Defence Science Journal, Vol. 60, No. 1, January 2010, pp. 87-99

Handling Out-of-Sequence Data: Kalman Filter Methods or Statistical Imputation?

Bhekisipho Twala

Council of Scientific Industrial Research (CSIR), PO Box 395, Pretoria 0001, South Africa
E-mail: btwala1@csir.co.za, bhekisipho.twala@brunel.ac.uk

ABSTRACT

The issue of handling sensor measurements data over single and multiple lag delays also known as out-of-sequence measurement (OOSM) has been considered. It is argued that this problem can also be addressed using model-based imputation strategies and their application in comparison to Kalman filter (KF)-based approaches for a multi-sensor tracking prediction problem has also been demonstrated. The effectiveness of two model-based imputation procedures against five OOSM methods was investigated in Monte Carlo simulation experiments. The delayed measurements were either incorporated (or fused) at the time these were finally available (using OOSM methods) or imputed in a random way with higher probability of delays for multiple lags and lower probability of delays for a single lag (using single or multiple imputation). For single lag, estimates of target tracking computed from the observed data and those based on a data set in which the delayed measurements were imputed were equally unbiased; however, the KF estimates obtained using the Bayesian framework (BF-KF) were more precise. When the measurements were delayed in a multiple lag fashion, there were significant differences in bias or precision between multiple imputation (MI) and OOSM methods, with the former exhibiting a superior performance at nearly all levels of probability of measurement delay and range of manoeuvring indices. Researchers working on sensor data are encouraged to take advantage of software to implement delayed measurements using MI, as estimates of tracking are more precise and less biased in the presence of delayed multi-sensor data than those derived from an observed data analysis approach.

Keywords: Multi-sensor data, time delayed measurements, out-of-sequence measurements, Kalman filter, fusion, imputation, multi-sensor tracking

1. INTRODUCTION

Recent years have witnessed the emergence of several approaches to sensing applications. Many applications require the maintenance of a high fidelity estimate of the state of a dynamic system based on a sequence of noisy observations. Such applications demand the use of filtering mechanisms such as the KF[1,2], to fit the observation sequence to a given model of the system dynamics. The primary motivation for this study came from problems that arise in multiple target tracking with distributed sensor networks. One central problem in multiple target tracking is out-of-scope measurements (OOSM)[3-14], which is sometimes referred to as the problem of tracking with random sampling delays[15-17] and the problem of incorporating time delayed measurements[18].

Most of the work on tracking and filtering has been based on the assumption that measurements are available immediately to an agent. However, it is not difficult to conceive situations in which measurements are subject to non-negligible delays, such that the lag between measurement and receipt is of sufficient magnitude to have an impact on estimation or prediction[19]. This can be caused by a number of specific reasons:

- Communication delays from the sensor to the tracker;
- Different sensors observing the current state of the target at different times;
- Delays in sending tracks to the data fusion node (often because the sensor is a rotating radar measurement with specific time stamps) and
- Unsteady pre-processing times of the observed data, depending on the system load, can vary from one measurement to another. Delayed measurements can create difficulties, especially for discrete time filtering.

Delayed measurements can be classified into two categories, constant delays, and random delays. Constant delays involve measurements being delayed by the same constant lag. In this way, measurements are never observed out-of-sequence, these are simple and consistent. Such behaviour could be induced, for example, by a constant bandwidth restriction on a sensor network. In contrast, random delays provide a number of possibilities, including that measurements are delayed with a constant probability but with a fixed lag, or constant probability with a random lag. Such problems could arise as a result of intermittent bandwidth restrictions on a sensor network. All models of random delay have the potential to cause OOSM. This study was inclined in solving the problem of random delays, hence, OOSM.

Another problem related to OOSM is that of incomplete data or missing values. In fact, skipping the correction

Received 22 September 2009

step and proceeding straight to the prediction step when using a KF for estimations purposes is the equivalent of considering the delayed measurement as missing[20-22]. The presence of missing values is commonplace in large real-world databases. This has become one of the most important problems in academic research since most learning systems and statistical analyses in the early stages were not designed to handle missing data (incomplete vectors). There are several reasons why there are missing values in data. An item could be missing because it was unavailable or arises by default in data recording activities. Missing values could also occur because of confusing questions in the data gathering or because of sensor malfunction. In some situations, the missingness could be caused by the relationships between the attribute variables themselves. That is, the information that is missing on a given attribute variable could be as a result of its relation to values of other attribute variables in the data set. An extreme case is that the missing value could be as result of its relation to an unobserved (missing value) in the data set.

In this paper, the resilience of OOSM, on the one hand, and imputation, on the other hand, to the various forms of imperfections in sensor data has been explored to increase the awareness of the impact delayed measurements could have, when building robotic prediction models. The main focus is on the application of OOSM and MI to multi-target tracking prediction.

This study is significant because of the following reasons:

- OOSM is an emerging technology that can aid in the handling of delayed measurements in single- or multi-target tracking predictions
- The delayed measurements problem is related to the incomplete data problem, hence, the use of imputation procedures (single or multiple imputation) could be utilised;
- MI has an advantage over single imputation strategies due to the fact that it overcomes the under-representation of uncertainty about which value to impute (i.e., single imputation methods underestimate the true variance of the values these are attempting to fill-in or impute);
- Due to the lack of adequate tools to deal with delayed measurements, machine learning techniques have been used to tackle such problems, including either single- or multiple-target tracking or navigation prediction.

Most of the work on tracking and filtering has been based on the assumption that measurements are immediately available to an agent. However it is not difficult to conceive situations in which measurements are subject to non-negligible delays, such as the lag between measurement and receipt is of sufficient magnitude to impact on estimation. In such situations, the classical assumption, that observations are available immediately, is easily violated[19].

A direct solution to the OOSM problem is simply to ignore and discard the OOSM in the tracking process more like the listwise deletion is a standard default approach for dealing with missing data in most statistical packages. This solution leads obviously to a loss of the information contained in the discarded OOSM. To avoid this drawback, several alternative methods proposed are available in the literature to deal with the OOSM problem, especially for random delays. It is also striking that almost all of the methods proposed to handle delayed measurements have in common that delayed measurements are always ultimately incorporated into the filtering process.

In the time delays context, one common approach is related to solving a partial differential equation and boundary condition equations which do not have an explicit solution in general[23-27].

For the case of discrete time systems (and especially for random delays), the problem has been investigated via a standard Kalman filtering[28] and by augmenting the system accordingly[28-30]. Matvev and Savkin[31] consider an iterative form of state augmentation for random delays with a random lag. Larsen[32], *et al.* address the OOSM problem by recalculating the filter through the delayed period. In the same context, Larsen[32], *et al.* propose a measurement extrapolation approximation using past and present estimates of the KF and calculating an optimal gain for this extrapolated measurement. Thomopoulos and Zhang[17] examine the case of random delay under the name of the fixed sampling and random delay filter, that is shown to be equivalent to constraining the lag to a value of 1. Alexander[1] and later Larsen[32], *et al.* suggest using the delayed measurements to calculate a correction term and adding this to the filter estimate. Zhang[34], *et al.* proposed algorithms that try to minimise the information storage in an OOSM situation. Challa[9], *et al.* formulated the OOSM problem in a Bayesian framework. The above methods are described in more detail in Section 3.

2. PROBLEM STATEMENT

The presentation is based on the KF equations for a discrete linearised time-varying system with state vector x_k, input vector u_k, and output vector y_k. KF is the optimal recursive data processing algorithm for a discrete linear system corrupted with noise in the states and measurements. What a KF requires is knowledge of the system and measurement dynamics, a statistical description of the system and measurement noises, uncertainty in the dynamic models and any available information about the initial conditions of the variable of interest. Based on this knowledge, it gives the optimal estimate of the state variables under observations[1,35]. Since its inception KF has become a subject of extensive research and application, particularly in areas of autonomous, assisted navigation or target tracking. For more detailed study of KFs probabilistic origin see Maybeck[36].

