Application of Nonlinear Time Series Analysis to the Prediction of Silicon Content of Pig Iron

Matias WALLER and Henrik SAXÉN

Lic. Tech. (Chemical Engineering), Doctoral Student, Heat Engineering Laboratory, Åbo Akademi University, Biskopsgatan 8, FIN-20500 Åbo, Finland. E-mail: matias.waller@abo.fi
1) Heat Engineering Laboratory, Åbo Akademi University, Biskopsgatan 8, FIN-20500 Åbo Finland.
E-mail: henrik.saxen@abo.fi

(Received on September 28, 2001; accepted in final form on November 30, 2001)

1. Introduction

Applying nonlinear time series analysis to data from industrial processes can serve several purposes; often the goal is to develop a model with predictive properties superior to those of a linear model, but the analysis can also serve as a means for gaining deeper understanding of the dynamics exhibited by the process. By applying some methods presented in two recent monographs, the current study examines the deterministic nature of the blast furnace and whether the dynamics, based on observations of hot metal silicon content, can be ascribed to a nonlinear source. A comparison is made between the predictive properties of nonlinear time series models implemented as radial basis function networks and (time variant) linear time series models, as well as parametrically linear (time variant) finite impulse response models.

2. Data Sets and Methods

For the current study, data extending over a one-year period from the two blast furnaces of Rautaruukki Steel, Raahe, Finland, is used. Both are one taphole furnaces with working volumes of approximately 1000 m³, using coke and oil as reducing agents and as iron-bearing materials sinter and pellets. Blast furnace 1 (BF1) is equipped with a bell-top while a bell-less charging equipment is used for blast furnace 2 (BF2). As the source for observations, ladle-wise analyses of the pig iron silicon content are considered, i.e., 1–10 observations of the output for each tap. Since the objective of time series analysis is to study, and predict, the evolution of a signal, it is necessary to order the observations. Due to unknown, varying, or not easily determined flow conditions in the hearth of the furnace, the choice of ordering is not obvious. However, the simple approach that assumes that the outflow of hot metal follows the order in which it was produced was adopted here, partly motivated by the fact that it facilitates the implementation of the models on-line. With a more detailed description of flow conditions in the hearth at hand, a more realistic and physically motivated ordering could possibly be employed. However, it appears as if the signal, in a statistical sense, quite well can be described in the discrete-event domain, with the events being the ladles of hot metal in the order they were filled. The “time” series thus obtained serve as the case study of the present investigation.

Since the pig iron silicon content is controlled, it could, apart from shorter segments (e.g., 10 taps), be expected that it is stationary even though time variant considerations are important for successful predictions with parametrically linear models. Due to possible variations on longer time scales, it is still important to statistically test the signal for stationarity. The applications of recurrence plots and a simple predictive scheme revealed no evidence against stationarity, so methods for stationary time series analysis were applied.

As a test for nonlinearity, the method of surrogate data is applied, using:

\[ S(k) = \sum_{i=1}^{N} \frac{\sum_{t=1}^{N} (y(t)-y(t-k))^2}{\sum_{t=1}^{N} (y(t)^2)} \quad \text{for } k = 1, 2, \ldots, d \]

as a discriminating statistic, which requires no intricate choices of parameters and reflects the time irreversibility of the signal—a strong signature of nonlinearity. With a collection of 39 surrogate data sets, this indeed reveals that with \( \alpha = 95\% \), both time series studied appear to stem from a nonlinear source.

Most of nonlinear time series analysis is performed in what is known as the reconstructed phase space. The time delayed values of the scalar time series \( \{y(t)\} \) are typically used for the reconstruction

\[ z(t) = (y(t-(d-1)\tau), y(t-(d-2)\tau), \ldots, y(t)) \]

where \( d \) is the dimension of the reconstructed state space, i.e., the embedding dimension, and \( \tau \) is the delay. The celebrated embedding theorems state that the time delay embedding provides a one-to-one image of the original state space \( x \) if \( d \) is large enough.

2 The choice of \( d \) and \( \tau \) are, however, crucial for the success of practical implementations; \( \tau \) should be large enough so that \( y(t-\tau) \) contains new information compared to \( y(t) \), but \( \tau \) should not make them completely independent. A common guideline is to choose the delay that corresponds to the first minimum of the average mutual information. For a given delay, the concept of false nearest neighbours is commonly evaluated as a means to obtain a suitable dimension.

3. The Case Study

The rather infrequent measurements of the silicon content suggest a short delay, \( \tau \), and this was verified by exam-

\[ *1 2/(1-\alpha) - 1 \] is the number of surrogate data required for a two-sided test (i.e., the value obtained for the original signal can be larger or smaller than all other values) where \( 1-\alpha \) is the significance level for an erroneous result.

\[ *2 \] The theorems assume that the measurement function \( y = h(x) \) is continuously differentiable and that there are no periodical orbits with a fundamental period of \( t \) or \( 2t \) where \( t \) is the sampling interval and only a finite number of periodical orbits with fundamental period \( pt, p > 2 \).
in the average mutual information yielding $1 \leq \tau \leq 3$. For these different delays the ratios of false nearest neighbours were calculated suggesting $d=4$ or $d=5$, but the analysis also implied quite noisy data. Thus, these $\tau$ and $d$ were merely used as guidelines for further analysis.

