World Academy of Science, Engineering and Technology
International Journal of Electronics and Communication Engineering
Vol:5, No:7, 2011

Model-free Prediction based on Tracking Theory and Newton Form of Polynomial

Guoyuan Qi , Yskandar Hamam, Barend Jacobus van Wyk, and Shengzhi Du

***Abstract*—**The majority of existing predictors for time series are model-dependent and therefore require some prior knowledge for the identification of complex systems, usually involving system identification, extensive training, or online adaptation in the case of time-varying systems. Additionally, since a time series is usually generated by complex processes such as the stock market or other chaotic systems, identification, modeling or the online updating of parameters can be problematic. In this paper a model-free predictor (MFP) for a time series produced by an unknown nonlinear system or process is derived using tracking theory. An identical derivation of the MFP using the property of the Newton form of the interpolating polynomial is also presented. The MFP is able to accurately predict future values of a time series, is stable, has few tuning parameters and is desirable for engineering applications due to its simplicity, fast prediction speed and extremely low computational load. The performance of the proposed MFP is demonstrated using the prediction of the Dow Jones Industrial Average stock index. .

***Keywords*—**Forecast, model-free predictor, prediction, time series

I. Introduction

TIME-SERIES prediction is becoming increasingly more important due to the needs for predicting the future behavior of control engineering [1-4], biological, physiological, economical, sociological and other systems and applications. The salient characteristic of most existing prediction approaches is model-dependence, which means first modeling the dynamic system which produces the time series and then making predictions based on the dynamic system model. Model-dependent approaches can in general be classified into two categories: 1) Model identification or parameter estimation and 2) Function approximation.

In the first category, traditional autoregressive integrated moving average (ARIMA) models provide reasonable results for linear or nearly linear dynamic systems [4], but might result in poor performance in cases where the dynamics of systems are highly nonlinear or disturbed [5]. Discrete-time adaptive predictors and controllers have also been studied for linear systems with a set of unknown parameters and systems with known nonlinearities [6]. Prediction models are constructed from observed data using various criteria, such as least squares, maximum likelihood [7], gradient methods, Gaussian filtering, Hidden Markov Models [8], Kalman filtering, Adaptive polynomial filters [9], etc. Many of these approaches are in fact related to curve-fitting schemes and make predictions based on established curve models, also referred to as local model prediction schemes.

In the second category, as viable alternatives to traditionally statistical regression models, global function approximation approaches have been presented. In this type of approach, some suitable parameterized nonlinearities, such as fuzzy systems and neural networks, are tuned adaptively to approximate unknown portions of plant dynamics during training and application. Neural networks [10], Fuzzy systems [11] and neural-fuzzy paradigms [12] have been employed in prediction with promising results, but the achievable accuracy depends on the size and the form of constructed approximations and global training from long data series.

Varadan, et al. [13] have developed a Least Squares Genetic Programming (LS-GP) approach to reconstruct nonlinear dynamic systems to model a power-pool time series. An improved prediction approach termed the Markov–Fourier Gray Model (MFGM) was proposed by Su et al. [14]. The MFGM integrates a local approach, using recent data, and a global approach, using all available information. Comparisons were made with both global and local approaches. It was demonstrated that the MFGM outperformed many existing prediction schemes on the prediction of Taiwan and Dow Jones Industrial Average stocks over two years.

The dynamic modelling of a time series is based on the assumption that the time series is a product of an underlying dynamical system. Obtaining this underlying model is in itself a very complex issue involving the simultaneous use of many different criteria. As mentioned before, these model-dependent approaches in general need very complex mathematical structures, prior knowledge, long training cycles, are computationally intensive and in the case of time-varying systems use online adaptation. Additionally, since time series, such as stocks, chaos or highly non-linear engineering systems, are possibly generated by numerous sub-systems, obtaining valid models are problematic and computationally intensive.