The following lists the equations for notation purposes[37] and improved by Bar-Shalom[38], *et al.* and later by Julier and Uhlman[39].

2.1 System Description

The KF addresses the general problem of trying to estimate the state $x \in \Re^n$ of a discrete-time controlled process that is assumed to evolve over time t_{k-1} to t_k and governed by the linear stochastic difference equation

$$x(k) = F(k, k-1)x(k-1) + v(k, k-1) \quad (1)$$

where, $x(k)$ is the state vector at time k, $F(k,k-1)$ is the state transition matrix to time t_k from t_k-1 and $v(k,k-1)$ represents the (cumulative effect of the) process noise for this interval. The order of the arguments in both F and v is according to the convention for the transition matrices. Typically, the process noise has a single argument, but here the two arguments will be needed for clarity. The time τ, at which the OOSM was made, is assumed to be such that

$$t_{k-l} < \tau < t_{k-l+1} \quad (2)$$

This will require the evaluation effect of the process noise over an arbitrary non-integer number of sampling intervals. Note that l=1 corresponds to the case where the lag is a fraction of a sampling interval; for simplicity this is called the "1-step-lag" problem, even though the lag is really a fraction of a time step.

The measurement $z \in \Re^m$ and thus measurement or observational model is

$$z(k) = H(k)x(k) + w(k) \quad (3)$$

where z(k) is the observation vector, $w(k)$ is the observation noise vector and is the observation matrix. The noise vector $v(k,K-1)$ and $w(k)$ are assumed to be independent (of each other), white, and with normal probability distributions

$$p(w) \sim N(0, Q) \quad (4)$$

$$p(v) \sim N(0, R) \quad (5)$$

The process noise covariance Q(k) and measurement noise covariance $R(k)$ are mutually uncorrelated and they are given as

$$E[v(k,j)v(k,j)'] = Q(k,j) \; E[w(k)w(k)] = R(k) \quad (6)$$

Similarly to Eqn (1), one has

$$x(k) = F(k, \kappa)[x(\kappa) - v(k, \kappa)] \quad (7)$$

where, κ is the discrete time notation for τ. The above can be written backward as

$$x(\kappa) = F(\kappa, k)[x(k) - v(k, \kappa)] \quad (8)$$

where, $F(\kappa, k) = F(k, \kappa)^{-1}$ is the backward transition matrix.

2.2 Fusion of Time Delayed Measurements

Denoting a cumulative set of measurements $Z^k \triangleq \{z(i)\}_{i=1}^k$, the OOSM problem [up to time instance $t=t_k$, and excluding a measurement z(τ) with a time stamp $t_\tau < t_k$ as shown in Fig. 1 to reduces to the problem of computing the conditional mean estimate of the target state

$$\hat{x}(k \mid k) \triangleq E[x(k) \mid Z^k] \quad (9)$$

and its associated error covariance

$$P(k \mid k) \triangleq \mathrm{cov}[x(k) \mid Z^k] \quad (10)$$

Under the assumption that the initial state x_0 is Gaussian, the conditional mean estimate $\hat{x}(k \mid k)$ of the target state, which is optimal in the minimum variance sense, can be computed recursively using the KF. Also, it is assumed that a measurement z is collected and used to update the track at the time interval h. The basic KF algorithm can then be extended to multi-sensor systems where the data is assumed to arrive at known times and in correct time sequence.

Suppose that a given measurement corresponding from time τ (denoted with discrete time notation as κ),

$$z(\kappa) \triangleq z(\tau) = H(\kappa)x(\kappa) + w(k) \quad (11)$$

arrives with a certain delay after Eqns (9) and (10) have been computed, as shown in Fig. 1.

One faces the problem of updating the state estimate and its covariance with the delayed measurements, i.e., to compute

$$\hat{x}(k \mid \kappa) \triangleq E[x(k) \mid Z^k] \quad (12)$$

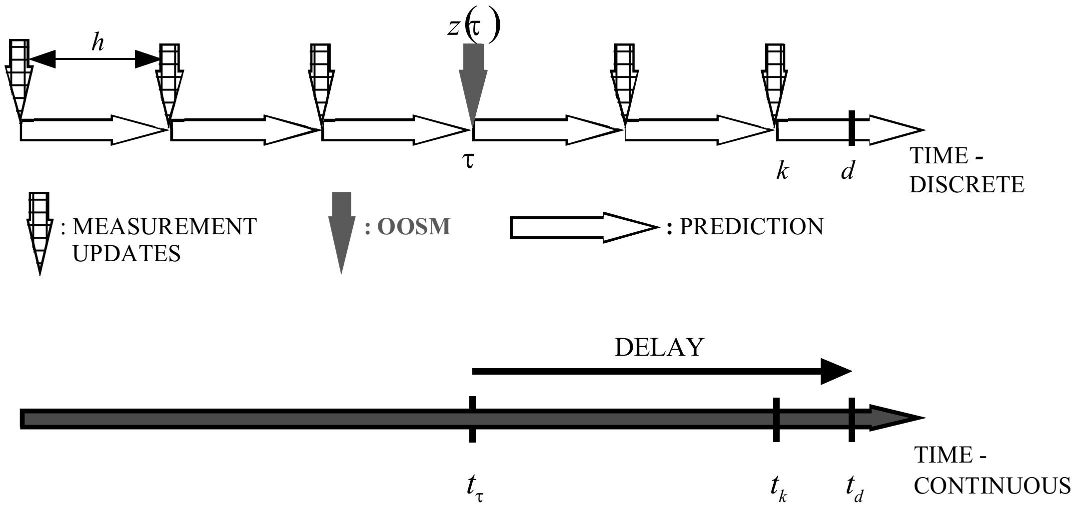


Figure 1. Out-of-sequence measurement.

and

$$P(k \mid \kappa) \triangleq \mathrm{cov}[x(k) \mid Z^{k}] \quad (13)$$

where $Z^{\kappa} \triangleq \{Z^{k}, z(\kappa)\}\}$ (14)

Equation (13) provides a simple, intuitive interpretation of the weight in the time delayed KF. The weight assigned to a measurement is a function of the degree to which the measurement is correlated with the current state of the system*. Therefore, the difficulty in implementing the time delayed KF is in calculating $P(k|\kappa)$. Solutions to the delay measurement problem are presented in Section 3.

3. SOLUTIONS TO DELAY MEASUREMENT PROBLEM

3.1 Out-of-Sequence Measurement

There have been a number of solutions proposed in the literature to solve this OOSM problem. Most existing solutions to the problem are based on retrodiction, where backward prediction of the current estimated state is used to incorporate the OOSMs at appropriate time instants. However, in recent years, some researchers have tackled the OOSMs problem without the need of backward prediction (See for example, Rhéaume[40], *et al.*). For the purpose of this study, the authors are more interested in the backward prediction solutions, which are described below.

Thomopoulos and Zhang[17] examined the case of random delays of the complete measurement vector which arrive out-of-sequence, where the lag was restricted to a value of 1. The measurements arrive at the fusion with random delays which can be due to queuing at the sensor buffer and to delays in the transmission time as well as in the propagation time. Optimal filters for the estimation of target tracks based on measurements of uncertain origin received by the fusion at random times and out-of-sequence were derived for the cases of random sampling, random delay, and both random sampling and random delay.

Alexander[33] and Larsen[32], *et al.* suggested using the delayed measurement to calculate a correction term and adding this to the filter estimate, again considering the complete observation vector being delayed. In the same context, Larsen[32], *et al.* further proposed a measurement extrapolation approach to ensure the optimality of the filter and at the same time address issues like changing measurement and state noise covariance matrices.

