Data driven approach for performance assessment of linear and nonlinear Kalman filters

Laya Das$^1$, Babji Srinivasan$^2$ and Raghunathan Rengaswamy$^3$

Abstract—A new technique is developed for assessing the performance of linear and nonlinear Kalman filter based state estimators. The proposed metric will indicate the performance of these state estimators which will be primarily influenced by: (i) difference between the model dynamics and process dynamics and, (ii) various approximations of the nonlinear plant dynamics used in nonlinear Kalman filters. Currently, there exists no such quantification method to analyze the performance of linear and nonlinear Kalman filters, a key requirement for improvement and a practical benchmark for comparison of these state estimation algorithms. The proposed technique uses the generalized Hurst exponent of the prediction errors (difference in measured output and a posteriori estimates) obtained from the state estimators to quantify the performance. This technique could be implemented on-line as it requires only plant operating data and the predicted outputs (from the linear and nonlinear Kalman filters) to assess the performance. Several simulation studies demonstrate the applicability of the proposed performance metric to both linear and non-linear Kalman filters.

I. INTRODUCTION

State estimation is a crucial task for all industries where there is large number of variables, internal to the system, whose knowledge is required for efficient control and fault diagnosis & monitoring. Accurate knowledge of the system state is necessary for knowing the margins of operating limits and safety of the overall plant. The celebrated Kalman Filter and its variants are largely used for the purpose of state estimation. Though these estimators produce highly reliable estimates, the filters themselves depend heavily on the model of the plant. When there is Model Plant Mismatch (MPM) [1] (i.e., a mismatch between plant and model dynamics), these filter provide poor estimates, or at times, divergent estimates, giving incorrect information to the plant operator. In practice such a mismatch can occur mainly due to: (i) changes in plant dynamics over time, a phenomenon primarily because of the change in plant operating conditions motivated by process economics, (ii) approximations made in various non-linear Kalman filters (usually in the covariance propagation step) [2] and, (iii) wear and tear of the process equipments. This MPM problem has been identified in several applications, such as the Model Predictive Controller [3], [4] and is an inevitable problem to the reliability of all model based state estimators.

Several techniques have been proposed to avoid mismatch between plant and model by careful selection of parameters and by suitably incorporating unmodelled effects in the noise covariance matrices [2]. There are also techniques such as innovation sequence test [5] that have been proposed to detect MPM from plant operating data and filter outputs. Such techniques use the properties of the optimal filtering dynamics to detect deviation from optimality. However, the applicability of these tests has been questioned [6], when the system matrix is not exactly known, which is precisely the problem in case of MPM. Moreover, mere detection of a mismatch will not help the plant operator, except for informative purposes. Quantification of the mismatch, on the other hand, would give an idea of the extent of deterioration of filter performance, and help the operator to take suitable actions. Recently, the application of scaled Hurst exponent as a measure of performance for linear control loops has been proposed in [7]. It was shown that the Hurst exponent based linear control loop measure compares favourably with the other existing linear control loop performance measures. The current work explores the use of generalized Hurst exponent for developing a quantification methodology to assess the performance of linear and nonlinear Kalman filter based state estimation techniques.

This paper is organised as follows: Section II introduces the framework of analysis and defines the problem statement. The effect of model plant mismatch on filter performance is also introduced. Section III introduces the Hurst exponent computed using Detrended Fluctuation analysis and describes the proposed methodology. Section IV provides results and discussion of the simulation studies performed on various benchmark systems (used in other articles). The article ends with a few concluding remarks along with the scope for future work in Section V.