Therefore, a model-free approach is both innovative and necessary. Here, the word ‘free’ implies that the potentially complicated modelling process to obtain a model for the dynamic system that produced the time series is avoided. It also implies that the predictor is independent of the model of the nonlinear system (the predictor however has its own model and structure). A model-free controller for nonlinear continuous systems has been proposed by the authors [15, 16]. This controller is independent of the models of the controlled system. Stability, convergence and robustness of the closed-loop system have been investigated and the proposed

Guoyuan Qi, Yskandar Hamam, Barend Jacobus van Wyk are with F’SATI and Department of Electrical Engineering, Tshwane University of Technology, Pretoria, 0001, South Africa (email: qig@tut.ac.za).

Shengzhi Du is with Department of Electrical and Mining Engineering, University of South Africa, Florida 1710, South Africa.

model-free controller has successfully controlled various uncertain chaotic systems and high order nonlinear systems. Similar to the proposed predictor, there is no need to obtain a model of the system being controlled. Very recently, we also proposed a model-free continuous observer for a nonlinear system [17, 18] which is very effective to estimate the states of a system without knowing the model.

Motivated by the model free controller and observer, a model-free predictor (MFP) is proposed for time series produced by nonlinear systems, such as stock markets and hyperchaos. The MFP is based on ideas from tracking theory. It has two adjustable parameters related to the precision requirement of prediction. It is theoretically demonstrated that the predictor is instantly stable. When it is utilized in practical applications, the MFP has the following desirable properties: simplicity, fast prediction speed and extremely low computational load. This paper is organized as follows. The problem being solved is described in section 2. The model-free predictor (MFP) is proposed in section 3 using tracking theory. Section 4 investigates the state-steady error of the MFP. Section 5 introduces the Newton form of interpolating polynomial. An alternative derivation of the MFP is proposed by using the property of Newton form of interpolating polynomial in section 6. Section 7 analyzes the similarity between the two derivations. Two numerical applications are discussed to verify the effectiveness of the proposed MFP in section 8. Section 9 concludes the paper.

II. Problem Statement

Assume that the most recent source data in a time series whose underlying dynamics are unknown, is given by

$$\begin{aligned} Y(k) &= [y(k-n+1),\ y(k-n+2),\ \cdots,\ y(k)]^{\mathrm{T}} \\ &= [z^{-n+1},\ z^{-n+2},\ \cdots,\ 1]^{\mathrm{T}}\, y(k), \end{aligned} \tag{1}$$

where z^{-1} denotes the delay operator, $y(k)$ represents the source datum, $1 \le k \le N$, N is the length of the data, and n is the information window which can be determined using the embedding dimension [19, 20]. The unknown future sequence of data which needs to be predicted is

$$\begin{aligned} Y(k+m) &= [y(k+1),\ y(k+2),\ \cdots,\ y(k+m)]^{\mathrm{T}} \\ &= [z,\ z^2,\ \cdots,\ z^m]^{\mathrm{T}}\, y(k). \end{aligned} \tag{2}$$

Here, m is the prediction step, termed the prediction window, $m < n$, and $1 \le k \le N$. Consider a data sequence

$$\begin{aligned} \hat{Y}(k+m) &= [\hat{y}(k+1),\ \hat{y}(k+2),\ \cdots,\ \hat{y}(k+m)]^{\mathrm{T}} \\ &= [z,\ z^2,\ \cdots,\ z^m]^{\mathrm{T}}\, \hat{y}(k), \end{aligned} \tag{3}$$

where $\hat{y}(k)$ is the predictive signal of $y(k)$.

Prediction is to obtain the vector $\hat{Y}(k+m)$ as an estimate of $Y(k+m)$ based on $Y(k)$.