One other approach by Matveev and Savkin[31] that addresses the random delayed measurements problem is called state augmentation. This approach was designed to handle a linear discrete-time partially observed system perturbed by white noises. The reduced-order linear unbiased estimator was designed via iterative state augmentation. By so doing, Matveev and Savkin[31] managed to solve the minimum variance state estimation problem and further showed how their proposed approach was exponentially stable under natural assumptions. Lu[41], *et al.* followed up this approach by proposing a variable dimension filter which handles only essential past states not just past states that are up to some maximum delay like Matveev and Sankin[23] approach.

Julier and Uhlman[39] considered the problem of applying a KF to estimate the state of a dynamic system using a sequence of observations that are not precisely time stamped. They argued that the problem has analogies with the identity ambiguity problem that arises in MTT applications. They further described a way in which multiple hypotheses and covariance union (CU) methods could be utilised for this kind of problem and compared them with the probabilistic data association filter (PDAF) method. Their results showed the PDFA yielding the most accurate results, however, at a higher computational cost. The strong dependence of PDAF on the accuracy of the likelihood model is another of its weakness. Although CU requires the evaluation of two KF updates, its advantage lies in its ability not to rely on specific assumptions as to the veracity of the likelihood model.

More recently, Zhang[34], *et al.* proposed two algorithms with three cases of different information storage for the state estimation update with OOSM. Both algorithms are optimal in the linear minimum mean square error sense for the information available at the time of update. Their proposed algorithms (based on the linear minimum mean-square error) try to minimise the information storage in an OOSM situation using different minimum storage of information concerning the occurrence time of single OOSMs. Further, they extend the single OOSM update algorithms to the case of arbitrarily multiple OOSMs.

Challa[9], *et al.* formulated and solved the OOSM problem in the Bayesian framework. They established that the solution involved the joint probability density of current and past states or the state corresponding to the delayed measurement. For the case of multiple delays, the authors show what the solution involves a Bayesian recursion for the joint probability density of an augmented state vector. Based on this, the augmented state Kalman filter (AS-KF) and its variable dimensions extension (VDAS-FK) have been proposed as the fundamental solution to this problem in the linear Gaussian case. AS-FK handles noise-target state cross-correlation implicitly and can be readily extended to handle clutter. The idea of VDAS-FK is that the augmented state only carries the current state and the past state for which there was a missing measurement. The filter will reduce to a normal KF if there is no OOSM. Further, a new augmented state probabilistic data association filter (AS-PDA) is proposed by Challa[9], *et al.* This filter is meant to deal with data association issues arising from the presence of clutter in the OOSM problem. Simulation results were used to demonstrate the effectiveness of these algorithms. The results show

* The result is algebraically the same as derived by Larsen[32], *et al.*, However, the interpretation is different. Larsen considered taking an observation and extrapolating its value forward to the current time step in the filter. Julier and Uhlmann[39] considered calculating the correlation backwards from the current time to the time when the observation was made.

the proposed solutions as computationally expensive when compared with existing methods but straightforward to implement and these also yield significant improvements in terms of performance.

A more principled way to handle this problem is to extend Challa's Bayesian formalism[9,36] to include uncertainties in the time delays. This is analogous to a problem that arises in multiple target tracking (MTT)[42]. MTT occurs when a tracking system receives an observation of one of several different targets, but the exact identity of the observed target is not known.

3.2 Multiple Imputation

Imputation is the substitution or replacement of some value of a missing data point or missing component of a data point[43-45]. Multiple imputation (MI) is one of the most attractive methods for general purpose handling of missing data in multivariate analysis described MI as a three-step process. First, sets of *M* plausible values (M=5 in Fig.1) for missing instances were created using an appropriate model that reflects the uncertainty due to the missing data. Each of these sets of plausible values was used to fill-in the missing values and create *M* complete datasets (imputation). Second, each of these *M* datasets can be analysed using complete-data methods (analysis). Finally, the results from the *M* complete datasets are combined, which also allowed the uncertainty regarding the imputation to be taken into account (pooling or combining).

For example, replacing each missing value with a set of five plausible values or imputations (as it was the case in our illustration in Fig. 2) would result to building five decision trees (DTs)[46], and the predictions of the five trees would be averaged into a single tree, i.e., the average tree is obtained by MI. MI retains most of the advantages of single imputation and rectifies its major disadvantages as already discussed.

There are various ways to generate imputations. Schafer[45] has written a set of general purpose programs for MI of continuous multivariate data (NORM), multivariate categorical data (CAT), mixed categorical and continuous (MIX), and multivariate panel or clustered data (PNA). These programs were initially created as functions operating within the statistical languages S and S-PLUS[47]. NORM includes an expectation maximisation (EM) algorithm for maximum likelihood estimation of means, variance, and covariances. NORM also adds regression-prediction variability using a Bayesian procedure known as data augmentation[48] to iterate between random imputations under a specified set of parameter values and random draws from the posterior distribution of the parameters (given the observed and imputed data). These two steps are iterated long enough for the results to be reliable for multiple imputed datasets[45]. The goal is to have the iterates converge to their stationary distribution and then to simulate an approximately independent draw of the missing values. The algorithm is based on the assumptions that the data come from a multivariate normal distribution and are missing at random (MAR). MAR essentially says that the cause of missing data may be dependent on the observed data but must be independent of the missing value that would have been observed.

Although not absolutely necessary, it is almost always a good idea to run the EM algorithm[50] before attempting to generate MIs. The parameter estimates from EM provide convenient starting values for data augmentation (DA). Moreover, the convergence behaviour of EM provides useful information on the likely convergence behaviour of DA. Therefore, EM estimates of the parameters are computed and then recorded the number of iterations required, say t. Then, a single run of DA algorithm of length tM using the EM estimates as starting values is performed, where M is the number of imputations required. The convergence of the EM algorithm is linear and is determined

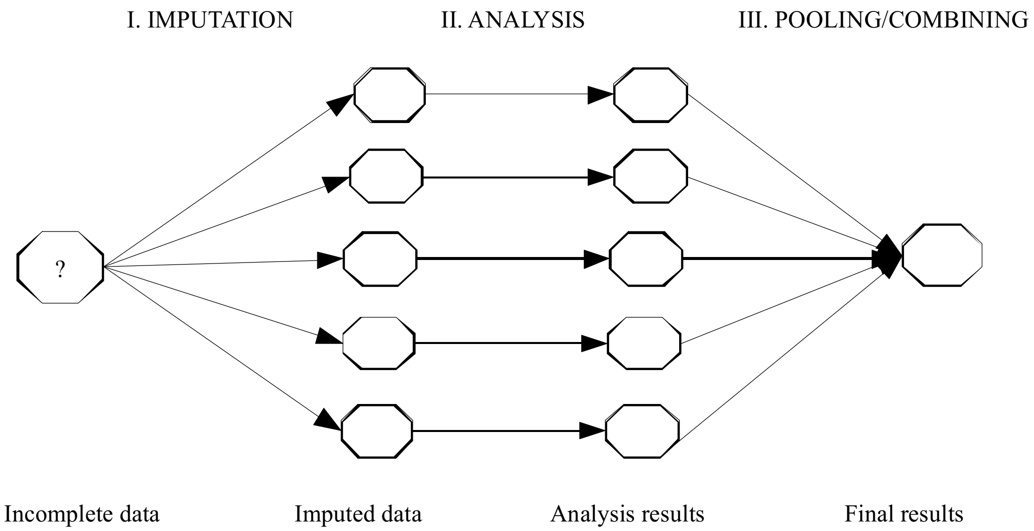


Figure 2. The three steps of multiple imputation.

by the fraction of missing information. Thus, when the fraction of missing information is large, convergence will be very slow due to the number of iterations required. However, for small missing value proportions convergence is obtained much more rapidly with less strenuous convergence criteria. The EM and DA processes are described below.

