# ADAPTIVE KALMAN FILTERING THROUGH FUZZY LOGIC

P. J. Escamilla-Ambrosio, N. Mort
Department of Automatic Control and Systems Engineering
University of Sheffield
Mappin Street, Sheffield S1 3JD, United Kingdom
Tel: +44 (0) 114 222 5619, Fax: +44 (0) 114 222 5661
E-mail: COP99PJE@Sheffield.ac.uk

**Abstract**– In this paper a development of an adaptive Kalman filter through a fuzzy inference system (FIS) is outlined. The adaptation is concerned with the imposition of conditions under which the filter measurement noise covariance matrix $R$ or the process noise covariance matrix $Q$ are estimated. The adaptive adjustment is carried out using a FIS based on the whiteness of the filter innovation sequence (IS) and employing the covariance-matching technique. If a statistical analysis of the IS shows discrepancies with its expected statistics then the FIS adjusts a factor through which the matrices $R$ or $Q$ are estimated. This fuzzy adaptive Kalman filter is tested on a numerical example. The results are compared with these obtained using a conventional Kalman filter and a traditionally adapted Kalman filter. The fuzzy-adapted Kalman filter showed better results than its traditional counterparts.

**Key words**: Adaptive Kalman filtering; fuzzy systems; filter adaptation; innovation sequence; covariance-matching technique.

## 1. Introduction

The Kalman filter is an optimal recursive data processing algorithm [Maybeck, 1979] that provides a linear, unbiased, and minimum error variance estimate of the unknown state vector $x_k \in \Re^n$ at each instant $k$ = 1,2,…, (indexed by the subscripts) of a discrete-time controlled process that is governed by the linear stochastic difference equation

$$x_{k+1} = A_k x_k + B_k u_k + w_k \qquad (1)$$

where $x_k$ is an $(n \times 1)$ system state vector, $A_k$ is an $(n \times n)$ transition matrix, $u_k$ is an $(l \times 1)$ vector of the input forcing function, $B_k$ is an $(n \times l)$ matrix, and $w_k$ is an $(n \times 1)$ process noise vector. The discrete vector measurement $z_k \in \Re^m$ is given by

$$z_k = H_k x_k + v_k \qquad (2)$$

where $z_k$ is a $(m \times 1)$ measurement vector, $H_k$ is a $(m \times n)$ measurement matrix, and $v_k$ is a $(m \times 1)$ measurement noise vector.

Both $w_k$ and $v_k$ are assumed to be uncorrelated zero-mean Gaussian white noise sequences with covariances

$$E\{w_k w_i^T\} = \begin{cases} Q_k, & i = k \\ 0 & i \neq k \end{cases} \qquad (3)$$

$$E\{v_k v_i^T\} = \begin{cases} R_k, & i = k \\ 0 & i \neq k \end{cases} \qquad (4)$$

$$E\{w_k v_i^T\} = 0 \qquad \text{for all } k \text{ and } i \qquad (5)$$

where $E\{\cdot\}$ is the statistical expectation, superscript $T$ denotes transpose, $Q_k$ is the process noise covariance matrix, and $R_k$ is the measurement noise covariance matrix.

The Kalman filter algorithm has two groups of equations [Welch and Bishop, 1995],

i) Time update (or prediction) equations:

$$\hat{x}_{k+1}^- = A_k \hat{x}_k + B_k u_k \qquad (6)$$

$$P_{k+1}^- = A_k P_k A_k^T + Q_k \qquad (7).$$

These equations project, from time step $k$ to step $k$+1, the current state and error covariance estimates to obtain the *a priori* (indicated by the super minus) estimates for the next time step.

ii) Measurement update (or correction) equations:

$$K_k = P_k^- H_k^T [H_k P_k^- H_k^T + R_k]^{-1} \qquad (8)$$

$$\hat{x}_k = \hat{x}_k^- + K_k [z_k - H_k \hat{x}_k^-] \qquad (9)$$

$$P_k = [I - K_k H] P_k^- \qquad (10).$$

These equations incorporate a new measurement into the *a priori* estimate to obtain an improved *a posteriori* estimate.