In the existing literature some researchers have often utilized recent data to reconstruct a dynamic model of the form

$$y(k) = \varphi(c,\ y(k-1),\ y(k-2),\ \cdots,\ y(k-n)) + e(k), \tag{4}$$

where, φ is some nonlinear function, c is a parameter vector of the function, and $e(k)$ is the noise or the disturbance. Complete reconstruction of the dynamic model involves the estimation of φ, c, and n. Various techniques and complex mathematical structures have been proposed to solve the inverse problem of constructing a smooth map

$$\varphi : R^n \to R^1 \tag{5}$$

in terms of the time-series values $y(k)$. Generally, one needs to firstly find an approximate solution or a best fitting model using finite time-series data, and then use the obtained model to predict future values of the time series.

III. Model-Free Predictor

A new scheme for prediction termed model-free prediction (MFP) is proposed, which does not require the derivation of a model to fit the source data. By incorporating a zero tracking method into the prediction, we directly establish a dynamic relationship between the prediction and the given information. The predictor is independent of the model which produces the given data.

The MFP dynamic model between $\hat{y}(k+m)$ and the most recent data $Y(k)$ is usually expressed as

$$\hat{y}(k+m) = f(Y(k)), \tag{6}$$

where f should be such that $\hat{y}(k+m)$ accurately predicts $y(k+m)$ based on $Y(k)$.

Observation 1: If a prediction $\hat{y}(k+m)$ is obtained from the MFP, prediction data $\hat{y}(k+m-1), \cdots, \hat{y}(k+2),\ \hat{y}(k+1)$ are a delayed sequence of the prediction $\hat{y}(k+m)$, and were available before $\hat{y}(k+m)$.

Observation 2: $\hat{y}(k+m)$ predicting $y(k+m)$ is not a common tracking because it is implemented based on the *m*-step backward sequence $Y(k)$ but not present sequence $Y(k+m)$.

The schematic diagram of the proposed MFP.

From eq. (6), we get

$$\hat{y}(k+m) = f(Y(k)). \tag{7}$$

Function $f(\cdot)$ can be either nonlinear or linear. For the sake of simplicity in establishing and analyzing $f(\cdot)$, we choose it as a linear function having the form

$$\begin{aligned} f(Y(k)) &= f_0 y(k) + f_1 y(k-1) + \cdots + f_{n-1} y(k-n+1) \\ &= f(z) y(k). \end{aligned} \tag{8}$$

where

$$f(z) = f_0 + f_1 z^{-1} + \cdots + f_{n-1} z^{-(n-1)}. \tag{9}$$

Then we have

$$\hat{y}(k+m) = f(z) y(k). \tag{10}$$

The prediction error between $y(k+m)$ and $\hat{y}(k+m)$ is described as

$$e(k+m) = y(k+m) - \hat{y}(k+m) = \left[z^m - f(z)\right] y(k) \tag{11}$$

Taking the *z*-transform on both sides, we obtain

$$z^m E(z) = z^m \left(Y(z) - \hat{Y}(z)\right) = \left[z^m - f(z)\right] Y(z) \tag{12}$$

resulting in the error transfer function,

$$\frac{E(z)}{Y(z)} = z^{-m}\left[z^m - f(z)\right] = \frac{z^{n+m-1} - z^{n-1}f(z)}{z^{n+m-1}}, \tag{13}$$

also known as a linear Prediction Error Filter (PEF), where $z^{n-1}f(z)$ is function in *z*.

Note that this PEF transfer function has the following properties:

1) There are $n+m-1$ multiple poles of system at the origin, which means the predictive system designed is stable and output $\hat{y}(k)$ instantly predict the given input $y(k)$.

2) It is monic polynomial.

3) The coefficients of $z^{m+n-2}, \cdots, z^n$ are zeros.