II. PRELIMINARIES

Extensive research on linear as well as non-linear state estimation has resulted into an exhaustive inventory of model based state estimators. The linear Kalman filter proposed by Kalman [8] in 1960 set the stage for model based filtering and state estimation algorithms. It was extended to non-linear systems, resulting in the Extended Kalman Filter (EKF). Several other filters [9], [10], [11] have been proposed in the literature, which are summarised in Figure 1.
This article deals with the performance analysis of linear Kalman filter, extended & unscented Kalman filters (nonlinear state estimators) under the presence of model-plant mismatch. For linear systems with Gaussian noise corrupting the process and measurements, the linear Kalman filter is the best linear unbiased estimator (BLUE) [8]. For the class of non-linear systems, on the other hand, there is no theoretically optimal filter that can handle all kinds of non-linearity. However, the non-linear versions of the Kalman filter such as the Extended Kalman Filter (EKF) and the Unscented Kalman Filter (UKF) [12] are very commonly used for non-linear state estimation. A brief account of these three filtering algorithms on sampled discrete time systems along with the framework of analysis is given below.

Consider a continuous time system with measurements taken at discrete time instants. The discrete state space representation of such a system is as follows:

\[
\begin{align*}
x_{k+1} &= f(x_k, u_k, k) + v_k & (1) \\
y_{k+1} &= h(x_{k+1}, u_{k+1}, k+1) + w_{k+1} & (2)
\end{align*}
\]

where, \( x \in \mathbb{R}^n, u \in \mathbb{R}^p, y \in \mathbb{R}^m \) are the states, inputs and outputs of the plant respectively; \( v \in \mathbb{R}^n \) and \( w \in \mathbb{R}^m \) are the process and observation noises that corrupt the system respectively. The noises \( v(k) \sim \mathcal{N}(0, Q), w(k) \sim \mathcal{N}(0, R) \) are assumed to be identically and independently distributed white noises with zero mean and covariances \( Q \) and \( R \) respectively. In case of linear systems, the functions \( f \) and \( h \) are linear, and the above equations reduce to the following:

\[
\begin{align*}
x_{k+1} &= Ax_k + Bu_k + v_k & (3) \\
y_{k+1} &= Cx_{k+1} + Du_{k+1} + w_{k+1} & (4)
\end{align*}
\]

A. Kalman Filter based State Estimation Techniques

The linear Kalman filter operates in two steps - prediction and update/correction. In the prediction step, it uses the model and current estimate to predict the future estimate, called the \textit{a priori} estimate, \( \hat{x}_{k+1|k} \). In the update/correction step, after the measurement is available, the predicted value and the plant output measurement are suitably weighed to obtain the \textit{a posteriori} estimate, \( \hat{x}_{k+1|k+1} \). The weighing is done so as to minimise the error covariance of the state estimates. This predictor-corrector framework is followed by all Kalman filter based state estimation techniques.

The linear Kalman filter requires that the plant dynamics be completely linear in addition to the noise assumptions mentioned above. The linearity assumption prevents the application of this algorithm to non-linear systems. For such systems, (EKF) is one of the commonly used techniques [13]. EKF algorithm uses a first order Taylor series expansion of the non-linear functions \( f \) and \( h \) (Jacobian matrices) for approximating the error covariance matrices in the update step. When the non-linearity of the system is very high, the linearisation of the model introduces large losses leading to poor performance of EKF. Higher order approximations require the calculation of the Hessian matrices at every iteration leading to increased computational complexity and are therefore not preferred.

One of the popularly used nonlinear Kalman filters called Unscented Kalman Filter (UKF) is based on the idea that it is easier to approximate a probability distribution function than an arbitrary non-linear function [9]. To that end, the UKF approximates the probability distribution function of the state estimates. It generates a set of sigma points and uses the non-linear model to obtain an estimate as a weighted sum of the individual outputs. The weights can be tuned so as to capture higher order moments of the probability distribution function of the state estimates. This approach makes errors in estimating the mean and covariance at the fourth and higher orders in the Taylor series [9] for a continuous function. The EKF, on the other hand estimates the mean only upto the first order and the covariance upto third order. Given the comparable computational expenses of both the techniques, the UKF is expected to have a better performance than the EKF.

