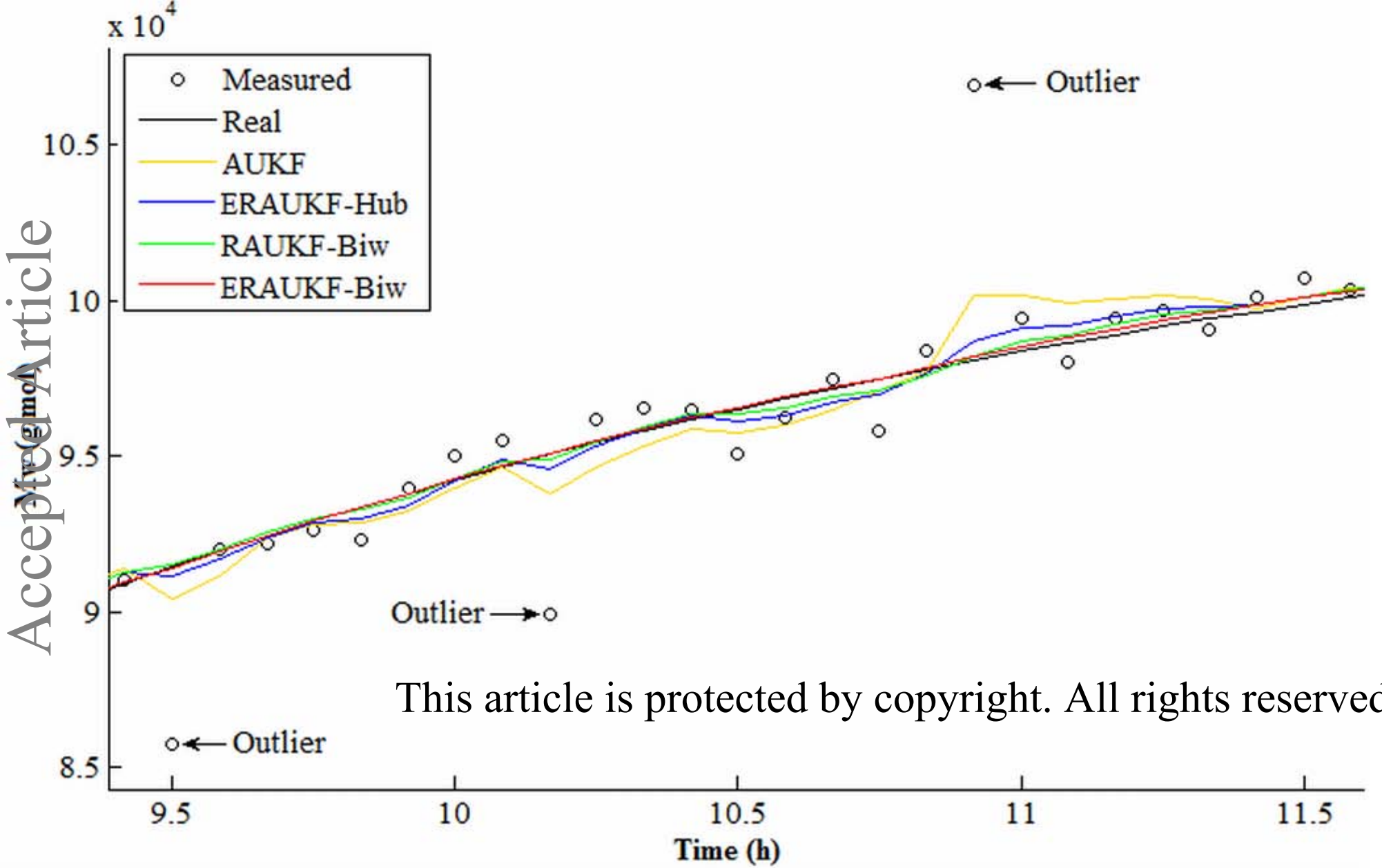


TABLE 1 Initial conditions and process perturbations

Variable	Value	Unit
$C_{MMA,0}$	0.254	(kmol · m^{-3})
$C_{VA,0}$	5.839	(kmol · m^{-3})
$C_{AIBN,0}$	2.012 × 10^{-3}	(kmol · m^{-3})
$C_{B,0}$	2.758	(kmol · m^{-3})
$C_{CTA,0}$	0.366	(kmol · m^{-3})
$C_{inh,0}$	0	(kmol · m^{-3})
$T_{r,0}$	350.426	(K)
$\Gamma_{MMA,0}$	0.822	(kmol · m^{-3})
$\Gamma_{VA,0}$	0.418	(kmol · m^{-3})
$\gamma_{0,0}$	5.401 × 10^{-3}	(kmol · m^{-3})
$\gamma_{1,0}$	119.705	(kg · m^{-3})
$\gamma_{2,0}$	4.177 × 10^{6}	(kg^2/kmol · m^{-3})
$C_{MMA,s,0}$	0.254	(kmol · m^{-3})
$C_{VA,s,0}$	5.839	(kmol · m^{-3})
$C_{AIBN,s,0}$	2.012 × 10^{-3}	(kmol · m^{-3})
$C_{B,s,0}$	2.758	(kmol · m^{-3})
$C_{CTA,s,0}$	0.366	(kmol · m^{-3})
$C_{inh,s,0}$	0	(kmol · m^{-3})
$C_{MMA,h,0}$	0.316	(kmol · m^{-3})
$C_{VA,h,0}$	7.266	(kmol · m^{-3})
$C_{AIBN,h,0}$	0	(kmol · m^{-3})
$C_{B,h,0}$	3.432	(kmol · m^{-3})
$C_{CTA,h,0}$	0	(kmol · m^{-3})
$C_{inh,h,0}$	0	(kmol · m^{-3})
$F_{1,0}$	4.99 × 10^{-5}	(kmol · s^{-1})
$F_{3,0}$	3.04 × 10^{-7}	(kmol · s^{-1})
$F_{1,perturbation}$	4.99 × 10^{-4}	(kmol · s^{-1})
$F_{3,perturbation}$	4.57 × 10^{-7}	(kmol · s^{-1})

