OUTPUT PREDICTION UNDER SCARCE DATA OPERATION. CONTROL APPLICATIONS.

P. ALBERTOS*, R. SANCHIS†, A. SALA*

*Departamento de Ingeniería de Sistemas y Automática. (DISA)
Universidad Politécnica de Valencia, Apdo. 22012, E-46071 Valencia, Spain

† Unitat Predepartamental de Tecnologia, Universitat Jaume I,
Campus de Penyeta Roja, 12071, Castellón, Spain,

Keywords: Output estimation, State estimation, Parameter estimation, Missing data, Unconventional sampling

Abstract. The problem of estimating the output in missing-data situations is addressed, in the case a process model is present. A simple algorithm is presented that uses the input-output model (difference equation), replacing the unknown past values by estimates when necessary. It is compared to state-space approaches such as time-varying Kalman filtering. The analysis of its convergence is carried out for the particular case of dual-rate scarce sampling patterns. The effects of disturbances are also studied. The use of extended order models allows the design of the desired error dynamics. Applications such as parameter estimation and the control under scarce data operation are outlined.

Corresponding author: Pedro Albertos, Tel. +34963879570, fax. +34963879579, email: pedro@aii.upv.es
1. INTRODUCTION

In many industrial applications the control signal is updated at a fixed rate $T$, but the output is measured with a different timing pattern and, sometimes, by various sensors, each one having a maybe different sampling rate, and reliability. In some practical cases the output is not available at every sampling time due to computer overload, communication errors, shared sensors or event-driven sensors. Different authors have dealt with the modelling of such systems when the measurement pattern is periodic (Albertos 1990), (Salt et al. 1993), (Araki 1993), based on the definition of a model that relates outputs measured at one rate with inputs updated at another rate. This allows, for example, to tackle the problem of the design of a dual-rate control system (Albertos et al. 1996). However, none of them try to explicitly estimate the outputs at the instants when they are unavailable, in order to apply standard control or parameter estimation techniques. Furthermore, the multirate approach cannot deal with random sampling.

This paper deals with the estimation of all the outputs of a process from scarcely sampled measurements, assuming that the process model is known. Although the convergence analysis of the algorithms here presented is based on the assumption of regular output availability (1 measurement is available every $N$ input periods), all of them are easily implemented and tested in the case of other irregular data availability patterns.

Output estimates can be used in conventionally sampled control and least squares (LS) parameter estimation schemes (Isermann, 1981). In this paper, the pseudo-recursive LS scheme introduced by (Albertos et al., 1992; Adams et al., 1994) is used with the presented missing data prediction algorithms. Unbiased convergence is difficult to guarantee in a general situation. Off-line Expectation-Maximization algorithms (Isaksson, 1993) can be an alternate approach.

The layout of this paper is as follows: the estimation problem is defined in section 2 and the simplest open loop predictor is described. A difference-equation algorithm based on a vector composed by estimates and measurements is presented in section 3, jointly with its stability analysis. The use of higher order models is shown to modify the error dynamics, allowing the design of overparameterized models to obtain a desired dynamics or noise filtering. The effects of measurement noise and state disturbances are also taken into account. Other approaches, such as state observers, Kalman filtering or state reconstruction, are described in section 4. Estimation of model parameters with missing-data is outlined in section 5, and a control application of the output predictors is suggested in section 6. A conclusion section stresses the main results.

2. PROBLEM STATEMENT

Consider a stable, observable, linear time-invariant SISO continuous time (CT) system described by:
\[
\dot{x}(t) = A x(t) + B u(t) \quad y(t) = x(t)
\]

where \( x \in \mathbb{R}^n \), \( y \in \mathbb{R} \), \( u \in \mathbb{R} \), the input being updated by a computer at a constant sampling period \( T \) (\( u_k = u(kT) \)) through a ZOH. If the output sampling and the input updating are synchronous, there exists a discrete time equivalent transfer function and a state space representation defined by:

\[
\begin{align*}
\left\{ \sum_{i=0}^{\infty} a_i q^{-i} \right\} y_k &= \sum_{i=0}^{\infty} b_i q^{-i} u_k \\
x_{k+1} &= A x_k + u_k \\
y_k &= x_k
\end{align*}
\]

where \( A = e^{AT} \) and \( b = \int_0^T e^{A \tau} d\tau \cdot b_c \).

Consider now that only a few output measurements are available due to sporadic sensor faults, dual-rate sampling (one measurement available every N input periods), random output availability, etc. A missing data sampling pattern is defined as the data obtained from the set of all inputs at sampling instants \{u0, u1, u2, ...\} and a subset of the full output set \{y0, y1, ...\}. If the probability of having \( n \) consecutive output samples is very small the pattern is named scarce data pattern. In this paper, scarce data are used either to estimate the outputs at the instants when they are unavailable, or to observe the state of the system. An algorithm giving the estimates of the output under these conditions will be called a missing-data predictor.

The simplest way to estimate the system output is the open loop structure, i.e. to dismiss the output measurements, obtaining the estimates from a model whose inputs are the same as those of the process. The equations can be implemented in state space form (3) or in input-output form (2). In the last case, the equation of the output estimate is:

\[
\hat{y}_k = -\sum_{i=1}^{\infty} \hat{\psi}_k_i + \sum_{i=1}^{\infty} \tau_k_i = \psi^{T}_k \cdot \theta
\]

where

\[
\begin{align*}
\psi_k_i &= -\gamma_{k-1} \cdots -\gamma_{i} \quad u_{k-1} \cdots u_{i} \\
\theta &= \begin{bmatrix} a_1 & \cdots & b_1 & \cdots & b_n \end{bmatrix} \end{align*}
\]

The dynamics of the estimation error \( e_k = \gamma_k - \hat{\gamma}_k \) is obviously that of the system. Therefore, the estimator is stable independently of the data availability. Nevertheless, if the system dynamics is slow or oscillating, then the estimation error transient is not adequate. The main drawback of this approach is the lack of robustness against modelling errors or external disturbances because no output feedback is considered.