The core idea of the EM algorithm is to introduce some unobserved variables Z, appropriate for the model under consideration, such that if Z were known the optimal value of θ could be computed easily. Then the complete conditional probability density (including the missing variables) can be written as:

$$L(\theta \mid X, Z) = \sum_{i=1}^{N}\sum_{j=1}^{M} z_{ij}\log f(x_i \mid z_i;\theta) f(z_i;\theta) \tag{15}$$

The usual approach is to regard Z as missing data and estimate it iteratively.

The intuition behind the EM algorithm is that one would like to maximise the complete data likelihood but it cannot be utilised directly, so we maximise its expectation, denoted by $Q(\theta \mid \theta^t)$, instead. As shown by Dempster[47], *et al.*, $L(\theta \mid X)$, the complete data likelihood can be maximised by iterating the following steps:

Step 1. Initialise parameters randomly. Set $t = 0$.

Step 2. *E*-step: Determine

$$Q(\theta \mid \theta^{(t)}) = E[L(\theta \mid X, Z) \mid X, \theta^{(t)}]$$

Step 3. *M*-step: Set $\theta^{(t+1)} = \arg\max_{\theta} \{Q(\theta \mid \theta^{(t)})\}$ where $\theta^{(t)}$ are the current parameter estimates in time step t.

Step 4. Iterate steps 2 and 3 until convergence

Assume that the data set $X = \{x_1, ..., x_N\}$ is divided into an observed X_{obs} and missing X_{miss} components, respectively. To handle missing values one can re-write the EM algorithm as follows:

Step 1. Initialise parameters randomly. Set t = 0

Step 2. *E*-step: Determine

$$Q(\theta \mid \theta^{(t)}) = E[L(\theta \mid X_{obs}, X_{miss}, Z) \mid X_{obs}, \theta^{(t)}]$$

Step 3. *M*-step: Set $\theta^{(t+1)} = \arg\max_{\theta} \{Q(\theta \mid \theta^{(t)})\}$

where $\theta^{(t)}$ are the current parameter estimates in time step *t*.

Step 4. Iterate steps 2 and 3 until convergence.

The expectation (*E*) step computes the expected values for the sufficient statistics given a model and values for model parameters θ, i.e., the expected value of the complete data likelihood wrt the missing data given the observed data and the current parameter estimates. The maximisation (*M*) step estimates the model parameters by maximising the likelihood using standard procedures, given complete data. The procedure iterates through these two steps until convergence is obtained. Convergence occurs when the change in parameter estimates from iteration becomes negligible. An important part of the EM algorithm is restoring error variability to the imputed values during the E-step. Replacing a missing value by an imputed value using the EM algorithm results in EM single imputation (EMSI).

DA (which resembles EM) follows the following process:

Step 1. Initialise parameters randomly. Set $t = 0$

Step 2. I-step: Given a current estimate $\theta^{(t)}$, select a value of the missing data from the conditional predictive distribution of $X_{miss}, X_{miss}^{(t+1)} \sim P(X_{miss} \mid X_{obs}, \theta^{(t)})$

Step 3. P-step: Conditioning on $X_{miss}^{(t+1)}$, draw a new value of θ from its complete data posterior, $\theta^{(t+1)} \sim P(\theta \mid X_{obs}, X_{miss}^{(t+1)})$. Through an iterative process two distributions are obtained, $P(\theta \mid X_{obs})$ and $P(X_{miss} \mid X_{obs})$. For a suitable large t, one can implement a DA algorithm by Tanner and Wong[48], which iterates between sampling θ^{t+1} from $P(\theta \mid X_{obs})$ and sampling $X_{miss}^{(t)}$ from $P(X_{miss} \mid X_{obs})$.

Step 4. Iterate steps 2 and 3 until convergence

The Imputation (I) step simulates a random imputation of missing data under assumed values of the parameters. The Posterior (*P*) step draws new parameters from a Bayesian posterior distribution based on the observed and imputed data. The procedure of alternately simulating data and parameters creates a Markov Chain (MC)[49] $X_{miss}^{(1)}, \theta^{(1)}, X_{miss}^{(2)}, \theta^{(2)}, ...,$ which eventually stabilises or converges in distribution to $P(X_{miss}, \theta \mid X_{obs})$. The procedure iterates through these two steps until convergence is obtained. The rate of convergence is related to the fraction of missing information. DA can be thought of as a small–sample refinement of the EM algorithm using simulation, with the imputation step corresponding to the E-step and the posterior step corresponding to the M-step. This approach has been followed in this paper, which is called Bayesian multiple imputation (BAMI).

MI has several desirable features:

- Introducing appropriate random error term into the imputation process which makes it possible for the method to get approximately unbiased estimates of all parameters,
- Repeated imputation allows one to get good estimates of standard errors;
- MI can be used with any kind of data and any kind of analysis without specialised software,
- MI saves money, since for the same statistical power, MI requires a smaller sample size than, say, listwise or case deletion, and
- Once imputations have been generated by a knowledgeable user, researchers can use them for their own statistical analysis.

However, certain requirements must be met for MI to have these desirable features.

- The data must be MAR.
- The model used to generate the imputed values must be 'correct' in some sense.
- The model used for the analysis must match up, in some sense, with the model used in the imputation.

The reader is referred to Schafer[45] and Allison[51] for a rigorous description of all these conditions.

4. EXPERIMENTAL SETUP

In this section the behaviour of the five proposed OOSM procedures against model-based imputation procedures have been explored. The four methods selected are based on the following KFs:

- Fixed sampling and random delay Kalman filter (FSRD-KF);
- Measurement extrapolation Kalman filter (ME-KF);
- State augmentation for random delays Kalman filter (SARD-KF), minimum storage Kalman filter (MR-KF) and Bayesian framework Kalman filter (BF-KF).
- The tracking performance is characterised by the root mean square error (RMSE) over 1000 Monte Carlo run for each specific scenario.

The root mean square deviation (RMSD) or RMSE is a measure of the differences between values predicted by a model or an estimator and the values actually observed from the thing being modelled or estimated.

The RMSD of an estimator $\hat{\theta}$ wrt the estimated parameter θ is defined as the square root of the mean squared error (MSE):

$$\text{RMSD} = \text{RMSE }(\hat{\theta}) = \sqrt{MSE(\hat{\theta})} = \sqrt{E(\hat{\theta}-\theta)^2} \qquad (16)$$

Following Bar-Shalom[4], two cases (process noise q = 0.1, and 4) corresponding to ë=0.3, and 2 were examined, i.e., the underlying target performs in a straight line motion, or was highly manoeuvring. Data was generated randomly for each run starting with a initial state

$$x(0) = [200 \text{ km}, \quad 0.5 \text{ km/s}, 100 \text{ km/s} \quad -0.08 \text{ km/s}] \qquad (17)$$

A two data point method[41] was used to initialise the filters with

$$P(0\,|\,0) = \begin{pmatrix} P_0 & 0 \\ 0 & P_0 \end{pmatrix} \text{where } P_0 = \begin{pmatrix} R & R/T \\ R/T & 2R/T^2 \end{pmatrix} \qquad (18)$$

for *a priori* error covariance or to form the initial error covariance for augmented state. Like in Challa[50], *et al.*, it is assumed that the OOSM can only have a maximum of one lag delay, and the data delay was uniformly distributed within the whole simulation period with probability P_r that the current measurement was delayed.

All statistical tests were conducted using the MINITAB statistical software program[53]. Analyses of variance, using the general linear model (GLM) procedure[54] were used to examine the main effects and their respective interactions. This was done using a 3-way repeated measures design (where each effect was tested against its interaction with the simulated dataset). The fixed effect factors were: OOSM and imputation methods; the probability of measurement delay; and the manoeuvring index.