In the above equations, $\hat{x}_k$ is an estimate of the system state vector $x_k$, and $P_k$ is the covariance matrix corresponding to the state estimation error defined by

$$P_k = E\{(x_k - \hat{x}_k)(x_k - \hat{x}_k)^T\} \qquad (13),$$

the term $H_k \hat{x}_k^-$ is the one-stage predicted output $\hat{z}_k$, and $(z_k - H_k \hat{x}_k^-)$ is the one-stage prediction error sequence, also referred to as the innovation sequence or residual, generally denoted as *r* and defined as:

$$r_k = (z_k - H_k \hat{x}_k^-) \qquad (14).$$

The innovation represents the additional information available to the filter in consequence to the new observation $z_k$. For an optimal filter the innovation sequence is a sequence of independent Gaussian random variables. The weighted innovation, $K_k[z_k - H\hat{x}_k^-]$, acts as a correction to the predicted estimate $\hat{x}_k^-$ to form the estimation $\hat{x}_k$; the weighting matrix $K_k$ is commonly referred to as the filter gain or the Kalman gain matrix.

The matrices $A_k$, $B_k$ and $H_k$ are assumed to be known. $Q_k$ and $R_k$ are nonnegative definite matrices whose values are also assumed known. The Kalman filter algorithm starts with initial conditions at $k = 0$ being: $\hat{x}_0^-$, and $P_0^-$. With the progression of time, as new measurements $z_k$ become available, the cycle estimation-correction of states and the corresponding error covariances can follow recursively ad infinitum.

**2. Statement of the problem**

The Kalman filter formulation as described previously assumes complete *a priori* knowledge of the process and measurement noise statistics, matrices *Q* and *R*. However, in most practical applications these statistics are initially estimated or in fact are unknown. The problem here is that the optimality of the estimation algorithm in the Kalman filter setting is closely connected to the quality of these a priori process noise and measurement noise statistics [Brown and Hwang, 1997; Mehra, 1970; Fitzgerald, 1971]. It has been shown that inadequate initial statistics of the filter will reduce the precision of the estimated states or will introduce biases to the estimates. In fact, wrong a priori information could cause practical divergence of the filter [Fitzgerald, 1971]. Additionally, insufficient a priori information and a frequently changing estimation environment will affect the accuracy of the Kalman filter. From the aforementioned it may be argued that using a fixed Kalman filter designed by conventional methods in a changing dynamic environment is a major drawback. From this point of view it can be expected that an adaptive estimation formulation of the Kalman filter will result in a better performance or will prevent filter divergence.

Different adaptive procedures have been devised [Mehra, 1972; Moghaddamjoo and Kirlin, 1989; Mohamed and Schwarz, 1999] since the development of the Kalman filter [Kalman, 1960]. The main advantage of the adaptive technique is its weaker reliance on the a priori statistical information. An adaptive filter formulation deals with the problem of having imperfect a priori information and provides an improvement in performance over the fixed filter approach.

The procedures used to adapt a Kalman filter can be classified into two main approaches: innovation-based adaptive estimation (IAE) and multiple-model-based adaptive estimation (MMAE) [Mohamed and Schwarz, 1999]. In the former the adaptation is made directly to the statistical information matrices *R* and/or *Q* based on the changes in the filter innovation sequence. In the second, a bank of Kalman filters runs in parallel with different models for the filter's statistical information. In both techniques the concept of utilising the new information available in the innovation (or residual) sequence is used but they differ in their implementation. In this work only the first approach will be examined, for the second approach the reader is referred to Brown and Hwang [1997].

The IAE approach is based on the improvement of the filter performance through the adaptive estimation of the filter statistical information, the matrices *Q* and/or *R*. The adaptation mechanism is based on the whiteness of the filter innovation sequence, Eq. (14).

The value of the innovation at the current instant *k* cannot be predicted from previous values. Therefore, the innovation represents the additional information available to the filter as a result of the new measurement $z_k$. For this reason the innovation sequence represents the information content in the new observation and is considered the most relevant source of information for the filter adaptation. The occurrence of bad data first shows up in the innovation vector. In this way the innovation sequence reports the discrepancy between predicted and actual measurement. If all prerequisites are met, the innovation sequence is a zero-mean white noise sequence [Dall, 1998

The adaptation procedure in this work is concerned with the imposition of conditions under which the filter statistical information matrices *R* or *Q* are estimated via the available new information given by the filter innovation sequence. We note that these matrices are considered as constants in the conventional Kalman filter.

**3. Adaptive Kalman filtering**

3.1. Adaptive estimation of the measurement noise covariance matrix *R* with *Q* fixed.

The covariance matrix *R* represents the accuracy of the measurement instrument. The enlargement of the covariance matrix *R* for measured data means that we trust this measured data less and more on the prediction. Assuming that the noise covariance matrix *Q* is completely known, an algorithm to estimate the measurement noise covariance matrix *R* can be derived.