In the next sections, several approaches to construct missing data predictors will be discussed. The variable \( N \) is used throughout the text to define the number of input periods between two consecutive measurements, for
dual-rate patterns.

3. OUTPUT ESTIMATION WITH MIXED REGRESSION.

3.1 Definition of the algorithm.

In order to overcome the problems of the open loop algorithm, but preserving its simplicity, a simple modification is proposed. Instead of directly using the vector $\theta$ on equation (4), the available measurements will be first used to update that vector, replacing the output estimates by them. This produces an output feedback to the algorithm. By defining:

$$\theta_a = a_1 \cdots a_n^T; \quad \theta_b = b_1 \cdots b_n^T$$

(6)

the algorithm can be written as:

$$\hat{y}_k = y(k-1)^T \cdot \theta + y(k-1)^T \cdot \theta$$

(7a)

$$\bar{y}_k = (1 - r_k) \cdot \hat{y}_k + r_k \cdot y_k$$

(7b)

$$\psi_n(k) = -r_k \cdots - \psi_{n-1} \cdots \psi_n(k) = \hat{y}_k \cdots \hat{y}_{n-1}$$

(7c)

where $r_k$ is the coefficient of data availability ($r_k = 1$ if measurement available, $r_k = 0$ otherwise). $\hat{y}_k$ is set equal to the measurement when available, otherwise is left as the difference equation output estimate. A block diagram (the indicated estimation error $e_k$ is only available if $r_k = 1$) is shown in fig. 1.

The main advantage of this algorithm is its appealing simplicity. It can be implemented with a very low computer overhead. The drawback is that the stability of the resulting error dynamics is not guaranteed. In fact, it depends not only on the data availability, but also on the process parameters and the sampling period. The following results develop this idea.

3.2 Stability analysis

In order to give insight into the problem, let us consider a 2nd order CT linear system with transfer function:

$$G(s) = \frac{k}{s^2 + \alpha s + \beta}$$

(8)

whose input is updated at period $T$, and whose output is measured every $N$ input periods. To simplify the notation, assume a canonical observability form in the state description (3) of the discretized system. The following result allows to test the predictor stability.

Lemma 3.1. For the process above, (8), the algorithm described by equations (7), has the implicit error dynamics given by:

$$e_{(k+1)} = \mathbf{Q} \cdot \mathbf{A}^N \cdot \mathbf{Q}^{-1} \cdot \mathbf{e}_{(k-N)}$$

(9)
where \( e_k \) = \( y_k - \hat{y}_k \), the subindex (2,2) means the element (2,2) of the matrix, and \( O \) is the observability matrix,

\[
O = \begin{bmatrix} \cdot & 1 \\ \cdot & c \end{bmatrix}
\]

The expression (9) can also be obtained as:

\[
e_{k-1} = R x \left( \frac{1 + q^{-1} \gamma A}{1 + \gamma A + q^{-1} \gamma A} \right) \cdot e_{k-(N+1)}
\]

(10)

where \( a_i \) are the parameters of the discrete ZOH equivalent transfer function (2) and \( IR_j(\text{H}(q^{-1})) = f_j \) is the \( j \)th term of the impulse response of \( \text{H}(q^{-1}) \).

Proof. As the state can be reconstructed from two immediate past measurements by:

\[
x_{k, \rightarrow, k-1} = A^1 \begin{bmatrix} y_k \- b u_k \- 1 \- 1 \\ y_{k-1, \rightarrow, k-1} \end{bmatrix}
\]

and the outputs at instants \( k \) and \( k-1 \) can be expressed in terms of \( x_{k,N:} \) by the usual discrete convolution formula

\[
y_k = c x_k = c A^N x_{k-N,} + f_1(u_k, \ldots, u_{k-N,})
\]

\[
y_k = c x_{k-1} = c A^{N-1} x_{k-N,} + f_1(u_k, \ldots, u_{k-N,})
\]

then

\[
\begin{bmatrix} y_k \- k-1 \- 1 \\ y_{k-1, \rightarrow, k-1} \end{bmatrix} = A^{-1} \begin{bmatrix} y_k \- k-1 \- 1 \\ y_{k-1, \rightarrow, k-1} \end{bmatrix} + (u_k, \ldots, u_{k-1})
\]

(11)

No measurements are available between \( y_{k,N} \) and \( y_k \), and therefore, the evolution of the output estimates in the algorithm (7) between instants \( k-N \) and \( k \) can be written in the same way:

\[
\begin{bmatrix} \bar{y}_k \- k-1 \- 1 \\ \bar{y}_{k-1, \rightarrow, k-1} \end{bmatrix} = A^{-1} \begin{bmatrix} \bar{y}_k \- k-1 \- 1 \\ \bar{y}_{k-1, \rightarrow, k-1} \end{bmatrix} + (u_k, \ldots, u_{k-1})
\]

where \( \bar{y} \) and \( \bar{y}' \) are defined in (7). Substracting the above expressions, and defining \( \bar{e}_k = y_k - \bar{y}_k \), it yields:

\[
\begin{bmatrix} e_k \- k-1 \- 1 \\ e_{k-1, \rightarrow, k-1} \end{bmatrix} = A^{-1} \begin{bmatrix} e_k \- k-1 \- 1 \\ e_{k-1, \rightarrow, k-1} \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \end{bmatrix}
\]

The error is zero at the measuring instants, i.e. \( \bar{e}_{k-1} = 0 \) and \( \bar{e}_k = 0 \). Hence, the above equation is:

\[
\begin{bmatrix} e_k \- k-1 \- 1 \\ e_{k-1, \rightarrow, k-1} \end{bmatrix} = A^{-1} \begin{bmatrix} 0 \\ 0 \end{bmatrix}
\]

Taking into account that at the instants when no measurement is available, \( e_j = \bar{e}_j \), equation (9) is obtained.

