Dynamic reconstruction of instrumented rocket motor data

D. H. Smith*

(Received 8 August 2003, revised 19 March 2004)

Abstract

Data from instrumented rocket motors is subjected to dynamic reconstruction by time delay embedding, with a focus on local linear and global nonlinear prediction, and associated parameter estimation.

Contents

*DSTO Edinburgh, PO Box 1500, Edinburgh, South Australia 5111.
mailto:david.h.smith@dsto.defence.gov.au

See http://anziamj.austms.org.au/V45/CTAC2003/Smi2/home.html for this article, Published April 3, 2004. ISSN 1446-8735

1 Introduction

Measured time series data contains valuable information about the dynamics of an underlying process which can be extracted by time delay embedding techniques, with the potential to obtain fundamental properties and invariants [1]. On this basis, nonlinear predictive models can be constructed from the data, as demonstrated by Casdagli for a number of test systems [5]. This study is concerned with such models for the case of a thermally forced system in which forcing and response data are both available. Key aspects of the model construction are examined, including local linear and global nonlinear approximation, with particular attention paid to the singular value decomposition and its vital role in the overall process.

2 Dynamic reconstruction from data

2.1 Time delay embedding

For a measured data series, $(y_0, y_1, \ldots, y_j, \ldots, y_N)$, time delay embedding proceeds by constructing a pseudo phase space of time delay vectors [1]

$$\mathbf{x}_k = [y_k, y_{k+\tau}, \ldots, y_{k+(d-1)\tau}]^T,$$

which are characterised by two parameters, dimension d and delay τ. These are considered to evolve under a nonlinear discrete map

$$\mathbf{x}_{k+\tau} = [y_{k+\tau}, y_{k+2\tau}, \ldots, y_{k+d\tau}]^T = \mathbf{f}(\mathbf{x}_k)\,, \tag{1}$$

to generate a *trajectory matrix* with columns comprising the delay vectors arranged in chronological order. This constitutes a projection of the unknown system state vector onto a finite dimensional subspace, the dimension of which needs to be sufficiently large in order to properly capture the dynamics. According to the Takens embedding theorem [11], $d \geq 2m+1$ is sufficient if the true system attractor resides within an m dimensional manifold, allowing access to the system dynamics via delay coordinates, with the attractor manifold embedded in $\mathbb{R}^d$. If d is too small, nearest neighbours in the reconstructed phase space will be close due to projection, not dynamics, a fact underlying the false nearest neighbour criterion for determining embedding dimension [3]. Choice of time delay τ rests on information theoretic criteria, independent of the topological considerations associated with embedding dimension [1]. A basic guiding principle is the requirement that information in two successive delay coordinates be as independent as possible, without introducing excessive delay. Too small means too much correlation and too large means insufficient causal connection, with useful guidance provided by consideration of the quantity average mutual information [7].

In the laboratory experiment context, forcing is often applied to a system under study to induce particular behaviour for analysis. When the forcing data is also available, an extension of Takens theorem [4] involves delay vectors augmented with the input data u_j in the form

$$\mathbf{x}_k = [y_k, y_{k+\tau}, \ldots, y_{k+(d-1)\tau}, u_k, u_{k+\tau}, \ldots, u_{k+(d-1)\tau}]^T$$

which also evolve under a discrete map similar to (1). The particular data under consideration derives from thermal forcing applied to solid rocket motors containing internal sensors and is amenable to this type of reconstruction. Another attribute of the data requiring attention is the presence of irregular

sampling, which is dealt with by utilising "fuzzy" delay coordinates [2], in which successive values are chosen with spacing as close as possible to the nominal delay τ, without applying interpolation.