B. Effect of Model Plant Mismatch on State Estimator Performance

In Kalman filter based state estimation techniques, it is assumed that the model used in the filter is the best approximation of the plant dynamics (shown in functions in equations 1 and 2) and ideally simulates the plant behaviour. The filter uses these simulated data and measurement data with noise to produce estimates that are closer to the true values. Under these conditions, the prediction errors (difference in the measured output and textita posteriori estimates) from the filter are completely random, with no predictable component. This signifies that whatever component of the output that the filter is not able to predict using a posterior estimates is truly random & unpredictable, therefore indicating the optimality of the filter.

When there is MPM, the plant dynamics no longer satisfy equations 1 and 2. However, the model used in the filter still remains the same without accounting for changes in the plant. So, with time, the model falls short of the plant by an amount \( (\delta f, \delta h) \), depending upon how much the plant has deviated. It is the task of quantifying \( \delta f \) and \( \delta h \) that are addressed in this work, under the assumption that the initial estimates of the system are correct, though the assumption is not necessary for the proposed technique to measure the performance of the filter. Since the model is
no more representative of the plant, the prediction errors (as defined earlier) have a predictable component (time series for which a parametric model of type AR (Auto-Regressive) or MA (Moving Average) could be used to predict) without being truly random. This key fact is used subsequently for development of performance metric for Kalman filter based state estimators. Notice that in this work, it assumed that reasonable initial estimates of the system are available, failing which will result in performance degradation of state estimators and presence of predictable component in prediction errors.

In signal processing applications, the Auto-Correlation Function (ACF) is used as a tool for detecting the presence of repeating patterns in a signal [17]. It is well known that for a true random sequence, the ACF has a value only at zero lag while this is not obeyed for a time series which could be predicted using AR or MA model and their variants. However, in order for this approach to be useful, it should be able to provide quantitative information regarding the MPM to the plant operator. With ACF being a series of numbers, it will be difficult to obtain a single quantitative measure of performance. So, we propose to use the scaling exponent of the prediction errors as an index to quantify MPM.

III. THE PROPOSED METHODOLOGY

The on-line performance assessment of control loops was recently addressed in [7]. The quantification of model plant mismatch in this paper is done along similar lines. A brief overview of the scaling exponent is provided followed by the proposed approach for quantifying MPM.

A. The Scaling Exponent

The scaling exponent is a measure used in time series analysis to quantify the long term correlations in time series. It is used in several applications such as river data analysis for dam construction [14], fractal analysis [17], and biological signal processing [18]. The scaling exponent $\alpha$ of a time series $y(t)$ calculated using $y(t) \equiv a^\alpha y \left( \frac{t}{a} \right)$ is an indicative of the statistical self-similarity of the time series. A small unit of actual series is expanded in time and scaled along the magnitude by an appropriate factor ($a^\alpha$) and the statistical properties of the resultant time series are compared with those of the original time series. Typically for a stationary time series, the value of $\alpha$ ranges from 0 to 1. A value less than 0.5 indicates the series is anti-persistent, i.e., a high (low) value at an instant is more likely to be followed by a low (high) value. If the value of $\alpha$ is between 0.5 and 1, then it signifies a persistent time series, i.e., a high (low) value at an instant is more likely to be followed by a higher (lower) value. Both these types of series can be predicted by using a suitable AR and/or MA model. For a random series, its value is 0.5. So, a value of $\alpha$ other than 0.5 will indicate predictability.

Although the rescaled range analysis was originally proposed by Hurst [14], there are several other techniques [15],[16],[17] which have been proposed for estimation of this exponent. Of these, the Detrended Fluctuation Analysis (DFA) [18] is one of the most commonly used methods. The DFA algorithm is as follows:

1) The given time series $y(k)$ is integrated after mean removal: $Y(k) = \sum_{i=1}^{k} (y(k) - \langle y \rangle)$.

2) $Y(k)$ is then divided into windows of length $n$ and in each window, a least squares line $\hat{y}$ is fit.