An interesting possibility in nonlinear time series analysis is to examine the data for more qualitative characteristics, e.g., estimate Lyapunov exponents and correlation dimensions. A critical evaluation of estimates based on time series is essential\(^\text{[1]}\) and apparently several erroneous interpretations have been reported. For the data at hand, reliable estimates could not be obtained for neither a maximal Lyapunov exponent nor a correlation dimension. The nature of the measurements is a possible reason for the failure: The infrequent measurements, or the uneven distributions in the time domain, might form an inadequate description of the relevant dynamics. The manual control by experienced supervisors could also render it more fundamentally stochastic; applied subjective state evaluation cannot easily be quantified by deterministic models.

Given an appropriate dimension $d$ and delay $\tau$ for reconstructing the phase space, vast possibilities for constructing empirical nonlinear models still remain. A commonly used approach is to minimise the $(k\text{-step ahead})$ prediction error,

$$
\epsilon(k) = \hat{y}(t+k) - y(t+k),
$$

with respect to some function $h: \mathbb{R}^d \rightarrow \mathbb{R}$, i.e., \( \hat{y}(t+k) = h(z(t)) \).

A popular choice for $h(z(t))$ is an expansion that utilises radial basis functions (RBF), $h(z) = \sum_{i=1}^{m} w_i \phi(||z-c_i||)$, and it is often termed radial basis (function) network (RBFN). Typically, an RBFN is defined by the function ($\phi$), the number, $m$, of functions, the width of the functions (the norm), the centers, $c_i$, for each respective function and the weights, $w_i$, in the expansion. Although methods for a full optimisation with respect to the prediction error of all parameters defining the network have been suggested\(^\text{[13]}\), it is questionable whether such a difficult and computationally expensive task has any advantages over more “traditional” nonlinear methods, e.g., multilayer perceptron neural networks. With respect to the weights, however, the output is a linear function of the input, and the corresponding simple optimisation is a reason for using an RBFN. In order to avoid a nonlinear optimisation, the centers are often chosen to represent the input space by, e.g., using clustering algorithms\(^\text{[14,15]}\) or as a suitable subset of a large set spanning the input space.\(^\text{[16,17]}\) The width of the functions can be determined according to nearest neighbour heuristics,\(^\text{[14]}\) or alternatively a few different values can be evaluated.

The most promising networks, using the methods discussed above (excluding the full optimisation), were thus obtained by forward selection\(^\text{[16]}\) and use $\tau=1$, $d=3$ and $m=22$ (24) for BF1 (BF2). The variance of the prediction errors normalised by the variance of the data for the used test segment (“hidden” from the network during training) is given in Table 1.

| Type of model | RBFN | ARMA | FIR |
|---------------|------|------|-----|
| Blast furnace 1 | 0.641 | 0.815 | 0.667 |
| Blast furnace 2 | 0.262 | 0.268 | 0.509 |

Fig. 1. Prediction of pig iron silicon content in BF2 using RBFN (thick line) and ARMA (thin line) modelling. Actual observations are indicated by circles.

Fig. 2. Prediction of pig iron silicon content in BF1 using RBFN (thick solid line), FIR (thin solid line) and ARMA (thin dashed line) modelling. Actual observations are indicated by circles.

| Type of model | RBFN | ARMA | FIR |
|---------------|------|------|-----|
| Blast furnace 1 | 0.641 | 0.815 | 0.667 |
| Blast furnace 2 | 0.262 | 0.268 | 0.509 |

Table 1. Variance of prediction errors normalised by the variance of the data for the last 1000 data points of each record (the test segment).
the development of FIR models was motivated. For both furnaces, similar explanatory inputs were obtained for the FIR models: Indices reflecting the energy reserve in the furnace, the heat of the combustion zone, the permeability and for BF1, additionally, the heat losses from the tuyeres. Quantitative results are also reported in Table 1.

In order to make a just comparison, it should be noted that the time series models (linear and nonlinear) are calculated for one-step ahead prediction (i.e., approximately 30 min in advance) and are therefore not very useful for an actual implementation. The FIR models, on the other hand, provide a forecast that is obtained about an hour in advance. Similar time variant considerations in the estimate of the parameters were employed for all linear models (ARMA and FIR) so these are, in a parametric sense, quite comparable, but the nonlinear models are time invariant and the performance on a test segment shows that the RBFN:s still perform remarkably well.

However, a simple measure such as the variance of the prediction error does not reveal all important features. Some qualitative differences regarding potential for predicting sudden changes are illustrated in Fig. 2, where the time series models often lag behind, especially at the larger variations in the latter part of the figure. Despite occasional predictions of the time series models closer to actual observations, the FIR model clearly succeeds in producing more useful predictions of important trends, one explanation naturally being the information provided by the indices used as inputs.

Interestingly, the results for the test segments are quantitatively as well as qualitatively quite different for the two furnaces. Since the linear models are time variant and use more intricate regressors than the nonlinear models, a comparison between the three approaches is not unambiguous, but it seems unnecessary to pursue nonlinear techniques for BF2. For BF1, on the other hand, the ARMA model performs poorly and the RBFN model obviously captures features of the data that the time varying linear time series model cannot.

4. Conclusions

Some methods from nonlinear time series analysis have been applied to data from two blast furnaces. Although some results are inconclusive, suitable suggestions for the delay $\tau$ and the dimension $d$ for nonlinear (empirical) modelling have been obtained. The time invariant nonlinear models developed are seen to perform well in comparison with the time varying linear ARMA and FIR models. Apparently, the linear time series models fail to capture some features of the data sets that the radial basis function networks do. A clear nonlinearity was detected in the sig-