Here an IAE algorithm to adapt the matrix *R* has been derived. The technique known as covariance-matching [Mehra, 1972] is used to adapt the covariance matrix *R*. The basic idea behind this technique is to make the residuals consistent with their theoretical covariance [Mohamed and Schwarz, 1999]. The innovation sequence $r_k$ has a theoretical covariance,

$$S_k = H_k P_k^- H_k^T + R_k \qquad (15)$$

obtained from the Kalman filter algorithm. If it is noticed that the actual covariance of $r_k$ has discrepancies with its theoretical value, then adjustments have to be made to *R* in order to correct this mismatch.

To monitor the discrepancy of *S* and its actual value a new variable is defined. This variable is called Degree of Matching (DoM),

$$DoM_k = S_k - \hat{C}_{rk} \qquad (16).$$

Having available the innovation sequence $r_k$, its actual covariance $\hat{C}_{rk}$ in Eq. (16) is approximated by its sample covariance through averaging inside a moving estimation window of size *N* [Mohamed and Schwarz, 1999],

$$\hat{C}_{rk} = \frac{1}{N}\sum_{i=i_0}^{k} r_i r_i^T \qquad (17)$$

where $i_0 = k - N + 1$ is the first sample inside the estimation window. The window size, *N*, is chosen empirically to give some statistical smoothing.

*DoM* is used to indicate the degree of discrepancy between the theoretical value of the innovation covariance *S* and its actual value $\hat{C}_{rk}$. If *DoM* is around zero that means *S* and $\hat{C}_{rk}$ match almost perfectly, then no changes are needed. If *DoM* is greater than zero this means the actual value of $\hat{C}_{rk}$ is smaller than its theoretical value *S*, then an adjustment is needed. Conversely, if *DoM* is smaller than zero, this means the value of $\hat{C}_{rk}$ is greater than its theoretical value *S*, then an adjustment is needed too.

The basic idea of adaptation used by a FIS to adapt *R* is as follows. It can be appreciated from Eq. (15) that an increment in *R* will increment *S*, and vice versa. Thus, *R* can be used to vary *S* in accordance with the value of *DoM* in order to reduce the discrepancies between *S* and $\hat{C}_{rk}$. The next three general adaptation rules are defined:

1. If $DoM \cong 0$ (this means $\hat{C}_{rk}$ and *S* are equal) then maintain *R* unchanged.
2. If $DoM > 0$ (this means $\hat{C}_{rk}$ is smaller than *S*) then decrease *R*.
3. If $DoM < 0$ (this means $\hat{C}_{rk}$ is greater than *S*) then increase *R*.

And *R* is adjusted on this way

$$R_{k+1} = R_k + AdjR \qquad (18).$$

where *AdjR* is the factor that is added or subtracted from *R*. *AdjR* is the FIS output.

3.2. Adaptive estimation of the process noise covariance matrix *Q* with *R* fixed.

The covariance matrix *Q* represents the uncertainty in the process model. An increase in the covariance matrix *Q* means that we trust less the process model and more on the measurement. Assuming that the noise covariance matrix *R* is completely known an algorithm to estimate matrix *Q* can be derived.

The idea behind the process of adaptation of *Q* is as follows. Eq. (15) can be rewritten as:

$$S_k = H_k (A_k P_k A_k^T + Q) H_k^T + R_k \qquad (19)$$

and from Eq. (19) it may be deduced that a variation in *Q* will affect the value of *S*. If *Q* is increased, then *S* is increased, and vice versa. Thus, if a mismatch between *S* and $\hat{C}_{rk}$ is observed then a correction can be made through augmenting or diminishing the value of *Q*. The next three general adaptation rules are defined:

1. If $DoM \cong 0$ (this means $\hat{C}_{rk}$ and *S* are equal) then maintain *Q* unchanged.
2. If $DoM > 0$ (this means $\hat{C}_{rk}$ is smaller than *S*) then decrease *Q*.
3. If $DoM < 0$ (this means $\hat{C}_{rk}$ is greater than *S*) then increase *Q*.

Thus *Q* is adjusted in this way

$$Q_{k+1} = Q_k * AdjQ \qquad (20).$$

where *AdjQ* is a factor obtained with a FIS.

**4. Illustrative example**

To demonstrate the efficiency of the fuzzy-adapted Kalman filter approach, a simple numerical example is presented. The results are compared with those obtained with a Kalman filter without adaptation (KFWA) and a traditionally-adapted Kalman filter (TKF).