3) The fluctuation is then defined as:

$$F(n) = \left( \frac{1}{n} \sum_{i=1}^{n} (Y(i) - \hat{y}_i)^2 \right)$$

This is a function of the window size $n$.

4) Steps 2 and 3 are repeated for varying window sizes $n$.

5) A log-log plot of the fluctuation with the corresponding window size is made, and the slope of this plot is calculated, which gives the estimate of $\alpha$.

The technique gets its name from the removal of trends in each window before the calculation of the fluctuation. This technique could also be used for estimation of scaling exponents for certain type of non-stationary time series [19].

B. Proposed approach to quantify MPM

As mentioned in subsection II-B, the existence of predictable component in the prediction errors of the filter is key to the detection of MPM. The proposed approach to quantify MPM is as follows:

1) Compute the prediction errors $\epsilon(t)$, difference between the measured outputs from the sensor and the corresponding a posteriori estimates from the filter.

2) Estimate the scaling exponent ($\alpha$) of the time series $\epsilon(t)$ using DFA.

3) Under the assumption that the initial estimates and covariance matrices are reasonably correct, ideally with no MPM the $\alpha$ value of $\epsilon(t)$ is expected to be 0.5. Deviations of $\alpha$ values from 0.5 indicate the presence of MPM, with larger deviations indicating higher MPM. Thus, deviations from $\alpha = 0.5$ serve as a metric for quantifying MPM, in turn a measure of performance of Kalman filter based state estimation techniques.

A few noteworthy features of the method are listed below:

- As mentioned earlier, the model used in the filter is only the best approximation of the actual plant dynamics. So, in practice, one would never expect the value of $\alpha$ to be exactly 0.5. This is where one can bring in process information and set suitable thresholds $\alpha_l$ and $\alpha_u$ on the value of $\alpha$. Any deviation of the estimated scaling exponents from these thresholds ($\alpha - \alpha_l$ or $\alpha - \alpha_u$) can be used as an index indicating the extent of MPM.

- An important feature of this technique is that it requires only the plant output $y(k)$ and the a posteriori estimate of the filter $\hat{y}_{k|k}$ to detect and quantify model plant mismatch. This allows for on-line performance monitoring of the model based state estimators.

The proposed technique is demonstrated using both linear and non-linear systems in the next section.
IV. RESULTS AND DISCUSSION

Performance analysis of linear Kalman filter, the best unbiased estimator for all linear systems is first discussed. The demonstration of the proposed metric on EKF and UKF is considered next using non-linear systems. Without loss of generality, the free response of the system is considered. In each of the cases, the model fed to the filter was modified, which has the same effect as changing the plant dynamics. Such modifications were introduced in steps with increasing order of mismatch.

A. Linear System

With linear plant dynamics, equations 1 and 2 simplify to 3 and 4. Let us consider a spring mass damper (SMD) system with one mass, one spring and one damper, governed by the following equations:

\[ x' = \begin{bmatrix} x_1' \\ x_2' \end{bmatrix} = \begin{bmatrix} k/\Delta \end{bmatrix} x - \begin{bmatrix} b/\Delta \end{bmatrix} x + \begin{bmatrix} 0 \\ 1 \end{bmatrix} y \]

where \( m = 1.8kg \), the spring stiffness was \( k = 0.25Nm^{-1} \), and the damping coefficient was \( b = 0.15Nm^{-1}sec^{-1} \). A sampling frequency of 4 Hz was used. Noise was added to the process and observation equations for plant simulation, with covariances \( Q = blkdiag(0.005, 0.005) \) (blkdiag refers to block diagonal matrix) and \( R = 0.005 \). The system was initialized with true initial states \( x_0 = [3, 5]^T \) and error covariance \( P_0 = blkdiag(5, 5) \). To introduce mismatch, the mass, spring stiffness and the damping coefficient fed to the filter were changed. Two such sets of deviation were used. The scaling exponents of the filter errors were estimated for all the cases. The results are summarised in Table I. It can be seen clearly that with increasing MPM, the deviation of \( \alpha \) from 0.5 (for white noise) is higher indicating poor performance of the filter. This also compares favourably with the MSE (mean square squared error between the measured output and a posteriori estimate), a quantity which is minimized in Kalman based state estimators. The tracking of the filter for the cases of no mismatch and highest mismatch are shown in Figure 2, indicating the deterioration in tracking due to presence of MPM.