2.2 Local linear approximations

Given suitable trial values for dimension and delay, a useful first step in the reconstruction process is to consider a linearisation of the nonlinear map (1)

$$\mathbf{x}_{k+\tau} = \mathbf{f}(\mathbf{x}_k) \approx A_k \mathbf{x}_k + \mathbf{b}_k \tag{2}$$

where the local Jacobian A_k and vector $\mathbf{b}_k$ are estimated from the available data, specifically neighbours residing inside a ball in the pseudo phase space centred at $\mathbf{x}_k$, exclusive, and their images under the map [9]. In this sense the accuracy of (2) depends on the distribution of neighbours around $\mathbf{x}_k$, which in turn depends on the amount of available data. Ideally, many neighbours are required inside a small ball; however, as the number of neighbours available for estimation shrinks with ball radius a compromise must be met. This study will use a fixed number of neighbours, just enough to give an overdetermined system amenable to least squares coefficient estimation, resulting in variable ball radii and accuracy as the trajectory matrix is traversed.

2.3 Global nonlinear approximations

Appending nonlinear terms to (2) and considering element d of the nonlinear map, containing the new response value one step ahead

$$y_{k+d\tau} \approx b + \mathbf{a}^T \mathbf{x}_k + \sum_{j=1}^{M} \alpha_j \phi_j(\mathbf{x_k}) \tag{3}$$

where $\phi_j(\mathbf{x})$ comprise a set of basis functions. Choosing an estimation set from the trajectory matrix and assembling the equations (3) produces an

overdetermined set of linear equations with $m = 2d + M + 1$ columns

$$\mathbf{y} \approx A\theta \tag{4}$$

where the columns of A represent constant, linear and nonlinear terms, with number of rows n depending on the chosen estimation set. A convenient choice for ϕ_j is the family of radial basis functions [5], defined with respect to a set of centres $\mathbf{c}_j$, taken from the data,

$$\phi_j(\mathbf{x}_k) = \phi(\|\mathbf{x_k} - \mathbf{c}_j\|)\,,$$

where ϕ may take a variety of possible forms.

2.4 Coefficient estimation

Coefficients for the local and global approximation processes are estimated from the data via two variants of least squares estimation utilising the singular value decomposition [6]. Considering the over-determined system (4), the unique least squares minimiser is defined by

$$\theta^* = VS^{-1}U^T\mathbf{y} = \sum_{j=1}^{m} \frac{\mathbf{u}_j^T\mathbf{y}}{\sigma_j}\mathbf{v}_j \tag{5}$$

where the left singular vectors $\mathbf{u}_j$, singular values σ_j and right singular vectors $\mathbf{v}_j$ reside in the matrices U, S and V that constitute the decomposition

$$A = USV^T\,.$$

From the expansion (5), destabilisation effects of small singular values, or ill-conditioning, become immediately evident, obviating the need for regularisation. Retaining the first k terms of (5) yields a family of truncated least squares solutions, θ_k, $k = 1, \ldots, m$, possessing varying degrees of regularisation corresponding to removal of small singular values. L curves [8] neatly

unfold this process by considering parametric plots of solution norms against their residual counterparts, $\mathbf{r}_k = \mathbf{y} - A\theta_k$,

$$\mathbf{r}_k^T \mathbf{r}_k = \mathbf{y}^T \mathbf{y} - \sum_{j=1}^{k} (\mathbf{u}_j^T \mathbf{y})^2 , \quad \theta_k^T \theta_k = \sum_{j=1}^{k} \left(\frac{\mathbf{u}_j^T \mathbf{y}}{\sigma_j} \right)^2$$

indicating a balance between the conflicting demands of data/noise reproduction and smoothness by a characteristic "elbow", where regularisation error loses its dominance [6].

The LS formulation effectively treats the matrix A as error free, and more appropriate in the current context is the concept of total least squares [6], based on SVD of the augmented matrix

$$\begin{bmatrix} A & \mathbf{y} \end{bmatrix} = \bar{U} \bar{S} \bar{V}^T .$$

Truncated TLS solutions, filtered versions of their LS counterparts, involve a family of trios $(\hat{A}_k, \hat{\theta}_k, \hat{\mathbf{y}}_k)$, with $\hat{\theta}_k$ constructed from the final $m + 1 - k$ columns of $\bar{V}$, or null space basis for the augmented matrix, satisfying

$$\hat{A}_k \hat{\theta}_k = \hat{\mathbf{y}}_k .$$

L curves are also applicable, with residual norms now comprising contributions from the matrix and right hand side.