Consider the following linear system, which is a modified version of a tracking model [Paik and Oh, 2000; Chen and Chui, 1991],

$$\begin{bmatrix} x_{k+1}^1 \\ x_{k+1}^2 \\ x_{k+1}^3 \end{bmatrix} = \begin{bmatrix} 0.77 & 0.20 & 0.00 \\ 0.25 & 0.75 & 0.25 \\ 0.05 & 0.00 & 0.75 \end{bmatrix} \begin{bmatrix} x_k^1 \\ x_k^2 \\ x_k^3 \end{bmatrix} + \begin{bmatrix} w_k^1 \\ w_k^2 \\ w_k^3 \end{bmatrix} \quad (21a)$$

$$z_k = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_k^1 \\ x_k^2 \\ x_k^3 \end{bmatrix} + v_k \quad (21b)$$

with initial conditions $\hat{x}_0 = 0$, $P_0 = 0.01 I_3$, where $x^1$, $x^2$, and $x^3$ are the position, velocity and acceleration, respectively, of a flying object. In Eq. 21, the system and measurement noise sequences $\{w_k\}$ and $\{v_k\}$ are pseudorandom sequences (i.e., uncorrelated zero-mean Gaussian white noise sequences) with $Q = 0.02I_3$ and $R = 1$.

MATLAB code was developed to simulate the Kalman filter and the fuzzy logic inference system used to adjust the measurement noise covariance matrix $R$ and the process noise covariance matrix $Q$. The results obtained are presented in next sections.

4.1. Fuzzy Adaptation of $R$ and comparisons.

Five fuzzy sets have been defined for *DoM*: *NM* = Negative Medium, *NS* = Negative Small, *ZE* = ZEro, *PS* = Positive Small, and *PM* = Positive Medium; and five fuzzy sets have been defined for *AdjR*: *IL* = Increase Large, *I* = Increase, *M* = Maintain, *D* = Decrease, and *DL* = Decrease Large. Five fuzzy rules comprise the rule base,

1. If *DoM* = *NM*, then *AdjR* = *IL*
2. If *DoM* = *NS*, then *AdjR* = *I*
3. If *DoM* = *ZE*, then *AdjR* = *M*
4. If *DoM* = *PS*, then *AdjR* = *D*
5. If *DoM* = *PM*, then *AdjR* = *DL*.

The membership functions for *DoM* and *AdjR* are presented in figure 1. The model described by Eq. 21 was simulated for 500s with a sample time of 0.5s. $Q$ was fixed as $0.02I_3$. The actual value of $R$ is unity, but it has been assumed unknown. The starting value of $R$ was selected to be,

$$R_0 = 5R \quad (23)$$

The value of $R$ is continuously adjusted once the first value of $\hat{C}_{rk}$ was available. Recall that this last parameter is obtained from a moving estimation window of size $N$.

The following performance measures were adopted for comparison purposes:

$$J_1 = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(z_i - zv_i)^2} \quad (23)$$

$$J_2 = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(z_i - ze_i)^2} \quad (24)$$

where $z_i$ is the actual value of the position; $zv_i$ is the measured position; and $ze_i$ is the estimated position.

The traditional adaptation method proposed by Mohamed and Schwarz (1999) was used to adapt $R$ and $Q$. In this method $R$ or $Q$ are adapted using the following equations,

$$\hat{R}_k = \hat{C}_{rk} - H_k^T P_k^- H_k^T \quad (25)$$

$$\hat{Q}_k = K_k \hat{C}_{rk} K_k^T \quad (26),$$

where $\hat{C}_{rk}$ is obtained with Eq. 18. $H_k^T$, $P_k^-$, and $K_k$ are those obtained in the Kalman filter algorithm.

Table 1 shows the performance measures obtained for each of the three methods: KFWA – Kalman filter without adaptation; TKF – Traditionally-adapted Kalman filter; and FKF – Fuzzy-adapted Kalman filter. From experimentation it was noticed that the best results were obtained with a window size of 200 samples for the TKF and 50 samples for the FKF.