In order to make the steady-state error of the MFP be zero, we design the PEF transfer function as

$$\frac{E(z)}{Y(z)} = \frac{(z-1)^n g(z)}{z^{n+m-1}}. \tag{14}$$

For (14) to satisfy the PEF properties above, $g(z)$ is a monic polynomial with degree $m-1$ in *z*:

$$g(z) = z^{m-1} + a_1 z^{m-2} + \cdots + a_{m-1}. \tag{15}$$

Comparing eq. (13) and (14), we have

$$z^{n+m-1} - z^{n-1}f(z) = (z-1)^n g(z) \tag{16}$$

$f(z)$ is obtained as

$$\begin{aligned} f(z) &= z^{-(n-1)}\left(z^{n+m-1} - (z-1)^n \left(\sum_{i=0}^{m-1} a_i z^{m-1-i}\right)\right) \\ &= z^m - \left(\sum_{i=0}^{n}(-1)^i C_n^i z^{1-i}\right)\left(\sum_{k=0}^{m-1} a_i z^{m-1-k}\right) \\ &= \sum_{i=1}^{n+m-1} F_i z^{m-i} \end{aligned} \tag{17}$$

where $F_i = \sum_{k=0}^{i} a_k (-1)^{i-k-1} C_n^{i-k}$, $i = 1, \cdots, n+m-1,$ with $\begin{cases} C_n^{i-k} = 0, \text{ when } i-k>n, \\ a_k = 0, \text{ when } k \ge m. \end{cases}$

From property 3), we have m following simultaneous equations

$$F_i = \sum_{k=0}^{i} a_k (-1)^{i-k-1} C_n^{i-k} = 0, \;\; i = 1, \cdots, m-1. \tag{18}$$

From (18), a_i is recursively derived as follows

$$a_k = \sum_{j=0}^{k-1} a_j (-1)^{k-j-1} C_n^{k-j}, \;\; k = 1, \cdots, m-1. \tag{19}$$

From (17) and (18), we have

$$\begin{aligned} f(z) &= \sum_{i=m}^{n+m-1} F_i z^{m-i} \\ &= \sum_{i=0}^{n-1} F_{i+m} z^{-i} \\ &= \sum_{i=0}^{n-1}\sum_{k=0}^{i+m} a_k (-1)^{i+m-k-1} C_n^{i+m-k} z^{-i}, \end{aligned} \tag{20}$$

with $C_n^{i+m-k} = 0$ when $i+m-k>n,$ and $a_k = 0$ when $k \ge m$. Using (10) and (20), the MFP is given by

$$\hat{y}(k+m) = f(z)y(k), \tag{21}$$

where $f(z) = \sum_{i=0}^{n-1}\sum_{k=0}^{i+m} a_k (-1)^{i+m-k-1} C_n^{i+m-k} z^{-i}$, with $C_n^{i+m-k} = 0$ when $i+m-k>n,$ and $a_k = 0$ when $k \ge m$.

The MFP only has two parameters termed the predictive step m and the information window n. The latter can be easily determined by the FNN technique [17] based on the analysis of the given data. For convenience of application of (21), the following two special cases are given.

Case 1: One-step prediction, i.e. $m=1$. From (21), we have

$$\hat{y}(k+1) = \left(\sum_{i=1}^{n}(-1)^{i-1} C_n^i z^{-i+1}\right) y(k). \tag{22}$$

Case 2: Two-step prediction, i.e. $m=2$. From (21), we have

$$y(k+2) = \left(\sum_{i=2}^{n}(-1)^{i-1} C_n^i z^{-i+2} + C_n^1 \sum_{i=1}^{n}(-1)^{i-1} C_n^i z^{-i+1}\right) y(k) \tag{23}$$

IV. MFP Steady State Error

As is well known, both the type of PEF transfer function and the property of the input signal determine the accuracy of the steady-state error. Therefore, a general conclusion about steady-state errors is difficult to obtain due to the dependence upon the input signal. For example, the *z*-transforms of input signals $1(kT)$, kT, $(kT)^2$, $\sin(\omega kT)$, etc. are $z/(z-1)$, $Tz/(z-1)^2$ and $T^2 z(z^2+1)/(z-1)^3$, $z\sin(\omega T)/(z^2 - 2z\cos(\omega T)+1)$, etc. respectively. In practice, input signals are usually a combination of such typical signals. Despite these difficulties, we can still maintain predictive precision by adjusting the predictive step and information window based on a given input. Consider $y(t)$ to be the original continuous function of the sampled data $y(k)$, expanded in a Taylor series

$$y(t) = y(0) + \dot{y}(0)t + \frac{1}{2!}\ddot{y}(0)t^2 + \cdots + \frac{1}{n_1!}y^{(n_1)}(0)t^{n_1} + R_{n_1}(t) \tag{24}$$

Assumption 1: $R_{n_1}(t)$ in eq. (24) can be ignored.