Expression (10) can be easily derived. Being

\[
O = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \rightarrow O^{-1} = \begin{bmatrix} 0 & 1 \\ 1 & a_j \end{bmatrix}
\]

the observability matrix and its inverse, then:
\[
Q_{(A^N)^{-1}}^{(A^N)^{-1}} = \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 1 \\ \vdots \\ 1 \\ 0 \end{bmatrix} = A^N \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}
\]

that is, the \(N\)th term of the impulse response of:

\[
\frac{1+a_1q^{-1}}{1+a_1q^{-1}+a_2q^{-2}}.
\]

(12)

This lemma shows how the stability of the observation algorithm (7) depends on the data availability. But, contrarily to the first “logic” idea, a higher data pattern density does not mean a more stable observer. In fact, if \(N\) is sufficiently large, the above impulse response term is very small, and hence, the stability of the observer is guaranteed, for stable processes.

For a given data availability \(N\), the dynamics of the estimation error depends on the sampling period. If the poles of the CT transfer function \(G(s)\) are complex, the expression of the discrete parameters is:

\[
G(s) = \frac{k_0}{s^2 + \sigma^2 + \omega^2} \quad a_1 = -e^{-\sigma T} \cdot \cos(\omega T) \\
\quad a_2 = -\sigma
\]

Let us consider \(N=3\). The observer will be unstable if \(|IR_N| > 1\), i.e. if \(\sigma T < \ln(2\cos(\omega T)) / 3\). In the case \(N=4\), the error dynamics is unstable if \(\sigma T < \ln(|2-4\cos^2(\omega T)|) / 4\).

**Lemma 3.2.** Consider a 2nd order linear CT system whose input is updated at period \(T\), and whose output is sampled 1 every \(N\) input periods. The observation algorithm (7) is unstable for sufficiently small sampling period with any value of \(N>2\).

**Proof.** By lemma 3.1, the estimated error system is unstable if the \(N\)th term of the impulse response of the discrete system (12) is larger than 1. As \(\lim_{T \to 0} a_1 = -2\), \(\lim_{T \to 0} a_2 = 1\), for \(T \to 0\) the impulse response will approach:

\[
IR_N \to -\infty \quad (N=1)
\]

For \(N>2\) it is larger than 1. As the function \(IR_N(T)\) is continuous in \(T\), the lemma results. \(\square\)

For higher order systems, the dynamics cannot be reduced to a first order one. The next theorem extends the lemma result.

**Theorem 3.1.** The dynamics of the estimation error in algorithm (7), for system (2), when there is one measurement every \(N\) input periods, is determined by the equation:

\[
E(k + \varphi) = M \cdot E(k)
\]

(13)
where $M_c$ is obtained from matrix $M=OAXO^T$ by eliminating the rows and columns \{1+iN, $i=1,\ldots,\text{int}(n-jN)\}$, (thus, the size of matrix $M_c$ is $(d\times d)$, with $d=1-\text{int}(n-jN)$). $O$ is the observability matrix, as previously defined.

Furthermore, the matrix $M$ can also be constructed as:

$$M = XAX^{-1} = \begin{bmatrix}
I_d(N+1) & I_l(N+1) & \cdots & I_{e-1}(N+1)
\end{bmatrix}
\begin{bmatrix}
I_d(N+1) & I_l(N+1) & \cdots & I_{e-1}(N+1)
\end{bmatrix}^T
$$

(14)

where the term $I_l(j)$ is defined as:

$$I_l(j) = \mathcal{R} \left( \frac{(1+a_j q^{-1}\cdots+a_j q^{-1})q^{-1}}{1+(q^{-1})_j} \right)
$$

(15)

**Proof.** A set of $n$ outputs can be obtained from a previous set of other $n$ outputs and inputs. The idea is to reconstruct the state and then to apply the state dynamic equations forward, following the same procedure than in (11), leading to:

$$
\begin{bmatrix}
\begin{array}{c}
k+n+1 \\
k+n+2 \\
k+n+3 \\
k+n+4 \\
k+n+5
\end{array}
\end{bmatrix} = \begin{bmatrix}
A^{n-1} & \cdots & A^1 & \begin{bmatrix}
\begin{array}{c}
k+n+1 \\
k+n+2 \\
k+n+3 \\
k+n+4 \\
k+n+5
\end{array}
\end{bmatrix}
\end{bmatrix} +
\begin{bmatrix}
\begin{array}{c}
k+n+1 \\
k+n+2 \\
k+n+3 \\
k+n+4 \\
k+n+5
\end{array}
\end{bmatrix}
$$

(16)

The estimation equation (7) leads to an identical equation for the output estimates $\hat{y}$, but only if there are no new measurements between $y_{k+n}$ and $y_{k+j+n}$. As before, substracting the expressions, one obtains:

$$
\begin{bmatrix}
\begin{array}{c}
k+n+1 \\
k+n+2 \\
k+n+3 \\
k+n+4 \\
k+n+5
\end{array}
\end{bmatrix} = \begin{bmatrix}
A^{n-1} & \cdots & A^1 & \begin{bmatrix}
\begin{array}{c}
k+n+1 \\
k+n+2 \\
k+n+3 \\
k+n+4 \\
k+n+5
\end{array}
\end{bmatrix}
\end{bmatrix} +
\begin{bmatrix}
\begin{array}{c}
k+n+1 \\
k+n+2 \\
k+n+3 \\
k+n+4 \\
k+n+5
\end{array}
\end{bmatrix}
$$

(16)

At measuring instants, $\bar{e}_i = 0$, independently of the value of $e_i = \gamma_i - \gamma_i$. Therefore, the above equation is not valid if there is any measurement between instants $(k+n+1)$ and $(k+j+n-1)$. The last value could correspond to a new measurement because this value does not affect the equations of the previous $n-1$ values. In that case, $\bar{e}_{k+n+1} = 0$ and the value of $e_{k+j+n}$ is not valid. Therefore, if the measurements are obtained 1 every $N$ periods, then $j=N$ is the largest value of $j$ for which the equation is valid. Let us define a column error vector $\bar{E}(k+1) = \pi_{k+1} \cdots \pi_{k+j}$ such that the first component is zero (there is a measurement at instant $k+n$).