Table 1

| Performance measure | KFWA | TKF | FKF |
|---|---|---|---|
| $J_1$ | 0.9478 | 0.9478 | 0.9478 |
| $J_2$ | 0.3882 | 0.3556 | 0.3541 |

From table 1 it is seen that the best adaptation of $R$ is made by the FKF. Figure 2 shows the outputs obtained from the KFWA. Notice the discrepancy between the actual innovation covariance $\hat{C}_{rk}$ and its theoretical value $S_k$. In this case because both $Q$ and $R$ are fixed *DoM* remains a large value. Figure 3 shows the outputs obtained from the TKF. In this case once the first value of $\hat{C}_{rk}$ is available the adjustment of $R$ is continuously made. It can be seen how *DoM* remains at a very small value while $R$ almost reaches its true value. The filter performance improvement is obvious. Figure 4 shows the outputs obtained from the FKF. As in the TFK case, once the first value of $\hat{C}_{rk}$ is available the adjustment of $R$ starts. In this case because the window size is smaller the adaptation starts earlier. Observe how $S_k$ and $\hat{C}_{rk}$ remain almost equal (DoM is around zero) and $R$ oscillates around its true value. Here it is deduced that the best performance is obtained because the adaptation is achieved in a gradual manner and starts early. The same widow size was tried in the TFK but in that case no improvement in performance was obtained. In fact, for the TKF the best performance was obtained with the window size of 200 samples as reported earlier.

4.2. Fuzzy Adaptation of *Q* and comparisons.

Five fuzzy sets have been defined for *DoM* with the same labels but different membership functions than those in the previous case. In this case only three fuzzy sets have been defined for *AdjQ*: *I* = Increase, *M* = Maintain, and *D* = Decrease. Thus, five fuzzy rules comprise the rule base,

1. If *DoM* = *NM*, then *AdjQ* = *I*
2. If *DoM* = *NS*, then *AdjQ* = *I*
3. If *DoM* = *ZE*, then *AdjQ* = *M*
4. If *DoM* = *PS*, then *AdjQ*= *M*
5. If *DoM* = *PM*, then *AdjQ* = *D*.

The membership functions for the variables *DoM* and *AdjQ* are presented on figure 5. The model described by Eq. 21 was simulated for 500s with a sample time of 0.5s. In this case *R* was fixed as unity. The actual value of *Q* is $0.02I_3$, but it has been assumed as unknown. The starting value of *Q* was selected to be,

$$Q_0 = 5Q \qquad (23).$$

The value of *Q* was continuously adjusted ones the first value of $\hat{C}_{rk}$ was available. Table 2 shows the performance measures obtained for the three methods. In this case, from experimentation it was noticed that the best results were obtained with a window size of 100 samples for both TKF and FKF. As can be seen, the FKF has the best performance as in the previous case.

Table 2

| Performance measure | KFWA | TKF | FKF |
|---|---|---|---|
| $J_1$ | 0.9478 | 0.9478 | 0.9478 |
| $J_2$ | 0.3891 | 0.3528 | 0.3489 |

Figure 6 shows the outputs obtained from the KFWA. Here, as in the previous case, a discrepancy between the actual innovation covariance $\hat{C}_{rk}$ and its theoretical value $S_k$ is observed. *DoM* is not as large as before but it remains close to zero. Figure 7 shows the outputs obtained from the TKF. In this case once the first value of $\hat{C}_{rk}$ is available the adjustment of *Q* is continuous. This adjustment is observed in the variation of *S*, which is attempting to follow $\hat{C}_{rk}$. *DoM* remains around zero. However a good improvement in filter performance is observed. Figure 8 shows the outputs obtained from the FKF. It can be seen that in this case $S_k$ follows $\hat{C}_{rk}$ a little more precisely than in the previous case. This gives a result that is a better improvement in the filter performance than before.

## 5. Conclusions

In this paper a fuzzy inference system to adapt the measurement noise covariance matrix *R* or the process noise covariance matrix *Q* of a Kalman filter have been presented. This method uses the covariance-matching technique to determine if adjustments to *R* or *Q* are needed. An example showing the efficiency of this method was presented. In this example, only five rules were needed to carry out the adaptation in each case. It was observed that better performance was obtained with the fuzzy-adapted Kalman filter than that obtained with both KFWA and TKF. It is relevant to notice that, in the case of the adaptation of R, the window size used to calculate the current innovation covariance is 3 times smaller in the FKF than that used in the TKF. However, for the adaptation of *Q* the same window size can be used. The dependence on the window size for good adaptation in both TKF and FKF is an open line to investigate. In this case only single-input, single-output (SISO) FISs were used and good results were obtained. However, better results can be expected if more than one input to the FISs is used. This is another line of research. The system used to explore the efficiency of the proposed method was simple, so the design of an alternative method for a more sophisticated system may be needed to give more support to the efficiency of the approach.

## REFERENCES


Brown, R. G. and Hwang P. Y. C. (1997*). Introduction to random signals and applied kalman filtering whit matlab exercises and solutions*. Third edition, John Wiley & Sons, New York.