3 Application to data

Figure 1 gives temperature input and response data comprising 7270 points over 190 days, with irregular sampling and time normalised with respect to the data length [10]. Based on false neighbour and mutual information calculations, trial delay and dimension values $(d, \tau) = (3, 0.08)$ will be used for the subsequent reconstruction, producing a trajectory matrix with 1761 columns.

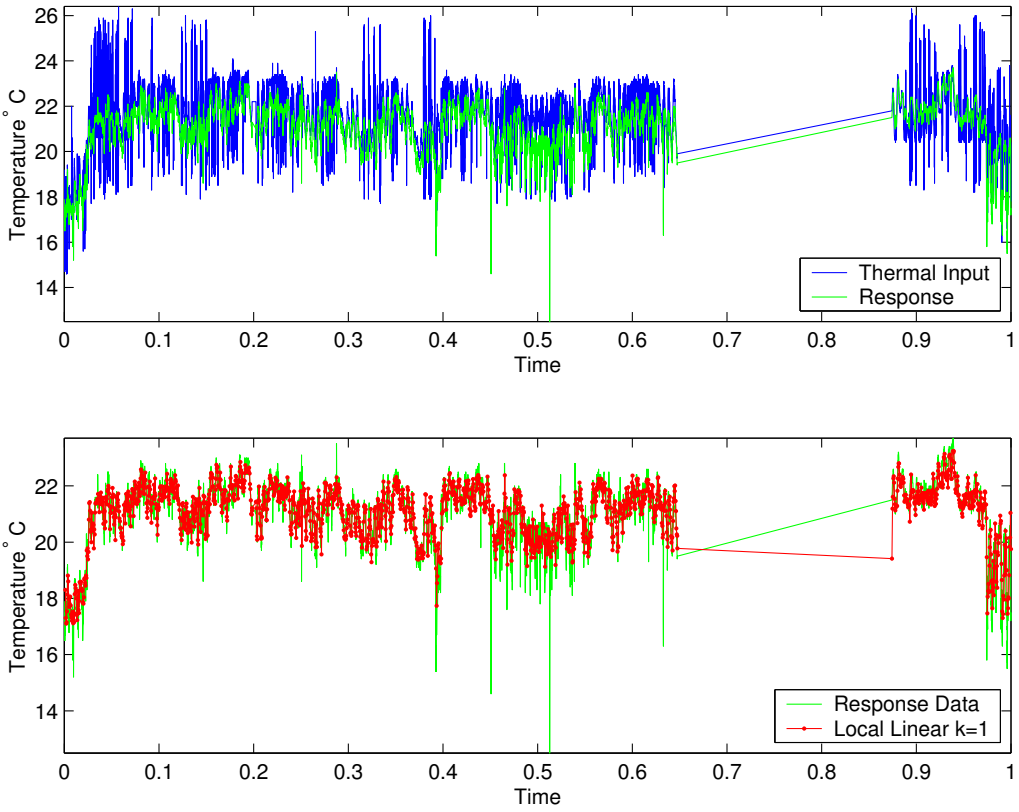

FIGURE 1: Temperature input and response data series, comprising 7270 points with irregular sampling. Local linear response prediction with $(d, \tau) = (3, 0.08)$ and SVD truncation $k = 1$ is also shown alongside the data.

It is emphasised that this serves for demonstration purposes and a more comprehensive study demands comparisons across a range of delay and dimension values. For this particular choice of embedding parameters, traversing the trajectory matrix and applying local linear prediction produces singular value spectra featuring a distinct gap separating a single dominant value from its counterparts, indicating a rank one tendency. Results of this prediction are also shown in the Figure alongside the true data, demonstrating its smoothing and spike reduction capacity.