4.1 Experiment I

To empirically evaluate the performance of the five OOSM methods and EMSI in terms of RMS error, an experiment on simulated data (as explained in sub-section 4.1) was used. This experiment was carried out to rank individual OOSM methods and also assess the impact of delayed measurements (at various time and distance intervals) on a single delay against single imputation in terms of position error.

Results of Experiment I

Experimental results on the effects of delayed measurements on one lag delay in terms of the RMS position error have been described. The behaviour of these methods has been explored for distance and time intervals. From these experiments the following results have been observed:

Main Effects: All the main effects were found to be significant at the 5 per cent level of significance (F = 37.17, df = 5 for OOSM methods and EMSI; F = 6.195, df=1 for probability of measurement delay; F=9.39, df=1 for manoeuvring index; p < 0.05 for each effect).

From Fig. 3, BF-KF has been found to be the overall best technique for handling delayed measurements on a one lag with an excess error rate of 5.6 per cent, closely followed by EMSI, FSRD-KF and MR-KF, with excess error rates of 6.1 per cent, 8.2 per cent and 8.5 per cent, respectively. The worst technique was SARD-KF, which exhibits an error rate of 9.9 per cent. Tukey's multiple comparison tests showed no significant differences between ME-KF and SARD-KF (on one hand) and FSRD-KF and MR-KF (on other hand). The significance level for all the comparison tests was 0.05. All interaction effects were found to be insignificant at the 5 per cent level of significance. Hence, not discussed. No interaction effects were found to be significant at the 5 per cent level. Hence, not discussed.

Figures 4 and 5 show simulation results where the performance of OOSM and imputation methods for single delay over 1000 runs have been compared. The following observations have been made:

- SARD-KF and ME-KF have similar RMS error performance (on one hand) with FSRD-KF, MR-KF and EMSI achieving similar performances (on other hand). However, the

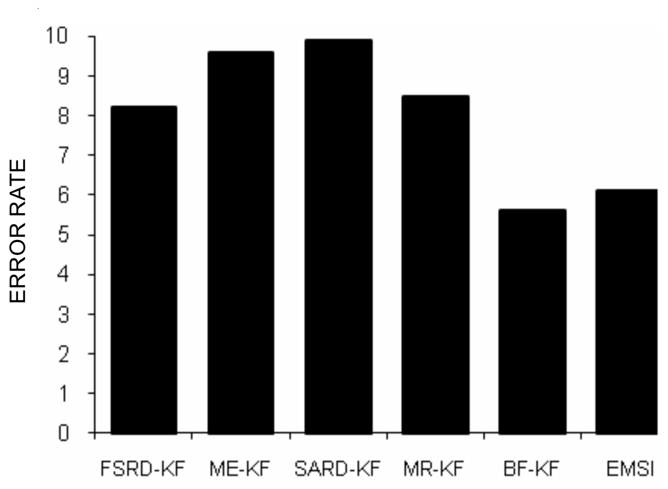


Figure 3. Overall means for OOSM and single imputation methods.

latter methods always achieve lower RMS error rates. BF-KF achieves higher accuracy rates at all time levels. This was the case for non-manoeuvring target tracking (Fig. 4).

- For manoeuvring target tracking, BF-KF (once again) outperforms all the methods with serious competition from EMSI. The differences in performance are mostly prominent at higher probabilities of measurement (Fig. 5).
- For both types of manoeuvres, increases in probability of measurement delay (P_r) was associated with increases in performance differences between methods. In fact, the performance by all the methods degrades with increases in probability of measurement.
- The accuracy of BF-KF and EMSI was achieved at a higher computational cost in terms of minutes (Table 1). Both methods take about twice (in some situations thrice) to compute compared to the others.

4.2 Experiment II

The main objective of this experiment was to compare the performance of OOSM and imputation methods for multiple delays, especially the top two OOSM methods that exhibited higher accuracy rates in the previous experiment. These are FSRD-KF and BF-KF. Also, the authors presumed that it would be interesting to test the effectiveness of multiple imputation (a procedure for handling incomplete data) against methods what have been proposed to deal with the delay measurement problem.

Results of Experiment II

Main Effects: All the main effects were found to be significant at the 5 per cent level of significance (F = 54.8, df = 2 for OOSM and multiple imputations methods; F = 11.62, df = 1 for the probability of measurement delay; F = 12.93 df = 1 for manoeuvring index; $p < 0.05$ for each).

Figure 6 shows the average results of 12000 experiments (3 OOSM and multiple imputation methods x 2 probability of measurement delay x 2 manoeuvring index) which summarise the accuracy of each method. It further shows that BAMI has the best accuracy throughout the entire spectrum in terms of the probability of measurement and manoeuvring index. Tukey's multiple comparison tests showed significant differences between BAMI and the other individual OOSM methods at the 5 per cent level. Once again, no interaction effects were found to be statistically significant at the 5 per cent level.

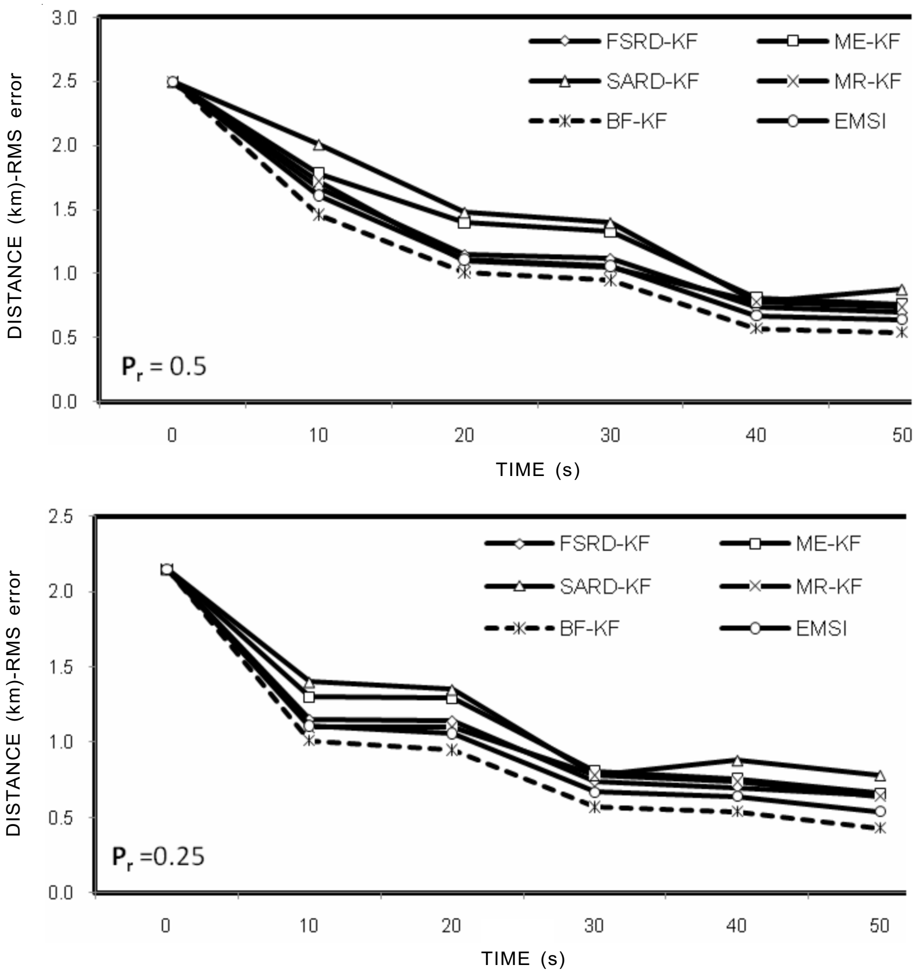


Figure 4. RMS performance of a straight line motion target with single delay OOSM (P_r=0.5 and 0.25; maneuvring index=0.3).