In that case, elements \{1+iN, $i=1,\ldots,\text{int}(n-jN)\}$ in that vector are also null, and matrix $OAXO^T$ can be reduced by eliminating those columns. Furthermore, the same components of vector $\bar{E}(k+1)$ are also null, and therefore,
the same set of rows can be eliminated from the matrix. Thus, defining \( E(k) \) as the non zero elements of \( E(k) \):

\[
E(k) = [e_{k+n-1}, ..., e_{k+n_N-1}, ..., e_{k+n_{2N-1}}, e_{k+n_{2N-1}}, ...]
\]

i.e., a vector with length \( n-1 - \text{int}((n-1)/N) \), the equation can be written as:

\[
E(k + \nu) = M_c \cdot E(k)
\]

with \( M_c \) obtained by eliminating the \( \{1+iN, i=0, \ldots, \text{int}((n-1)/N)\} \) rows and columns from matrix \( OA^N O^T \).

The equation (14) is easy to derive, considering the particular form of the inverse of the observability matrix (in canonical form) and the definition in (15), so the first column elements of matrix \( OA^N O^T \) can be written as:

\[
M_{c(i,j)} = A^{N_i} - \begin{bmatrix} 0 & \cdots & 0 \end{bmatrix} 1 = \begin{bmatrix} 0 \end{bmatrix}(N + \nu - 1)
\]

similarly, the elements of the column \( i \) are:

\[
M_{c(i,j)} = A^{N_i} - \begin{bmatrix} 0 & \cdots & 0 \end{bmatrix} 1 = \begin{bmatrix} a_i \cdots a_{i-1} \end{bmatrix} (N + \nu - 1)
\]

and the result follows.

\( \square \)

**Example:** For a 2nd order system, with a data availability \( N=3 \), the dynamics of the observer is given by:

\[
e_{k+1} = \begin{bmatrix} 0^2 & 0 \end{bmatrix} e_{k-4}
\]

Therefore, if the sampling period is small and \( |a_1, a_2| > 1 \) the observer is unstable.

For a 4th order system and \( N=2 \Rightarrow d=2 \):

\[
M_c = \begin{bmatrix} 0^2 & 0 \end{bmatrix} \begin{bmatrix} -2 & -4 \\ 1 & 0 \end{bmatrix}
\]

The stability of the observer depends on the eigenvalues of the matrix (13) and it is easy to show that the terms \( I_d(j) \) and the parameter values decrease as the sampling period increase for stable processes. This means that for a sufficiently large period, the matrix \( M_c \) will have very small values, and the observer will be stable. Obviously, the opposite is also true, if the sampling period is sufficiently small, then the terms \( I_d(j) \) are very large, and the observer will easily become unstable. Another conclusion can be derived from (13). For a given sampling period, the terms \( I_d(N-j) \) decrease as \( N \) increases, and therefore, for a sufficiently large \( N \), the observer will be stable.

3.3 Disturbed Processes

Assume that some state and output disturbances are present, such that the system equations (3) are now:

\[
x_{k+1} = 4x_k + ru_k + \nu \theta
\]

\[
y_k = x_k \quad y_c(k) = x_k + \nu
\]

\[
(18)
\]

\[
(18)
\]
where $\omega_k \in \mathbb{R}$ and $\nu_k \in \mathbb{R}$ and $y_m(k)$ is the measurement, corrupted by the noise $\nu_k$. The effect of the disturbances on the prediction error $e_k$ of algorithm (7) is stated in the following lemma.

**Lemma 3.3.** Assume that the algorithm (7) is applied to the system (18), where $\omega_k$ and $\nu_k$ are white noise signals. Then, the dynamic equation of the output estimation error is linear with respect to the noise variables, and expressed by:

$$
\dot{E}(k + \nu)_{2 \times n} = OA^N O^{-1}_{2 \times 2 \times n} E(k)_{2 \times n} +
+ \gamma(N)_{2 \times 2 \times n, n, \lambda} G(N)_{2 \times \lambda} + O^N O^{-1}_{2 \times 2 \times n} \nu,
$$

where:

$$
G(N) = 
\begin{bmatrix}
g & cAg & \cdots & \cdots & cA^{N-2}g & \\
0 & \cdots & 0 & cAg & \cdots & cA^{N-1}g
\end{bmatrix}
$$

For a proof of the lemma, see (Albertos et al., 1997). This result can be used to determine the filtering characteristics of the output predictor.

3.4 Higher order model for observation.

A very simple modification can be introduced to make the predictor (7) stable and to reach an adequate error dynamics or disturbance filtering even in the case of small sampling period.

To modify the observer dynamics (13) a higher order model can be used in the mixed vector algorithm (7). This model must be obtained by multiplying the numerator and denominator of the simple transfer function model by an adequate polynomial. The idea is to obtain a new model, such as:

$$
\frac{dq^{-1} + \cdots + dq^{-N}}{1 + dq^{-1} + \cdots + dq^{-N}} = \frac{(b_0 q^{-1} + \cdots + b_0 q^{-N}) E(q^{-1})}{(1 + b_0 q^{-1} + \cdots + b_0 q^{-N}) E(q^{-1})} = \frac{E(q^{-1})}{C(q^{-1})}
$$

and then apply (7) to it.