Chen, G. and Chui, C. K. (1991). *IEEE Trans. Aerospace and Electronic Systems*, 27 (1), pp. 149-154.

Dall, Larsen (1998). *Optimal Fusion of Sensors*. PhD Dissertatio, Department of Automation, Technical University of Denmark.

Fitzgerald, R. J. (1971). Divergence of the Kalman filter. *IEEE Trans. Automatic Control*, AC-16 (6), pp. 736- 747.

Hsiao, C-Y. (1999). Modified gain fuzzy kalman filtering algorithm. *JSME International Journal Series C*, 42 (2), pp. 363-368.

Kalman, R. E. (1960). A new approach to linear filtering and prediction problems. *Trans. ASME, J. Basic Eng.* 82, pp. 35-44.

Maybeck, P. S. (1979). *Stochastic models estimation and control*. Volume 1, Academic Press, New York.

Mehra, R. K. (1970). On the identification of variances and adaptive kalman filtering. *IEEE Trans. Automatic Control*, AC-15 (2), pp. 175- 184.

Mehra, R. K. (1972). Approaches to adaptive filtering. *IEEE Trans. Automatic Control*, AC-17, (Oct), pp. 693- 698.

Moghaddamjoo, A. and Kirlin, L. R. (1989). Robust adaptive kalman filtering with unknown inputs. *IEEE Trans. Acoustics, speech, and signal processing*, 37 (8), pp. 1166-1175.
Mohamed, A. H. and Shwarz, K. P. (1999). Adaptive kalman filtering for INS/GPS. *Journal of Geodesy*, 73 (4), pp. 193-203.
Paik, B.S. and Oh, J.H. (2000). Gain fusion algorithm for decentralised parallel Kalman filters. *IEE Proc. Control Theory Appl.*, 147 (1), pp. 97-103.
Welch, Greg and Bishop, Gary (1995). *An Introduction to the Kalman Filter*. The University of North Carolina at Chapel Hill, TR95-045.

**FIGURES**

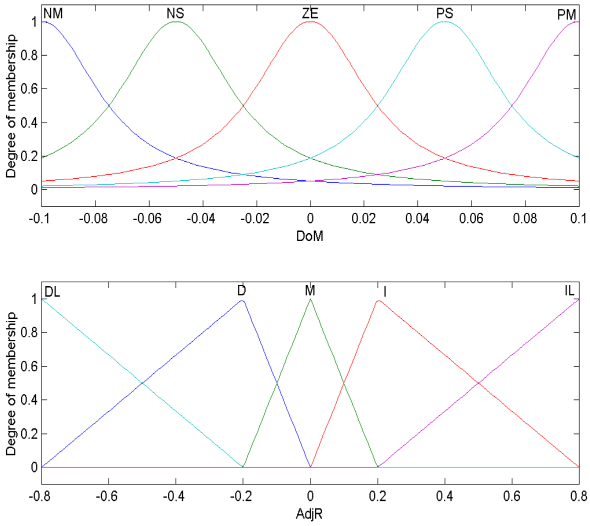


Figure 1. Membership functions for *DoM* and *AdjR*.

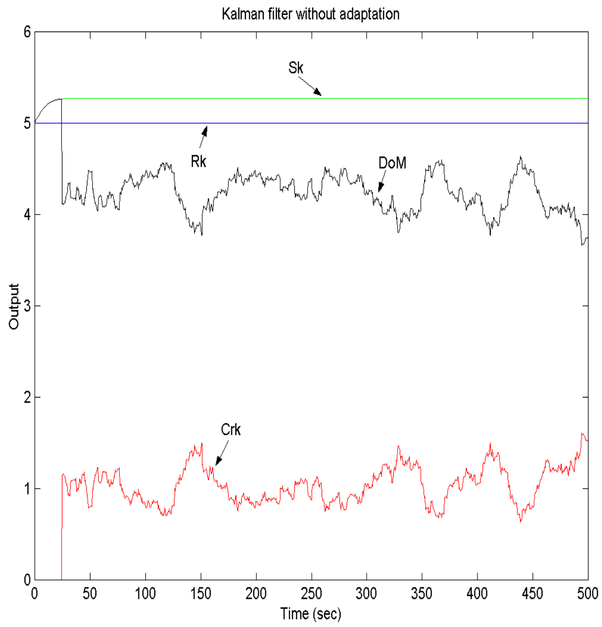


Figure 2. Kalman filter without adaptation, *Q* and *R* fixed.