Table 1. Computational comparison of OOSM and single imputation methods (single delay) in minutes

	Methods					
P_r	FSRD-KF	ME-KF	SARD-KF	MR-KF	BF-KF	EMSI
0	2.67	2.78	3.71	2.01	6.54	6.78
0.25	3.15	3.40	4.43	2.57	6.54	6.78
0.5	3.74	3.97	4.78	2.64	6.55	6.78

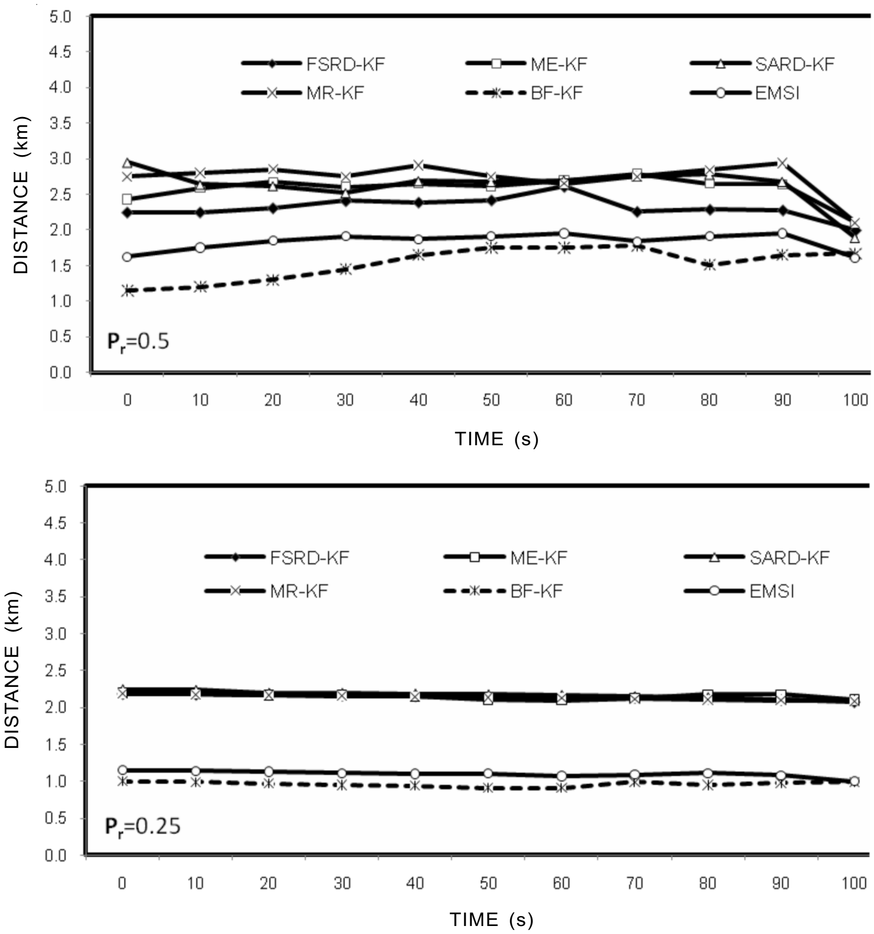


Figure 5. RMS performance of a highly manoeuvring target with single delays OOSM (P_r=0.5 and 0.25; manoeuvring index=1.0).

From these experimental results, the following observations were made:

- For non-manoeuvring tracking, all the three methods significantly reduce accuracy at all time and distance levels. Otherwise, all the methods show a very good fit when no measurements are delayed). In fact, at lower distance levels (between 20 s and 90 s) BAMI and BF-KF compare favourably. Overall, BAMI achieves the highest accuracy rates as a method for handling delayed observations, followed by BF-KF and FSRD-KF, respectively (Fig. 7).
- For manoeuvring tracking, there appears to be no difference in performance between BAMI and BF-KF, especially when the probability of delay increases. For lower probabilities, the difference in performance was quite prominent as shown in Fig. 8. Nonetheless, BAMI still outperforms FSRD-KF in terms of RMS error.
- Overall, the computational cost of BAMI was about three times that of FSRD-KF and almost one-and-a-half time that of BF-KF. (Table 2). P_r

5. DISCUSSION AND CONCLUSIONS

The major contribution of the paper is the use of simulation experiments to demonstrate the effectiveness of OOSM algorithms to handle delayed measurements. The referred techniques are well known, but the extensive empirical evaluation of these methods is an original contribution. Furthermore, imputation procedures are not of widespread use in sensor data fusion, so showing the possibility of using the techniques on handling OOSM

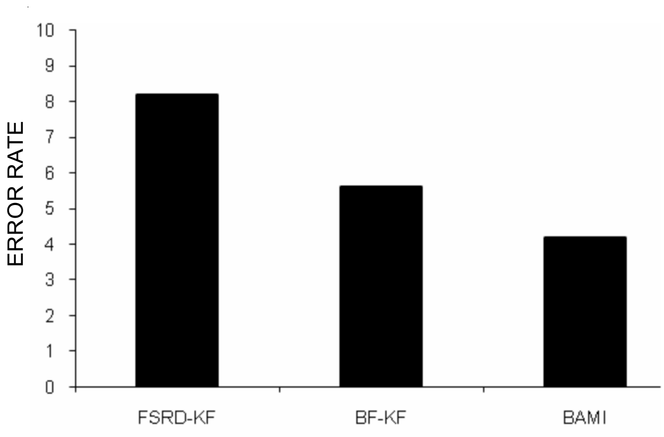


Figure 6. Overall means for OOSM and multiple imputation methods (multiple delays).

data is it another contribution for robotics learning.

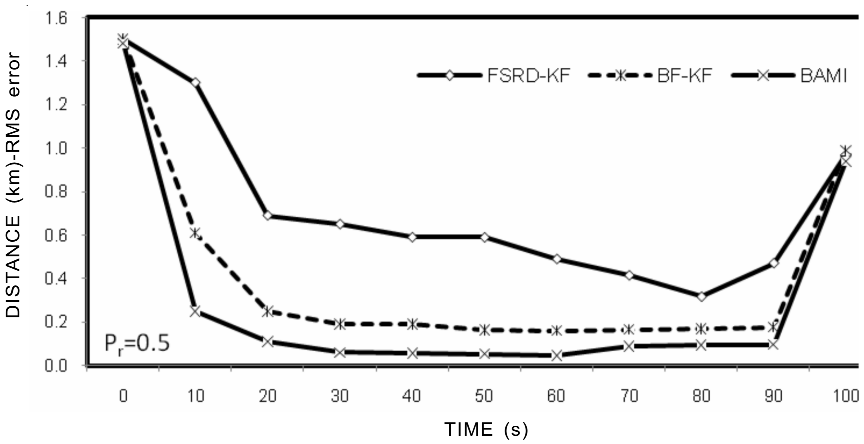
The empirical study is based on simulated data, and the results suggest that imputation strategies can be successfully applied to deal with delayed measurements. Based on preliminary evidence, it has been found that EMSI performs comparable with BF-KF for single delay measurements data while BAMI achieves higher accuracy rates for multiple delayed measurements. The good performance of BAMI could be attributed to it variance averaging benefit even if it came at a high computational cost.

Table 2. Computational comparison of OOSM and multiple imputation methods (multiple delay) in minutes

	Methods		
P_r	FSRD-KF	BF-KF	BAMI
0	3.19	8.14	10.67
0.25	4.43	8.14	10.67
0.5	5.99	8.14	10.67

The results further show the impact on the performance of methods is caused by the probability of measurement delays. Bigger positional error rates were achieved by methods for high probability delays with bigger performance differences among methods. Also, given that the performance of each method varies by probability of measurement delay, it appears that the treatment of delayed measurements not only heavily depends on the probability of measurement delay but on the range of manoeuvring target tracking. The worst performance achieved by methods is for non-manoeuvring target tracking. This was a rather surprising result, which is in not in accordance with statistical theory which considers missing completely at random (MCAR) as easier to deal with and IM data as very difficult to handle[45].