The main difference is that the new vector of past outputs is larger than in the simple $\psi_{y,k}$ model, and hence it contains more measurements. This can produce a stabilizing effect if the polynomial $E(q^{-1})$ is appropriately chosen. This extension to the algorithm does not significantly increase its complexity.

The design of the polynomial $E(q^{-1})$ must be based either on the error dynamics given by (13) to lead to adequate eigenvalues of matrix $M$, or, if disturbances are considered, on adequate filtering in (19).
As an example, consider the 2nd order system, with a data availability \(N=3\). If the order of the model is increased by using a polynomial \(E(q^{-1})=1+\beta q^{-1}\) the observer dynamics is given by (13) with \(n=3\) and \(N=3\), and the system with denominator:

\[
1+c_1q^{-1} + c_2q^{-2} + c_3q^{-3} = (1+a_1q^{-1} + a_2q^{-2})(1+\beta q^{-1}) = + (a_1 + \beta t^- + (a_2 + a_1\beta t^- + a_3\beta t^-)
\]

This leads to:

\[
M_c = OA^O A_1^{-1} = \begin{bmatrix} \beta^2 & \gamma^2 \beta & a_2 \beta + \gamma a_3 \beta \\ -\beta \cdot t^- & -\gamma \beta \cdot t^- & -a_2 \beta \end{bmatrix}
\]

whose eigenvalues are minimum for:

\[
\beta = -\frac{a_2}{a_1} \rightarrow \text{Eigenvalues}(M_c) = \left\{ \frac{a_2}{a_1}, \frac{a_2}{a_1} \right\}
\]

If the sampling period is small, the above eigenvalues tend to -0.5, leading to a stable observer.

The last example shows how by extending the model transfer function with cancelling poles and zeros, the observer can be made stable and sufficiently fast. The order of the extended model depends on both the process order and the data availability. If in the last example a value of \(N=4\) is considered, the resulting matrix (14) for the extended order model is:

\[
\begin{bmatrix} (a_2 + a_1\beta)(a_1 + \beta t^- - a_2\beta + a_3\beta t^-) \\ (a_1\beta + a_1^2\beta - a_2\beta) \\ a_1^2\beta + a_2a_3\beta \end{bmatrix}
\]

whose eigenvalues can not be made smaller than 1 for any value of \(\beta\) if the sampling period is small (\(a_1\) close to -2, \(a_2\) close to 1). Therefore, in this case, a model order higher than 3 is needed to reach a stable predictor.

In order to do the extension in a more systematic way, it seems reasonable to choose the extended system order such that the number of free design parameters coincides with the number of eigenvalues to be assigned. Defining \(n_e\) as the order of the extended model, it is easy to show that the number of eigenvalues of matrix \(M_c\) is

\[
n_e - 1 - \inf\left( \frac{b-1}{N} \right)
\]

while the number of coefficients in the polynomial \(E(q^{-1})\) is \(n_e-n\). The heuristic criterion to select \(n_e\) should then be

\[
\inf\left( \frac{b-1}{N} \right) = -1 \Leftrightarrow \ n-1)N+1 \leq b < uN+1
\]

In general, selecting the order of the extended model as \(n_e=N(n-1)+1\) results in a prediction error dynamics whose eigenvalues can be made stable by proper choice of the factors \(\beta_i, i=1,\ldots,(N-1)(n-1)\). This is stated in the following theorem.
Theorem 3.2. The eigenvalues of the dynamics of the estimation error in algorithm (7), for system (2), when there is one measurement every \( N \) input periods can be set to zero by using an extended model of order \( n_e = N(n-1)+1 \).

Proof. The proof is based on showing that the extended model can be selected such that the prediction error reduces to zero in a finite number of sampling periods. This dead-beat behaviour obviously means that all the eigenvalues are null. Assume that the extended denominator has the form:

\[
C(q^{-1}) = 1 + q^{-1} + N_{-1}q^{-N_{-1} - 1} + \cdots + q^{-n(N-1)-1}
\]

Assume also that the initial estimates are \( \hat{y}_1, \hat{y}_2, \ldots, \hat{y}_{(n-1)N} \), where \( \hat{y}_{jN+1} = y_{j+1} \), \( j=1, \ldots, n-1 \) are measurements. The form of \( C(q) \) implies that \( \hat{y}_{(n-1)N+2} = y_{(n-1)N+2} \) even not being measured, because it only depends on past measurements (no disturbances are assumed present). The same can be applied to \( \hat{y}_{jN+2}, j = 1, \ldots, 2(n-1) \), and recursively for any \( \hat{y}_{jN} \) after a finite number of iterations, implying that all the eigenvalues are null.

To complete the proof it is necessary to show that the extended denominator can be selected with the previous form. This is true if the system is the discrete ZOH equivalent of a CT observable system such that no CT pole has an imaginary part multiple of \( 2\pi/T \). In that case, the output at any instant can be obtained as a function of any \( N \) past outputs, i.e. there exists a model \( C(q) \) with the previous special form. But this model contains the poles of the original one, and therefore, there exists a polynomial \( \beta(z) \) such that a solution to \( C(q^{-1}) = \beta(q^{-1})A(q^{-1}) \) can be found by solving a system of linear equations.