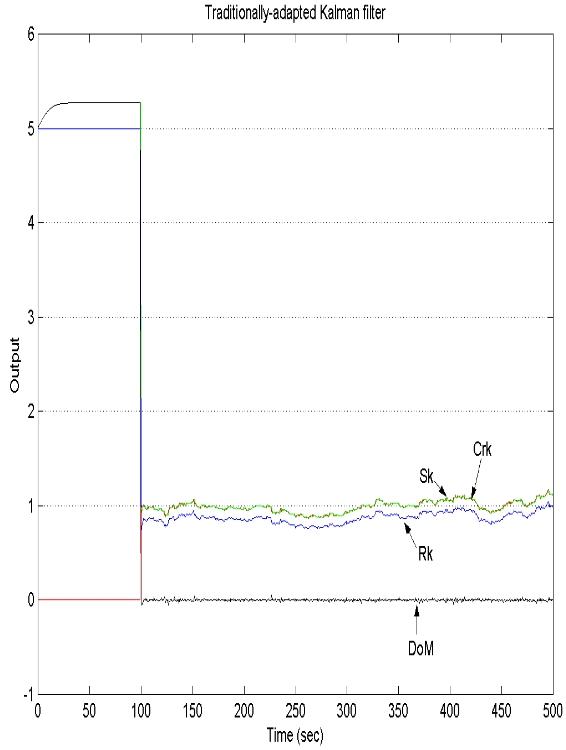


Figure 3. Traditionally-adapted Kalman filter, *Q* fixed.

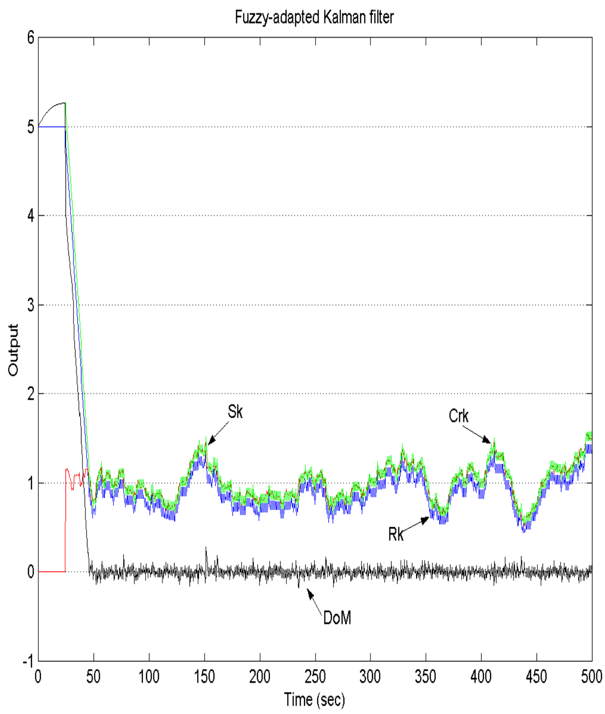


Figure 4. Fuzzy-adapted Kalman filter. *Q* fixed.

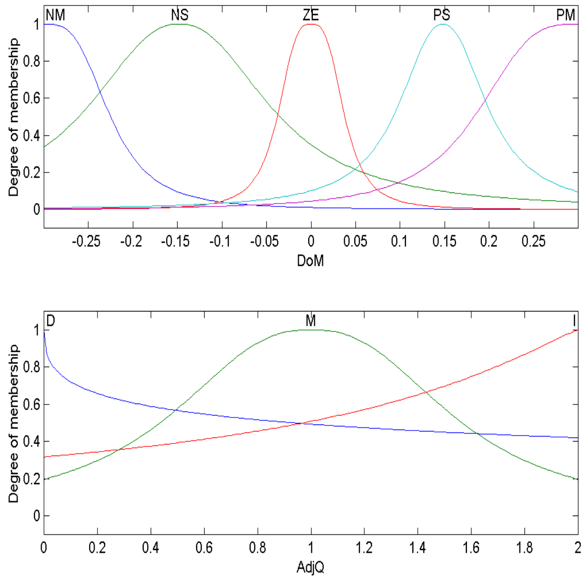


Figure 5. Membership functions for *DoM* and *AdjQ*.