Global nonlinear prediction is now considered with radial basis functions $\phi = r$ on an estimation set comprising the first half of the trajectory matrix, with centres taken from the whole matrix at equal intervals, producing a regression matrix with 879 rows and singular value spectra shown in Figure 2 for 16, 32 and 64 centres. In each case there is an initial steep decrease followed by a slower decay, with all values sufficiently large to avoid ill-conditioning threats, as confirmed in the accompanying right singular vector coefficient diagram. Increasing the number of centres lifts the curves slightly and slows the decay, while the absence of a distinct spectral gap is the characteristic signature of discrete ill-posed problems [6]. Associated L curves for LS and TLS approximation are shown in Figure 3, illustrating key differences between the two formulations. For the LS case there is a large residual drop between $k = 1$ and 2 at the start of the curve followed by relatively slow decay and little solution growth while the TLS case shows a steady residual decline at the expense of larger solution coefficients. With respect to the number of centres, the LS curves shift towards the origin as this increases, producing only marginal residual improvements and solution growth while TLS offers far more substantial residual reductions.

To complement these L curves, a useful visualisation of the fitting procedure is provided by the scatter plots of Figure 4, which show predicted values against their true data counterparts in the estimation set for 16 RBF centres at various truncation levels, with residuals manifested as departure from straight line behaviour. The plots represent selected points from the

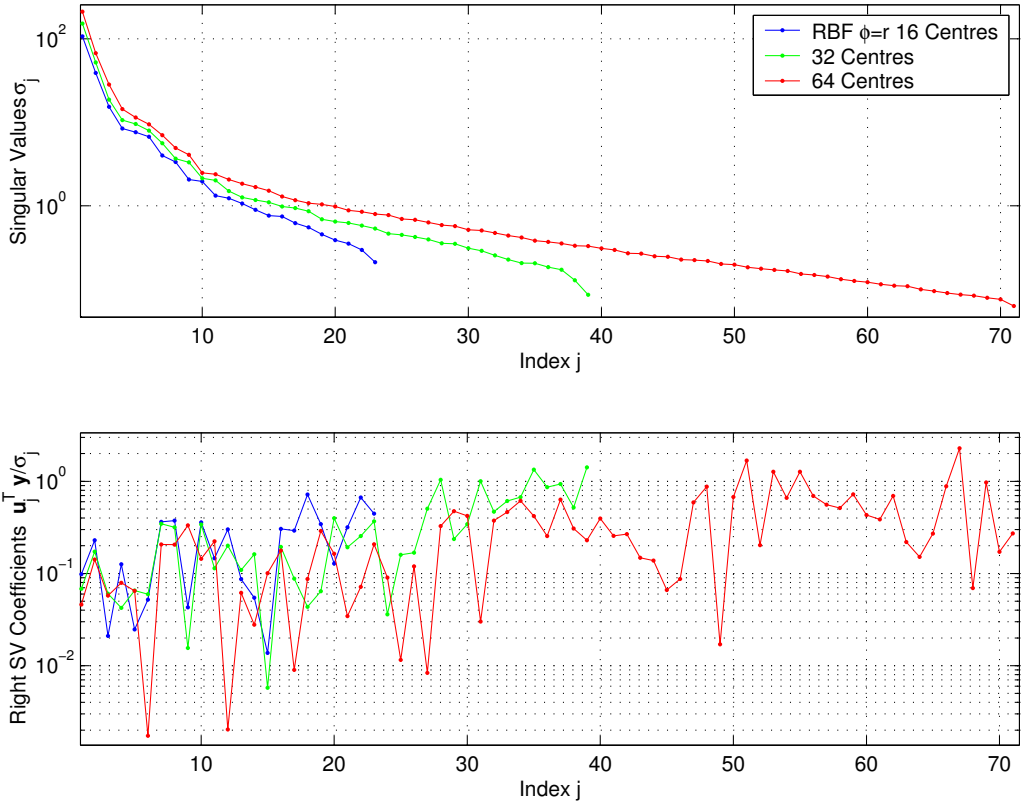

FIGURE 2: Singular value spectra and associated solution coefficients from (5) for RBF approximation with $\phi = r$ on an estimation set comprising the first half of the trajectory matrix and various numbers of centres.