In the last example, with \( n=2 \) and \( N=3 \), the resulting extended order is \( n_e=4 \). Applying equation (13) to the system with denominator:

\[
1 + c_0q^{-1} + c_1q^{-2} + c_2q^{-3} + c_3q^{-4} = (1 + a_0q^{-1} + a_1q^{-2} + a_2q^{-3} + a_3q^{-4} + a_4q^{-5})
\]

leads to the equation:

\[
\begin{bmatrix}
1 & -\beta^1 & -\beta^2 & -\beta^3 \\
-\beta^1 & 1 & -\beta^2 & -\beta^3 \\
-\beta^2 & -\beta^1 & 1 & -\beta^2 \\
-\beta^3 & -\beta^2 & -\beta^1 & 1
\end{bmatrix}
\begin{bmatrix}
\gamma_1 \beta^1 + \gamma_2 \beta^2 + \gamma_3 \beta^3 + \gamma_4 \beta^4 \\
\gamma_2 \beta^2 + \gamma_3 \beta^3 + \gamma_4 \beta^4 \\
\gamma_3 \beta^3 + \gamma_4 \beta^4 \\
\gamma_4 \beta^4
\end{bmatrix}
= \begin{bmatrix}
\alpha_1 \beta^1 + \alpha_2 \beta^2 + \alpha_3 \beta^3 + \alpha_4 \beta^4 \\
\alpha_2 \beta^2 + \alpha_3 \beta^3 + \alpha_4 \beta^4 \\
\alpha_3 \beta^3 + \alpha_4 \beta^4 \\
\alpha_4 \beta^4
\end{bmatrix}
\]

The eigenvalues of the above matrix are zero for \( \beta_1, \beta_2 \) obtained as outlined in the proof of the last theorem,:

\[
\beta_1 = \frac{a_2a_3}{a_3-a_1}, \beta_2 = \frac{a_2^2}{a_2-a_1}
\]

so the polynomial \( E(q) = 1 + \beta_1 q^{-1} + \beta_2 q^{-2} \) for finite-time response.

To consider the effect of disturbances, assume that a white measurement noise of variance \( \sigma_v \) is present. Take,
for example, a value of $a_1=-2, a_2=1$ (double integrator). The error equation is given by (19), with

$$M_{2424} = \begin{bmatrix}
-\beta^2 + \beta' \beta & \beta' - \beta \beta' + \beta \\
0 & -\beta' + \beta
\end{bmatrix}
$$

and

$$M_{(2,4,i)} = \frac{1}{\beta - \beta'} 2 \beta - 1^T$$

The variance-covariance matrix of the prediction error, $S_e$, can be obtained from the discrete Riccati equation:

$$S_e = M_{(2,4,2)} S_e M_{(2,4,2)}^T + M_{(2,4,1)} \sigma_e M_{(2,4,1)}^T$$

A design objective could be to minimize the diagonal elements of the above matrix while maintaining the eigenvalues of $M_e$ stable with an adequate time constant and damping factor. This problem can be solved numerically.

For the previous values of $\beta_1$ and $\beta_2$ that assign the eigenvalues to the origin, the variances are $\sigma_{11}=3.96 \sigma_e, \sigma_{22}=1.89 \sigma_e$. If for example the values $\beta_1=0.95, \beta_2=0.9$ are taken, the variances are $\sigma_{11}=1.21 \sigma_e, \sigma_{22}=1.1 \sigma_e$ while the eigenvalues are $\lambda_1=\lambda_2=0.85$.

The analysis of the general case is complex. If the order of the extended polynomial is selected as $n_e=N(n-1)+1$ the dynamics can be made stable with small eigenvalues. However, the minimization of the variance of the estimates may be a difficult problem because the equation is nonlinear in the parameters. Nevertheless, the calculation of the extended parameters is carried off-line, and therefore, numerical computation can be used. The on-line implementation is as simple as the non extended order case.

In this section, a regular sampling is assumed, with a fixed number of input periods between measurements. The results can be easily extended to the case of time varying intersampling periods following a periodic pattern. If the measurements availability($N$) is random then two approaches could be investigated. One of them is defining a fixed extended model of order $n_e=(n-1)N_{max}+1$ such that for some positive definite matrix $P$, the matrix $M_e(N)^T P M_e(N) - \sigma$ is negative definite for all possible values of $N$. The other option is to define a time varying extended model, updated every time a measurement is present, based on the relative position of the last measurements. In any case the stability is not easy to be guaranteed.

4. OTHER APPROACHES

Another observation method that can be applied is the state linear observer. The canonical observability representation of the system is used to build an state observer that uses only measured outputs. The output estimates can then be obtained from the state estimates. The difficulty, as before, is that only some measurements are available. If the sampling pattern is regular (1 of $N$) then the gain of the observer can be set up by pole.
placement with the matrix $A^N$ as process matrix. The observer equation should be:

$$\dot{x}_k = (A^N x_{k-1} + t^{N-1} bu_{k-1} + \ldots + u_{k-1} + \ldots + u_{k-1},)$$(20)

and the output estimates:

$$\hat{y}_{k+j} = \Phi^{j} \hat{x}_k + t^{j-1} bu_{k+j-1} + \ldots + u_{k+j-1}, \quad j = 1, \ldots, N-1$$

The current observer has been proposed to use the latest of the scarce available measurements.

The above procedure works well if the measurements pattern is regular, because in that case, the matrix $A^N$ does not change from one iteration to another, and a constant gain $L$ is adequate. If the number of periods between measurements change with time ($N$ is time varying), then the matrix $A^N$ may suffer an important change from one iteration to the next. The gain $L$ can then be designed as a constant matrix (based on the average value of $N$, or by robust design for the possible values of $N$), or it can be updated every iteration according to the value of matrix $A^N$. In any case, the stability is not easy to be guaranteed.