<table>
<thead>
<tr>
<th>Mismatch</th>
<th>MSE ( \alpha )</th>
</tr>
</thead>
<tbody>
<tr>
<td>No MPM</td>
<td>0.0013 0.5154</td>
</tr>
<tr>
<td>Small MPM</td>
<td>0.0018 0.6374</td>
</tr>
<tr>
<td>Large MPM</td>
<td>0.0035 0.7130</td>
</tr>
</tbody>
</table>

B. Non-Linear System

The first non-linear system considered is a batch reactor process with two gas phase reversible reactions, originally published in [20] and recently used in [2] for analysing and preventing failure of EKF by proper selection of parameters and covariance matrices. The states are the concentrations of the gases while the measurement is the total pressure in the chamber. The governing equations of the system are:

\[
\begin{align*}
\dot{x}_1 &= -k_1 x_1 + k_2 x_2 x_3 \\
\dot{x}_2 &= k_1 x_1 - k_2 x_2 x_3 - 2k_3 x_2^2 + 2k_4 x_3 \\
\dot{x}_3 &= k_1 x_1 - k_2 x_2 x_3 + k_3 x_2^2 - k_4 x_3 \\
y &= RT (x_1 + x_2 + x_3)
\end{align*}
\]

The system was initialized with true initial states of \( x_0 = [0.5, 0.05, 0]^T \) and state error covariance \( P_0 = blkdiag(0.5^2, 0.05^2, 4^2) \). The true parameters \( k_1 = 0.5, k_2 = 0.05, k_3 = 0.2, k_4 = 0.01, RT = 32.84 \) were used in the simulations. Noise was added to the system with covariances \( Q = blkdiag(0.001^2, 0.001^2, 0.001^2) \) and \( R = 0.25^2 \). A sampling frequency of 4 Hz was used. The same parameters were fed to both extended and unscented Kalman filters. The reaction rate constants were changed to introduce mismatch. The results are summarised in Table II. Notice that when there is no MPM, as expected UKF performs better compared to EKF based on the MSE and \( \alpha \) values. However, in this case, it turns out that UKF is much sensitive to MPM compared to that of EKF in terms of MSE and \( \alpha \) values (Table II). Thus, the proposed approach could possibly help identify a better state estimator for a process under various levels of MPM. The tracking of both EKF and UKF for the cases of no and highest mismatch are shown and compared in Figure 3.

<table>
<thead>
<tr>
<th>Mismatch</th>
<th>MSE EKF</th>
<th>MSE UKF</th>
<th>( \alpha ) EKF</th>
<th>( \alpha ) UKF</th>
</tr>
</thead>
<tbody>
<tr>
<td>No MPM</td>
<td>0.0435</td>
<td>0.0418</td>
<td>0.5487</td>
<td>0.3220</td>
</tr>
<tr>
<td>Small MPM</td>
<td>0.0460</td>
<td>0.0483</td>
<td>0.5984</td>
<td>0.7114</td>
</tr>
<tr>
<td>Large MPM</td>
<td>0.0498</td>
<td>0.0518</td>
<td>0.6630</td>
<td>0.7988</td>
</tr>
</tbody>
</table>

The second system is again a gas phase reversible reaction taken from [20]. The states are the partial pressures of the reactant and product while the measurement is the total pressure. The system is governed by the following equations:

\[
\begin{align*}
\dot{x}_1 &= -2k x_1^2 \\
\dot{x}_2 &= k x_1^2 \\
y &= x_1 + x_2
\end{align*}
\]