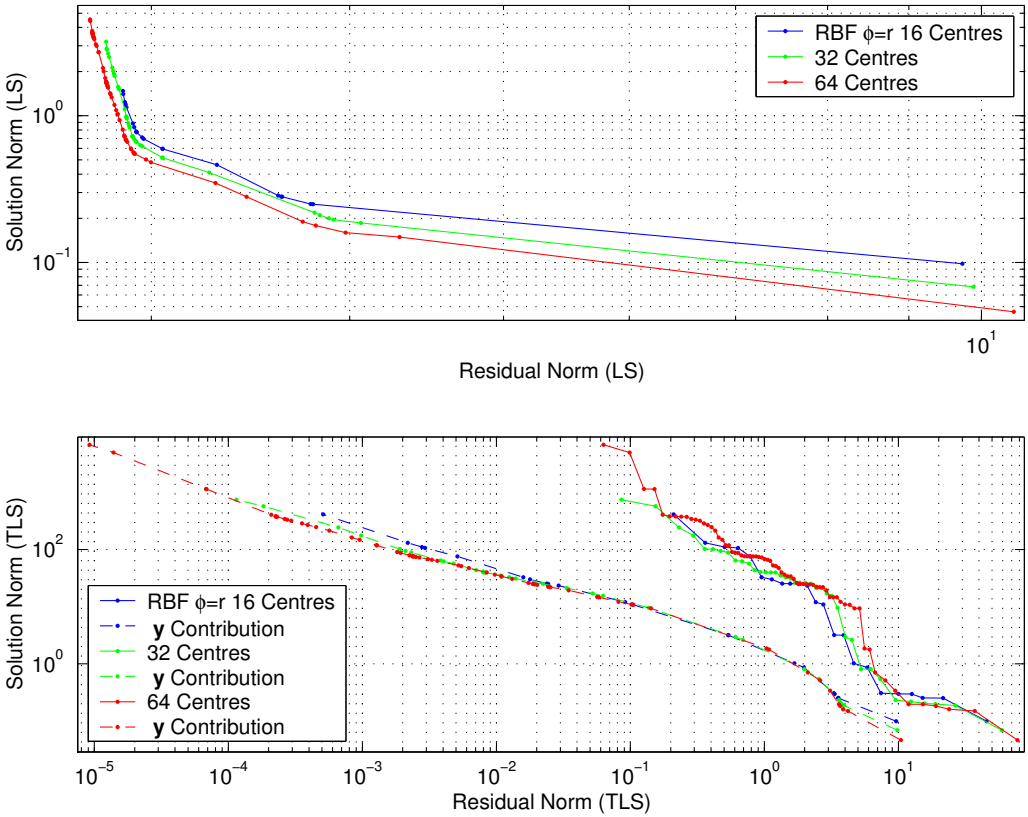

FIGURE 3: L curves for LS and TLS corresponding to Figure 2, showing LS residual stagnation while TLS declines at the expense of a large solution.