Stochastic disturbances can be considered in the observer scheme defining the gain $L$ as a Kalman filter. Assume that the system equations are of the form (18), where $v_k$ and $\theta_k \in \mathcal{R}$ are uncorrelated white noise signals of zero mean and variance $\sigma_v$ and $\sigma_{\theta}$ respectively. Even if some measurements are missing, the standard recursive Kalman filter can be used to obtain an optimally filtered state estimate. The missing outputs are equivalent to a measurement with infinite variance. The effect of the unmeasured outputs can be modelled by a disturbance $\eta_k$ of time varying variance, $\sigma_{y}(k)$. Defining the measured output as $y_m$, the equation is $y_m(k) = x_k + \eta_k$, where $\sigma_{y}(k) = \sigma_v$ at the measuring instants and $\sigma_{y}(k) = \infty$ otherwise. The infinite variance represents the absence of measurement, i.e. a null knowledge of the output at that instant. The observer should be now

$$\dot{x}_k = \Phi^{k} x_{k-1} + u_{k-1} + \eta_{y}(k) \Phi_{x\theta}(k) - (A^{N} x_{k-1} + u_{k-1},)$$

The gain of the Kalman filter that gives the optimum state estimate is given by the well known equations:

$$Q(k+1) = \Phi(k) A^T + \tau \sigma_v \gamma^T$$

$$P(k+1) = \gamma(k+1) e^{T} (Q(k+1) e + \sigma_{\theta} k + 1)^{-1}$$

The previous equations define a time varying observer gain. When a measurement is not available, $L(k)=0$. This is the approach used in (Isaksson, 1994) to predict the missing outputs to implement a recursive parameter estimation algorithm. If the sampling is regular, with constant $N$, it is possible to obtain a constant gain $L$ to be applied when data are available assuming the system is in stationary state, either by repeated iteration of the
above equations or, equivalently, by solving:

\[
Q = 4^N PA^{N^T} + \Phi W F^T
\]

\[
L = \frac{1}{\sqrt{c}} \Phi Q c^T + \sigma_{\nu}^{-1}
\]

\[
P = \phi Q
\]

where

\[
F = A^{N+1} g \ldots A g \; \; \; g \cdot \Omega_{k-1} = \omega \cdot \; \; \ldots \; \; \omega_1 \cdot \;
\]

\[
W = \frac{1}{\sqrt{\Omega}} \cdot \Omega \cdot \; \; \Omega^T = \Lambda \gamma \sigma, \ldots, \sigma_{y_k, y_t}
\]

In order to compare the performance of the linear observer and the algorithm described in section 3, let us consider the system whose ZOH discrete equivalent is \(G(z)\) but, due to modelling errors the available model is \(\hat{G}(z)\):

\[
G(z) = \frac{0.012g^- + 1.011g^-}{1 - 0.8g^- + 0.82g^-}, \quad \hat{G}(z) = \frac{0.012g^- + 1.011g^-}{1 - 0.8g^- + 0.822g^-}.
\]

The input to the system is a known sinusoidal signal of amplitude 1. An output disturbance of \(\sigma_y = 0.05\) and a state disturbance of \(\sigma_u = 0.05\) are assumed. The prediction error in steady state for both algorithms with the eigenvalues assigned to the origin is shown in figure 2. The performances are almost identical. The transient behaviour of the state observer (not shown in the figure) is worse than that of the mixed regression algorithm (peak error of 0.8 and 0.4 respectively).

The prediction error variance can be reduced in both algorithms by adequately choosing the eigenvalues of the predictors, as shown in figure 3 (the linear observer with poles assigned at 0.35 has a very similar stationary behaviour than the mixed regression predictor). The transient deteriorates, but is still better for the mixed regression (peak of 0.7 against 1.5).

In this example, the optimal stationary Kalman filter does not improve the performance obtained by pole assignment, due to the modelling error, as shown in figure 3.

By these simulations it is shown that both predictors can attain similar performances on the transient responses and disturbance filtering. The main difference is that the linear state observer is easier to design (for example by pole placement), but implies a higher computational cost on the implementation.

Another possible method to estimate the output consists of reconstructing the state from past measurements and then to simulate forward the system equations. This is almost always possible for an observable system whose order is \(n\), so that an expression for the output at instant \(k\) can be obtained in terms of the last \(n\) measured outputs and all the inputs from instant \(k-I\) till the sampling period of the oldest of those \(n\) measurements.

The advantage with respect the algorithm described in section 3 is that only real measurements are used in this case, so the observer is always stable (in fact, it has a dead-beat behaviour). Notwithstanding, there is a
considerable increase of the computational load. Another problem of this approach is the lack of robustness against modelling errors or disturbances, which are not taken into account in the reconstruction.

If the estimation of the outputs is to be carried out off-line, other approaches are possible. In (Isaksson, 1993) the fixed-interval smoothing algorithm is described. It consists of running a Kalman filter forward in time followed by a fixed-point smoother backwards.

5. PARAMETER ESTIMATION UNDER SCARCE DATA OPERATING CONDITIONS

These output prediction algorithms can be used as a basis to modify the standard recursive parameter estimation algorithms to deal with missing data situations. The disturbed basic system (2) can be expressed by the difference equation:

\[ y_k = y_{-}^{T} \theta \cdot v = y_{-}^{T} (k-1)^{T} \theta + y_{-}^{T} (k-1)^{T} \theta + v \]  

(22)

where \( y_{k} \) is the output of an stochastic process, \( \theta_{k} \) are parameter vectors defined as in (6), and \( y_{k} \) (regression vector) is: \( y_{k} = \psi_{k} = y_{-} \ldots y_{-(k+1)} u_{k} \ldots u_{-(k+1)} \).

If \( y_{k} \) is a white noise and persistent excitation conditions for the input hold, \( \theta \) can be estimated from a set of observations by LS and RLS algorithms (Ljung, 1987) with full output availability and with non-scarce missing data patterns.