TABLE 2 Mean square error reduction (%) of efficient and robust augmented unscented Kalman filter regarding the method by Chang et al[22] (random errors)

Method	$\overline{M}_W$		
	Case 1	**Case 2**	**Case 3**
RAUKF-Hub	10.36	9.59	8.69
RAUKF-Wel	8.86	8.74	8.85
RAUKF-Cor	9.72	8.76	9.70
RAUKF-Biw	10.54	10.25	8.93

TABLE 3 Mean square error reduction (%) of robust augmented unscented Kalman filter regarding the method by Chang et al[22] (outliers)

Method	$\overline{M}_W$		
	Case 1	**Case 2**	**Case 3**
RAUKF-Hub	11.51	12.28	12.43
RAUKF-Wel	11.13	11.60	12.31
RAUKF-Cor	12.76	12.52	12.68
RAUKF-Biw	11.79	11.40	11.35

TABLE 4 Mean square error of $\overline{M}_W$ (random errors)

Method	$\overline{M}_W$		
	Case 1	**Case 2**	**Case 3**
AUKF	$3.045\ 8 \times 10^5$	$2.314\ 4 \times 10^5$	$2.568\ 0 \times 10^5$
RAUKF-Hub	$3.503\ 3 \times 10^5$	$2.878\ 8 \times 10^5$	$3.220\ 1 \times 10^5$
RAUKF-Wel	$3.327\ 3 \times 10^5$	$2.563\ 2 \times 10^5$	$2.921\ 9 \times 10^5$
RAUKF-Cor	$3.310\ 8 \times 10^5$	$2.535\ 1 \times 10^5$	$2.909\ 3 \times 10^5$
RAUKF-Biw	$3.084\ 8 \times 10^5$	$2.359\ 1 \times 10^5$	$2.767\ 1 \times 10^5$
ERAUKF-Hub	$3.061\ 0 \times 10^5$	$2.334\ 7 \times 10^5$	$2.582\ 1 \times 10^5$
ERAUKF-Wel	$3.053\ 2 \times 10^5$	$2.321\ 6 \times 10^5$	$2.573\ 8 \times 10^5$
ERAUKF-Cor	$3.049\ 5 \times 10^5$	$2.320\ 2 \times 10^5$	$2.571\ 1 \times 10^5$
ERAUKF-Biw	$3.046\ 0 \times 10^5$	$2.314\ 8 \times 10^5$	$2.568\ 7 \times 10^5$

TABLE 5 Mean square error of $\overline{M}_W$ (outliers)

Method	$\overline{M}_W$		
	Case 1	**Case 2**	**Case 3**
AUKF	$1.070\ 1 \times 10^6$	$8.028\ 5 \times 10^5$	$9.261\ 3 \times 10^5$
RAUKF-Hub	$4.015\ 8 \times 10^5$	$2.933\ 7 \times 10^5$	$3.498\ 1 \times 10^5$
RAUKF-Wel	$3.754\ 9 \times 10^5$	$2.858\ 4 \times 10^5$	$3.298\ 9 \times 10^5$
RAUKF-Cor	$3.676\ 1 \times 10^5$	$2.798\ 1 \times 10^5$	$3.267\ 1 \times 10^5$
RAUKF-Biw	$3.460\ 4 \times 10^5$	$2.641\ 0 \times 10^5$	$3.162\ 4 \times 10^5$
ERAUKF-Hub	$3.667\ 3 \times 10^5$	$2.768\ 0 \times 10^5$	$3.250\ 3 \times 10^5$
ERAUKF-Wel	$3.436\ 0 \times 10^5$	$2.627\ 0 \times 10^5$	$3.033\ 7 \times 10^5$
ERAUKF-Cor	$3.348\ 3 \times 10^5$	$2.592\ 1 \times 10^5$	$3.005\ 2 \times 10^5$
ERAUKF-Biw	$3.209\ 1 \times 10^5$	$2.427\ 5 \times 10^5$	$2.883\ 1 \times 10^5$

TABLE 6 Mean square error reduction (%) of robust techniques regarding augmented unscented Kalman filter (outliers)

Method	$\overline{M}_W$		
	Case 1	**Case 2**	**Case 3**
RAUKF-Hub	62.47	62.71	62.23
RAUKF-Wel	64.91	64.40	64.38
RAUKF-Cor	65.65	65.15	64.72
RAUKF-Biw	67.66	67.11	65.85
ERAUKF-Hub	65.73	65.52	64.90
ERAUKF-Wel	67.89	67.28	67.24
ERAUKF-Cor	68.71	67.71	67.55
ERAUKF-Biw	70.01	69.76	68.87

TABLE 7 Mean square error reduction (%) of efficient and robust augmented unscented Kalman filter regarding robust augmented unscented Kalman filter (outliers)

Method	$\overline{M}_W$		
	Case 1	**Case 2**	**Case 3**
ERAUKF-Hub	8.68	7.54	7.09
ERAUKF-Wel	14.44	12.25	13.28
ERAUKF-Cor	16.62	13.42	14.09
ERAUKF-Biw	20.09	18.91	17.58