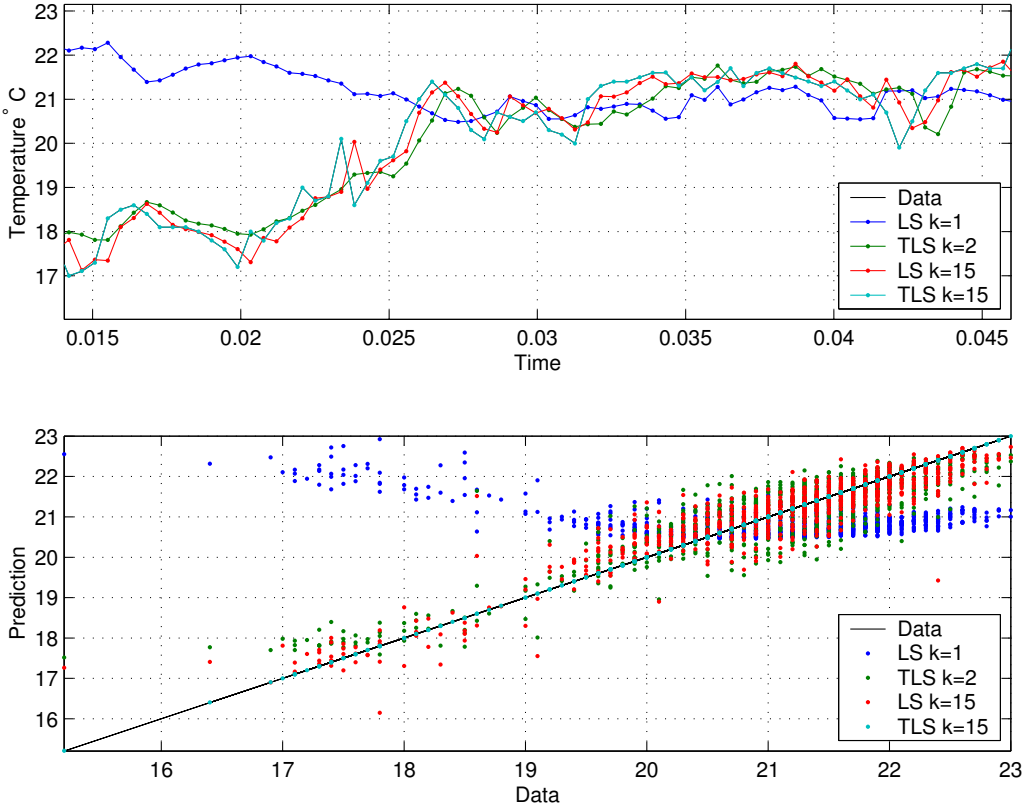

FIGURE 4: Scatter plots of predicted values against their data counterparts in the estimation set, accompanied by a segment of the corresponding time domain data, showing convergence of the TLS result and LS stagnation.

corresponding L curves in Figure 3, and illustrate superior convergence behaviour of the TLS result as k is increased while its LS counterpart stagnates after a few terms. In addition, the appearance of distinct vertical bands at regular intervals reveals a fundamental property of the measurement device not immediately evident in the time domain. Accompanying the scatter plot is a segment of the corresponding time domain picture, in which the TLS prediction at $k = 15$ is indistinguishable from the data, unlike the LS result.

Moving outside the estimation set, out of sample prediction is considered from both recursive and non-recursive viewpoints, the latter of which constitutes one step ahead prediction. Figure 5 gives a selection of these results at different truncation levels, showing a stark difference between the two approaches, with one step predictions tracking the data far more effectively. For the recursive calculations, LS and TLS results at high truncation, or small k, produce similar behaviour as shown for $k = 2$ and 5, however at larger values the bigger TLS solutions, as indicated by their L curves, produce instability and divergent behaviour while the LS results remain stable.

4 Summary and conclusions

Local linear and global nonlinear prediction based on time delay embedding have been applied to measured data from a thermally forced solid rocket motor, yielding satisfactory results from a small number of basis functions relative to the data size, and demonstrating the SVD as a valuable tool. Comparisons between LS and TLS coefficient estimation for the global case have shown very similar behaviour under heavy truncation, near the start of the L curves, while the TLS result displayed a tendency for strong coefficient growth as the truncation is reduced, with the potential to destabilise recursive out of sample prediction. Much scope exists for further investigation with respect to embedding parameters, local prediction variants, alternative nonlinear basis functions, and application to other data sets.