From both experiments, there exists threats to the validity of the results. Potential threats include the use of simulated data, which could have involuntarily introduced biases, especially if those measurements considered as



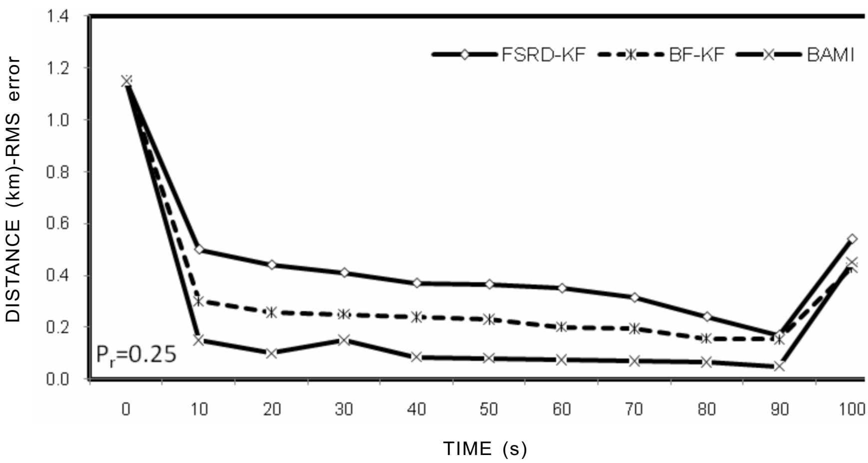


Figure 7. RMS performance of a straight line motion target with multiple delays OOSM (P_r=0.5 and 0.25).

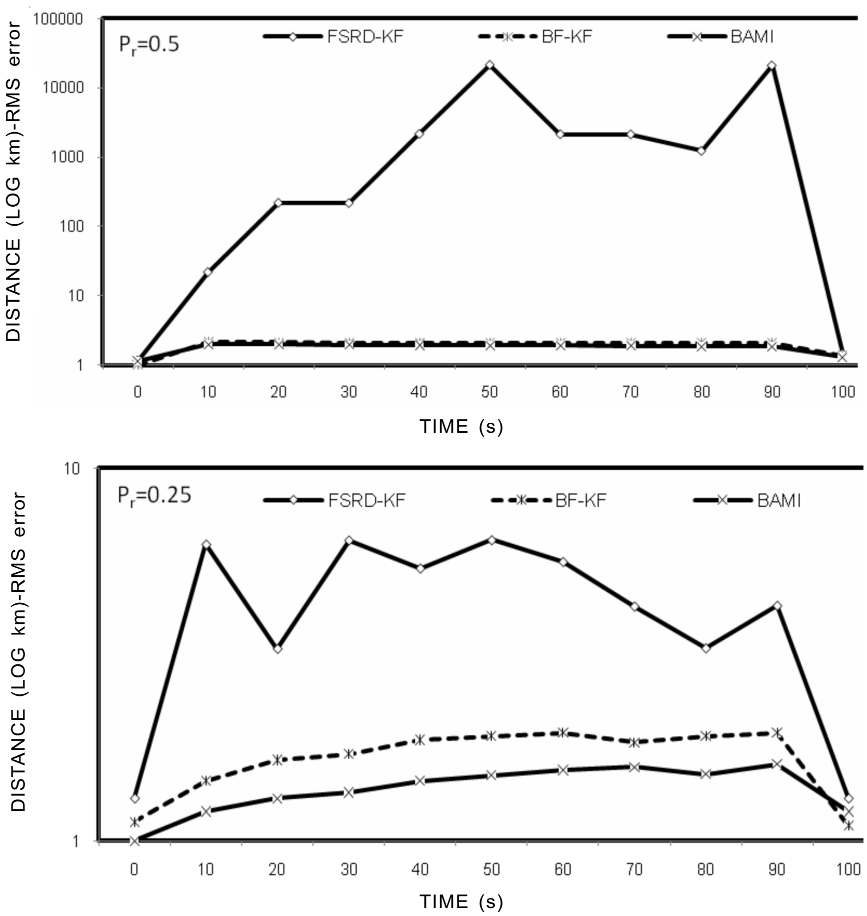


Figure 8. RMS performance of a highly manoeuvring target with multiple delays OOSM (P_r=0.5 and 0.25).

being delayed contained important information and they were not delayed, as we assumed. However, the experimental results were carefully validated. For example, the experiments were conducted under the supervision of a domain expert who had a deep understanding for sensor data. This was a time-consuming exercise on ourselves and the expert.

The issue of determining whether or not to apply an imputation strategy to a given sensor dataset given that there are delayed measurements, must be considered. For the work described here, the data were simulated. Unfortunately, this type of information is rarely known for most real-world applications. In some situations, it may be possible to use domain knowledge to determine the mechanism generating the delayed measurements. For situations where this knowledge is not available, the conservative nature of the consensus dictates that the measurements will be delayed randomly.

To sum up, this paper provides the beginnings of a better understanding of the relative strengths and weaknesses of model-based imputation strategies to handle delayed measurements. It is hoped that it will motivate future theoretical and empirical investigations into incomplete data and related (soft-) prediction, and perhaps reassure those who are uneasy regarding the use of imputed data in software prediction.

6. FUTURE STUDIES

The OOSM and imputation methods were applied on only one dataset. This work could be extended by considering a more detailed simulation study using much more balanced additional types of datasets or even smaller datasets required to understand the merits of imputation. In addition, using as many datasets as possible in comparative simulation study would enable a more sound generalisation of the results. This work could also be extended to a comparative evaluation of datasets with artificially simulated missingness against original datasets. The authors leave the above issues to be investigated in the future.

ACKNOWLEDGEMENTS

The work was funded by the CSIR under project MDSARR1. The author would like to thank Chris Jones for his helpful discussion and the anonymous reviewers for their useful comments.

REFERENCES

1. Kalman, R.E. A new approach to linear filtering and prediction problems. *Transactions ASM and J. Basic Eng.*, 1960, 33-45
2. Brown, R.G. & Hwang, P.Y.C. Introduction to random signals and applied Kalman filtering. Ed. 2, John Wiley & Sons, Inc. 1992.