The true initial states of the system (used for initialization) were \( x_0 = [3, 1]^T \) with \( k = 0.16 \) and the state error covariance matrix is \( P_0 = blkdiag(4, 4) \). Noise added to the process had covariance matrices \( Q = blkdiag(0.001^2, 0.001^2) \) and \( R = 0.02^2 \). A sampling frequency of 100 Hz was used. The same parameters were fed to both the state estimators. Mismatch was introduced by manipulating the reaction rates in the model. The parameter \( k \) was varied to result in a deviation from plant dynamics. The results are summarised in Table III. The tracking of both the filters for the cases of no and highest mismatch are shown and compared in Figure 4.
**Fig. 2.** Performance of Linear Kalman Filter with and without MPM (a) Output of plant and filter, i.e., distance of the mass (same as the first state of the system) (b) Second state of the plant and filter, i.e., velocity of the mass

**Fig. 3.** Performance of EKF and UKF with and without MPM for the Batch reactor process governed by equation 6. (a) Output of plant and filters, i.e., total pressure (b) Concentration of species A (c) Concentration of species B (d) Concentration of species C

**TABLE III**

<table>
<thead>
<tr>
<th>Mismatch</th>
<th>MSE</th>
<th>MEK</th>
<th>UKF</th>
<th>$\alpha$</th>
</tr>
</thead>
<tbody>
<tr>
<td>No MPM</td>
<td>$3.75 \times 10^{-4}$</td>
<td>$3.60 \times 10^{-4}$</td>
<td>0.5496</td>
<td>0.4779</td>
</tr>
<tr>
<td>Small MPM</td>
<td>$3.76 \times 10^{-4}$</td>
<td>$3.63 \times 10^{-4}$</td>
<td>0.5608</td>
<td>0.5476</td>
</tr>
<tr>
<td>Large MPM</td>
<td>$3.82 \times 10^{-4}$</td>
<td>$3.78 \times 10^{-4}$</td>
<td>0.5854</td>
<td>0.7916</td>
</tr>
</tbody>
</table>

**C. Key outcomes from proposed methodology**

As expected, introducing mismatch into a model based state estimator affects its performance, which reflects in the prediction errors. This is true for linear and nonlinear Kalman filters. The predictability property in the prediction errors is exploited to detect as well as quantify mismatch using the scaling exponent. It is noteworthy that upon increasing the mismatch in the model, the performance degrades further and the scaling exponents also consistently deviate further away from 0.5 (value for white noise). This can be clearly seen from the simulation results summarized in Tables I, II and III. Further, the mean squared error values also compare favourably with the trends of the scaled exponent. Since the proposed method only depends on the routine operating data, it is easy to do on-line performance assessment and monitoring of the state estimation algorithms.

Another important outcome of the proposed method lies in the comparison of various nonlinear state estimators under different levels of MPM, since there is no optimal filter for such systems. For example, in the systems used in this work, the UKF performed better than the EKF for the case of no mismatch, but when mismatch was introduced, in one of the cases EKF performed better than the UKF. This feature is specific to the systems considered in this work and the proposed methodology can be used to assess and compare
the performance of the Kalman filter based state estimation algorithms for given non-linear systems.

V. CONCLUSIONS AND FUTURE WORK

A new technique was proposed for assessing the performance of Kalman filter based state estimation approaches. The proposed metric computes the scaling exponent of the output errors to determine the predictability component in the output error data. The plant operating data and the model are the only information needed for the technique. This allows for on-line performance assessment as well as monitoring. Another important application is the comparison of the performance of various model based state estimation algorithms for non-linear systems. Our future work is focused towards extending the technique to various other constrained non-linear state estimators. Further, the possibility of improving the state estimates, after having detected and quantified the mismatch shall also be explored.

REFERENCES