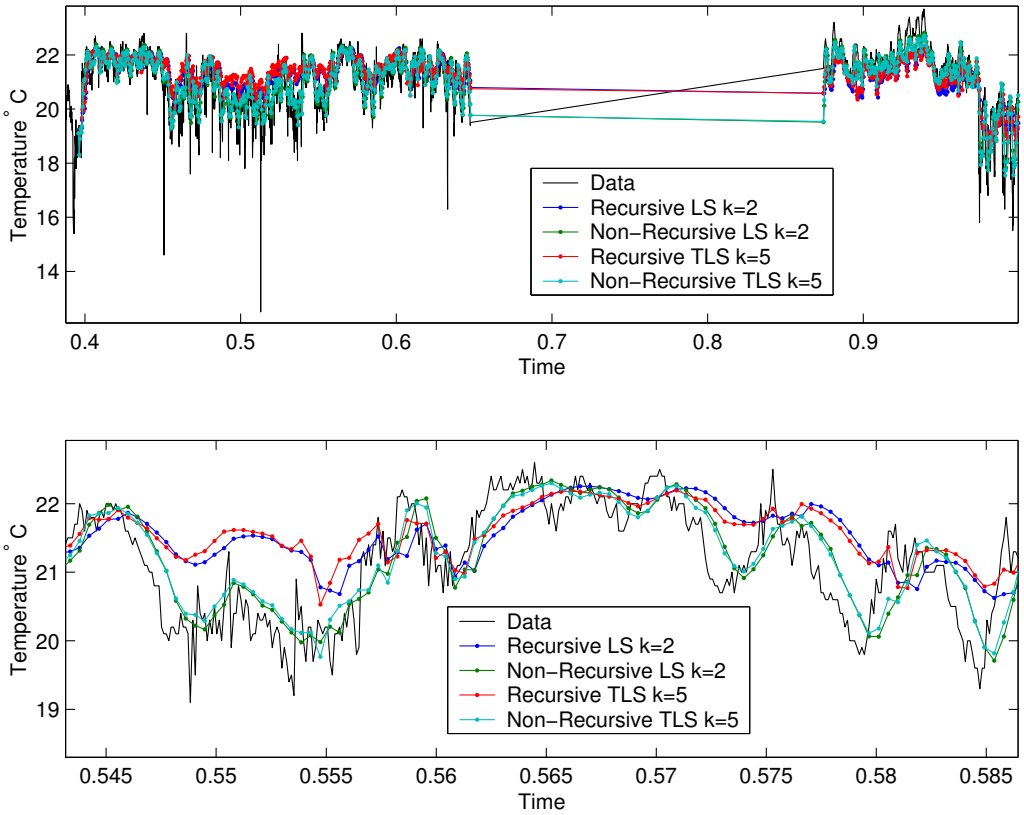

FIGURE 5: Out of sample recursive and non-recursive nonlinear prediction in the second half of the trajectory matrix at various SVD truncation levels.

References

[1] H. D. I. Abarbanel et al., "The analysis of observed chaotic data in physical systems", Rev. Mod. Phys. **65**, 4 p1331–1392 (1993). C121, C122

[2] J. L. Breeden and N. H. Packard, "Nonlinear analysis of data sampled nonuniformly in time", Physica D **58** p273–283 (1992). C123

[3] L. Cao et al., "Determining the minimum embedding dimension of input-output time series data", Int. J. Bif. Chaos **8** 7, p1491–1504 (1998). C122

[4] M. Casdagli, "A Dynamical Systems Approach to Modeling Input-Output Systems", in "*Nonlinear Modeling and Forecasting*", eds M. Casdagli and S. Eubank, Addison-Wesley (1992). C122

[5] M. Casdagli, "Nonlinear Prediction of Chaotic Time Series", Physica D **35** p335–356 (1989). C121, C124

[6] R. D. Fierro, G. H. Golub et al., "Regularisation of Truncated Total Least Squares", SIAM J. Sci. Comp. **18** 4 p1223–1241 (1997). C124, C125, C127

[7] A. M. Fraser, "Information and Entropy in Strange Attractors", IEEE Trans. Inf. Theory **35** 2 p245–262 (1989). C122

[8] P. C. Hansen and D. P. O'Leary, "The Use of the L-Curve in the Regularization of Discrete Ill-Posed Problems", SIAM J. Sci. Comp. **14** 6 p1487–1503 (1993). C124

[9] E. Kostelich, "Problems in estimating dynamics from data", Physica D **58**, p138–152 (1992). C123

[10] P. Macdowell, "New Service Life Methodologies for Solid Propellant Rocket Motors", DSTO-RR-0099 (1999). C125

[11] F. Takens, "Detecting strange attractors in turbulence", in Dynamical Systems and Turbulence, eds. D. A. Rand and L. S. Young, Lecture Notes in Mathematics, Volume 898, Springer (1981). C122