3. Bar-Shalom, Y. Update with out-of-sequence measurements in tracking: Exact solution. *In* Proceedings of the SPIE Conference on Signal and Data Processing of Small targets, Orlando, FL, USA, 2000, pp. 51-556.
4. Wang, H.; Kirubarajan, T.; Li, Y. & Bar-Shalom, Y. Precision large scale air traffic surveillance using and imm estimator with assignment. *IEEE Trans. Aerospace, Electro. Syst.*, 1999, **35**(91), 255-66.
5. Mallick, M.; Coraluppi, S. & Carthel, C. Advances in asynchronous and decentralized estimation. *In* Proceedings of the 2001 IEEE Aerospace Conference, Big Sky MT. 2001.
6. Mallick, M. & Marrs, A. Comparison of the KF and particle filter based out-of-sequence measurement filtering algorithms. *In* Proceedings of the 6th International Conference on Information Fusion, 8-10 July 2002, Cairns, Australia.
7. Mallick, M.; Krant, J. & Bar-Shalom, Y. Multi-sensor multi-target tracking using out-of-sequence measurements. *In* Proceedings of the 5th International Conference on Information Fusion, 8-11 August 2002, Annapolis, MD, USA.
8. Mallick, M.; Zhang, K. & Li, X.R. Comparative analysis of multiple-lag out-of-sequence measurement filtering algorithms. In Proceedings Signal and Data Processing of Small Targets, San Diego, CA, USA. 2003.
9. Challa, S.; Evans, R.H. & Wang, X. A Bayesian solution and its approximation to out-of-sequence measurement problems. *Information Fusion*, 2003, **4**, 185-99.
10. Bar-Shalom, Y.; Chen, H. & Mallick, M. One-step solution for the multistep out-of-sequence measurements problem in tracking. *IEEE Trans. Aerospace Electro. Syst.*, 2004, **40**, 27-37.
11. Bar-Shalom, Y.; Chen, H.; Mallick, M. & Washburn, R. One-step solution for the general out-of-sequence measurement problems in tracking. *In* Proceedings of IEEE Aerospace Conference, Big Sky MT. 2002.
12. Bar-Shalom, Y.; Li, X.R. & Kirubarajan, T. Estimation with applications to tracking and navigation. Wiley & Sons. 2001.
13. Bar-Shlaom, Y. & Li, X-R. Estimation and tracking: Principles, techniques and software. Artech House, MA, USA. 1993.
14. Blackman, S.S. & Ropoli, R. Design and analysis of modern tracking systems. Artech House, MA, USA. 1999.
15. Marcus, G.D. Tracking with measurements of uncertain origin and random arrival times. Department of Electrical Engineering and Computer Science, University of Connecticut, Storrs, USA. 1979. M.S. Thesis.
16. Hilton, R.D.; Martin, D.A. & Blair, W.D. Tracking with time-delayed data in multisensor systems. NSWCDD/TR-93/351, Dahlgren, VA. 1993.
17. Thomopoulos, S.C.A. & Zhang, L. Decentralized filtering with random sampling and delay. *Information Science*, 1994, **81**, 117-31.
18. Ravn, O.; Larson, T.D.; Andersen, N.A. & Poulsen, N.K. Incorporation of time delayed measurements in a discrete time Kalman filter. *In* Proceedings of the 37th IEEE Conference on Decision and Control, Tampa, Florida, USA, National Bureau of Standards, 1998. pp. 3972-977.
19. Ray, A. Output feedback control under randomly varying distributed delays. *J. Guidance, Control Dyna.*, 1994, **17**(4), 701–11.
20. Lewis, R. Optimal estimation with an introduction to stochastic control theory. John Wiley & Sons, Inc. 1986.
21. Challa, S.; Evans, R. & Wang, X. Target tracking in clutter using time delayed out-of-sequence measurements. *In* Proceedings of Defence Applications of Signal Processing (DASP). 2001.
22. Tasoulis, D.K.; Adams, N.M. & Hand, D.J. Selective fusions of out-of-sequence measurements. *Information Fusion*, 2009. [In press].
23. Kwakernaak, H. Optimal filtering in linear systems with time delays. IEEE Trans. Automatic Contr., 1967, 12, 169-73.
24. Richard, J.P. Time delay systems: A review of some recent advances and open problems. *Automatica*, 2003, **39**, 1667-694.
25. Mallick, M. & Bar-Shalom, Y. Non-linear out-of-sequence measurement filtering with applications to GMTI tracking. *In* Proceedings of SPIE Conference Signal and Data Processing of Small Targets, Orlando Florida, USA. 2002.
26. Zhang, H.; Zhang, D. & Xie, L. Necessary and sufficient condition for finite horizon H" estimation of time delay systems. *In* Proceedings of the 43rd IEEE Conference on Decision and Control, December, USA. 2003.
27. Zhang, H.; Zhang, D. & Xie, L. An innovation approach to H" prediction for continuous-time systems with application to systems with time delayed measurements. *Automatica*, 2003, **40**, 1253-261.
28. Anderson, B.D.O. & Moore, J.B. Optimal filtering. Prentice Hall. 1979.
29. Hsiao, F-H. & Pan, S-T. Robust Kalman filter synthesis for uncertain multiple time-delay stochastic systems. *J. Dyna. Sys, Meas. Control*, 1996.
30. Kaszkurewicz, E. & Bhaya, A. Discrete time state estimation with two counters and measurement delay. *In* Proceedings of the 35th IEEE Conference on Decision and Control. 1996.
31. Matveev, A. & Savkin, A. The problem of state estimation via asynchronous communication channels with irregular transmission times. *IEEE Automatic Contr.*, 2003, **48**(4), 670-76.
32. Larsen, T.; Poulsen, N.; Anderson, N. & Ravino, O. Incorporating of time delayed measurements in a discrete-time Kalman filter. In CDC '98, Tampa, Florida. 1998.
33. Alexander, H.L. State estimation for distributed systems with sensing delay. *In* SPIE Conference Proceedings

of Data Structures and Target Classification, 1991, 1470, pp. 103-11.
34. Zhang, K.; Li, X.R. & Zhu, Y. Optimal update with out-of-sequence measurements. *IEEE Trans. Signal Process.*, 2005, **53**(6), 1992-004.
35. Maybeck, P.S. Stochastic models, estimation and control. Vol. 1 Academic Press. 1979.
36. Orton, M. & Marrs, A.D. A Bayesian approach to multi-target tracking and data fusion with out-of-sequence measurements. *In* IEE International Workshop on Target Tracking Algorithms and Applications. 2001.
37. Geld, A. Applied optimal estimation. The MIT Press. 1974.
38. Bar-Shalom, Y. & Li, X.R. Multitarget-multisensor tracking: Principles and techniques. YBS Publishing, 1995.
39. Julier, S.J. & Uhlmann, J.K. Fusion of time delayed measurements with uncertain time delays. *In* Proceedings of the American Control Conference, 2005. pp. 4028-033.
40. Rhéaume, F. & Benaskeur, A. Out-of sequence measurements filtering using forward prediction. Technical Report TR 2005-484, Defence R & D Canada–Valcartier. 2007.
41. Lu, X.; Zhang, H.S.; Wang, W. & Teo, K.L. Kalman filtering for multiple time-delay systems. *Automatica*, 2005, **41**(8), 1455-461.
42. Bar-Shlaom, Y. & Fortmann, T.E. Tracking and data association. Academic Press, New York. 1988.
43. Rubin, D.B. Multiple imputation for nonresponse in surveys. John Wiley and Sons, New York. 1987.
44. Rubin, D.B. Multiple imputation after 18+ years. *J. Amer. Statistical Asso*., 1996, **91**, 473-89.
45. Schafer, J.L. Analysis of incomplete multivariate data. Chapman and Hall, London. 1997.
46. Breiman, L.; Friedman, J.H.; Olshen, R.A. & Stone, C.J. Classification and regression trees. Chapman and Hall Inc. 1984.
47. S-PLUS 6.2 for windows. MathSoft, Inc., Seattle, Washington, USA. 2003.
48. Tanner, M.A. & Wong, W.H. The calculation of posterior distributions by data augmentation. *J. Amer. Statistical Asso*., 1987, **82**, 528-50.
49. Gilks, W.R.; Richardson, S. & Spiegelhalter, D.J. Markov chain Monte Carlo in practice. Chapman and Hall, London. 1996.
50. Dempster, A.P.; Laird, N.M. & Rubin, D.B. Maximum likelihood estimation from incomplete data via the EM algorithm. *J. Royal Statist. Soc., Ser. B*, 1977, **39**, 1-38.
51. Allison, P.D. Missing data. Thousand Oaks, Sage, CA. 2001.
52. Challa, S.; Evans, R.J.; Wang, X. & Legg, J. A fixed lag smoothing solution to out-of-sequence information fusion problems. *Commu. Info. Sys*., 2002, **2**(4), 327-50.
53. MINITAB Statistical Software for Windows 9.0. MINITAB, Inc., PA, USA. 2002.
54. Kirk, R.E. Experimental design. Ed. 2. Brooks, Cole Publishing Company, Monterey, CA. 1982.