The modification of the standard RLS algorithm to account for the missing data can lead to a very simple Pseudo-RLS algorithm that is valid for on-line parameter estimation. In this scheme, the unknown regression vector elements of the usual RLS are substituted by an estimation of them. The general form of this algorithm is:

\[ y_{k} = y_{-}^{T} \theta \cdot v = y_{-}^{T} (k-1)^{T} \theta + y_{-}^{T} (k-1)^{T} \theta + v \]  

(23a)

\[ \gamma_{k} = \frac{P_{k}}{\lambda + \gamma_{k}^{T} P_{k} \psi_{k}} \]  

(23b)

\[ \theta_{k} = \theta_{k-1} + \gamma_{k} (y_{k} - y_{-}^{T} \theta_{k-1}) \]  

(23c)

\[ P_{k+1} = \frac{1}{\lambda} (I - \gamma_{k} \psi_{k} \psi_{k}^{T}) P_{k} \gamma_{k} + \frac{\lambda}{\lambda} (1 - \gamma_{k}) \]  

(23d)

where the factor \( r_{k} \) defines the data availability (1 if there is a measurement and 0 if not), and the function \( f \) (missing-data predictor) represents the output estimation method as a function of the model parameters, past output estimates and present and past input and output measurements. Various possibilities for that predictor have been discussed in the previous sections. In order to achieve P-RLS estimation convergence the predictor should be not only stable but quick enough in comparison with the parameter estimate dynamics. P-RLS estimators may be fooled by thinking that
its RLS prediction error is caused by a model error (thus modifying the parameter vector) whereas it is actually caused by a slow convergence of predictions to accurate ones from non-zero error initial conditions. For regular data availability, the stability of the predictor can be guaranteed with the state-space pole-placement observer (20) or the predictor defined in (7), as discussed in section 3, resulting in this case a simpler final algorithm. In any case, the convergence of the complete parameter estimation scheme is difficult to be guaranteed.

6. OTHER APPLICATIONS.

The proposed output predictor and parameter estimation algorithms can have multiple applications in missing data situations, such as the control with scarce measurements. In many industrial applications, the output of the system can not be measured at the desired sampling rate due to the use of slow sensors. If a simple digital controller (such as a PID) is to be used then it must work at the slow output sampling rate with the consequent loss of performance. The output estimator described in section 3 can be used to implement the digital controller at a faster rate with no significant increase on the computer cost, by using the predictor to generate estimates of the output at the fast rate (working in open loop between measurements). The block diagram is shown in Fig. 4.

The following simulation example will illustrate the applicability of the above control scheme. Consider the system with transfer function \( G(s) = \frac{5}{s^2 + 4s + 3} \), and assume that the output is available every 0.9 s. A white noise disturbance at the input and a white measurement noise of 2% will also be assumed. A digital PID controller is defined as

\[
\begin{align*}
    u(k) &= \rho(k-1) + \gamma_0 \epsilon(k) + \gamma_1 \epsilon(k-1) + \gamma_2 \epsilon(k-2) \\
    q_0 &= \zeta_p \left( \frac{T_d}{T} \right) ; q_1 = -\zeta_p \left( \frac{T_d}{T} - \frac{T}{T_d} \right) ; q_2 = \zeta_p \frac{T_d}{T}
\end{align*}
\]

At 0.9 s sampling time, the best dynamic performance that can be achieved is obtained by setting the controller constants to \( K_p = 0.9, T_d = 0.2, T_i = 1.6 \). If the controller is defined at \( T = 0.3 \) s with full measurements, then the step response that can be obtained (for \( K_p = 1.6, T_d = 0.2, T_i = 1.4 \)) is much better.

In order to approximate the fast rate response in the low sampling rate case, the missing-data predictor described in section 3 is used to estimate the intersampling measurements to apply the controller at the high rate. An extended order model has been used, the additional factor being \( E = 1 + 0.95 q^{-1} + 0.7 q^{-2} \), so the eigenvalues of the prediction error (13) are \( 0.53 \pm 0.1j \). The step response is very similar to the full measurements case, as shown in figure 5. The behaviour with respect the disturbance is similar for the three cases, i.e. the differences in the variances of the output in steady state are not significant.
7. CONCLUSIONS

In this paper, some predictors for estimation of the unmeasured outputs in missing-data situations are studied. The main result is the presentation of a very simple algorithm that uses the difference equation of the process, replacing the unknown past values with estimates when necessary. The analysis of convergence has been carried out for the case of regular but scarce sampling pattern. By extending the order of the difference equation by means of a polynomial an adequate error dynamics can be achieved without sacrificing simplicity. Effects of measurement as well as state disturbances have also been considered. More complex algorithms based on state reconstruction or observation are also analysed.

The application of the output predictors to system identification and control has also been outlined. Standard RLS techniques can be applied in scarce-data situations leading to Pseudo-RLS parameter estimation algorithms. Classical digital controllers (such as PID) are implemented in the case of missing-data or slow sensors with a very low increase in computer cost.

Acknowledgements. The authors wish to thank the reviewers for their helpful comments and suggestions.

8. REFERENCES


Fig. 1
Fig 5.

The graph shows the response over time for different time constants.

- T=900ms
- T=300ms
- T=900ms with predictor

The y-axis represents the response variable, and the x-axis represents time in seconds.
FIGURE CAPTIONS

Fig. 1. Output predictor block diagram.

Fig. 2. Prediction error in steady state.

(---) Linear observer. Eigenvalues (0,0).

( ) Mixed regression alg. : $\beta_1=0.61$, $\beta_2=0.278$. Eigenvalues (0,0).

Fig 3. Prediction error in steady state.

(---) Linear observer. Optimal stationary Kalman filter.

( ) Mixed regression alg. $\beta_1=0.9$, $\beta_2=0.75$. Eigenvalues (0.645±0.05j).

Fig. 4. Missing-data control structure.

Fig. 5. Step response of controlled system.