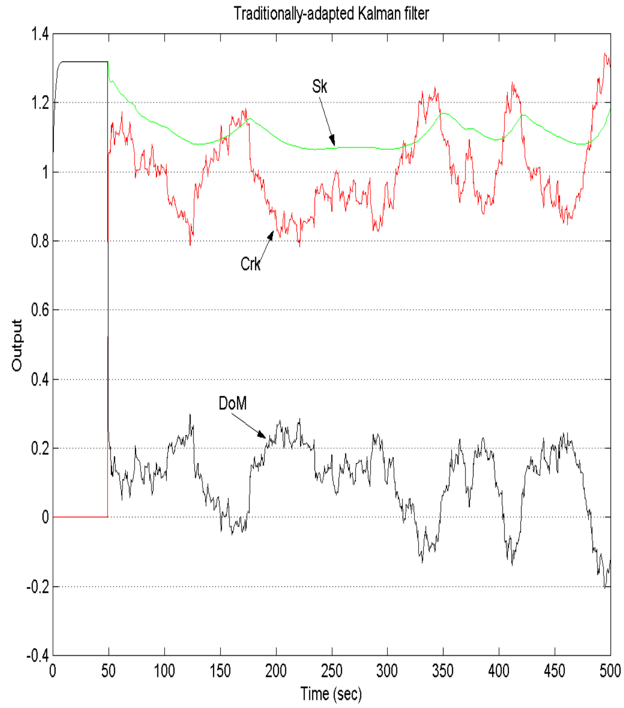


Figure 7. Traditionally-adapted Kalman filter, *R* fixed.

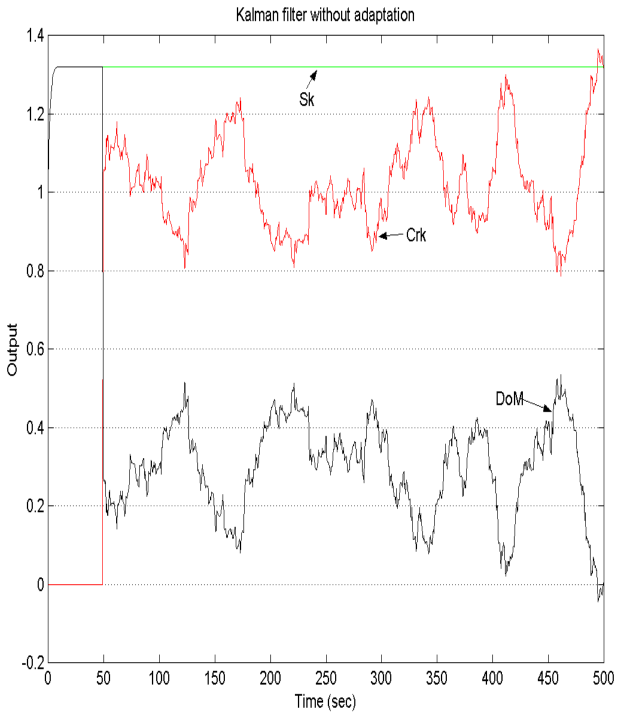


Figure 6. Kalman filter without adaptation, *Q* and *R* fixed.

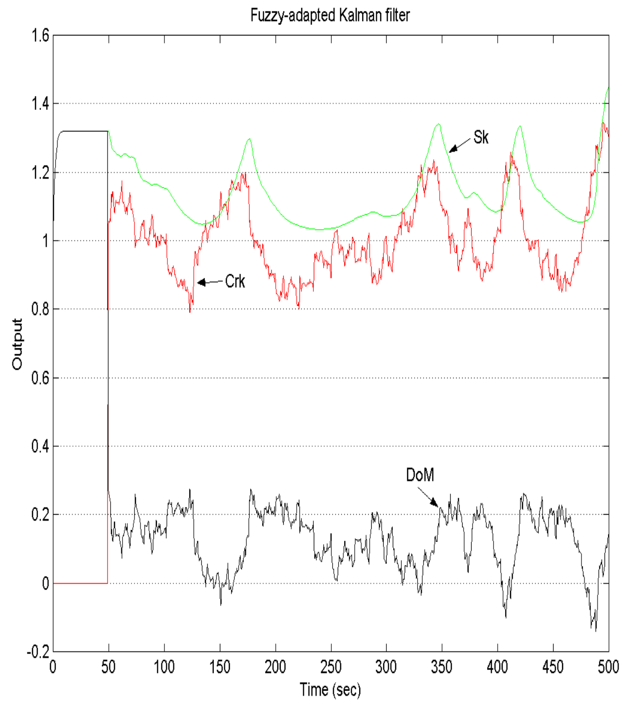


Figure 8. Fuzzy-adapted Kalman filter, *Q* fixed.