In fact, *Assumption 1* is reasonable in engineering processes. For signals, such as sine and cosine functions, the magnitudes of the higher derivatives approach zero.

The *z*-transform of the term t^{n_1} of which corresponding discrete form is $(kT)^{n_1}$ has the form

$$Z\left[y(kT)\right] = \lim_{a\to 0}(-1)^{n_1}\frac{d^{n_1}}{da^{n_1}}\left[\frac{z}{z-e^{-aT}}\right] = \frac{h(Tz)}{(z-1)^{n_1+1}}, \tag{25}$$

where $h(\cdot)$ is a polynomial function of Tz, T is the sample time. From (24) and (25), the *z*-transform of $y(k)$ under *Assumption 1* can be written as

$$Y(z) = \frac{\phi(Tz)}{(z-1)^{n_1+1}}, \tag{26}$$

where $\phi(Tz)$ is a polynomial of Tz, and the highest order of denominator of $Y(z)$ is n_1+1.

From (14), we have the steady-state error transform of the MFP as

$$E(z) = \hat{Y}(z) - Y(z) = \frac{(z-1)^n g(z)}{z^{n+m-1}} Y(z). \tag{27}$$

Setting $n > n_1$, under *Assumption 1*, from (26) and (27), and using the final value theorem, we obtain

$$e(\infty) = \lim_{z\to 1}(1-z^{-1})E(z) = \lim_{z\to 1}\frac{(z-1)^{n-n_1}}{z^{n+m-1}}g(z)\phi(Tz) = 0 \,.(28)$$

It implies that the predictive sequence $\hat{Y}(k+m)$ based on $Y(k)$ is able to approximately predict the real signal sequence $Y(k)$ with zero error. Since $E(z)$ has multiple poles at the origin, the approximating process becomes instant process, i.e.

$$e(k) = y(k) - \hat{y}(k) = 0 \text{ or } E(z) = 0 \qquad (29)$$

or

$$e(k+m) = y(k+m) - \hat{y}(k+m) = 0 \text{ or } z^m E(z) = 0 \qquad (30)$$

From (16) and (27), we have

$$z^{n+m-1} - z^{n-1} f(z) = (z-1)^n g(z) = 0 \qquad (31)$$

Lemma 1 (*Weierstrass Approximation Theorem*): If $y(t)$ is continuous on a finite interval $[a, b]$, then given $\varepsilon > 0$, there exists n_1 depending on ε and a polynomial $P_{n_1}(t)$ such that $\left|y(t) - P_{n_1}(t)\right| \le \varepsilon$ for all $t \in [a, b]$.

The *z*-transform of sampling polynomial $P_{n_1}(kT)$ of degree n_1 corresponding $P_{n_1}(t)$ has the same form of eq. (25). Therefore, we have that 1) if $y(k)$ is produced by a polynomial model $P_{n_1}(kT)$ of order $n_1 < n$, then the proposed MFP instantly results in zero prediction error; 2) if $y(k)$ is a non-polynomial nonlinear continuous signal, the MPF has prediction error $e(k) \le \varepsilon$ for any given $\varepsilon > 0$ by adjusting n.

Theorem 1: The proposed model-free linear predictor with a dynamic model as in (21) is independent of the model of the original nonlinear system producing the given sequence $Y(k)$. The MFP is stable. The MFP under *Assumption 1* makes the predictive sequence $\hat{Y}(k+m)$ instantly predict the real signal sequence $Y(k+m)$ with prediction error $e(k+m) \le \varepsilon$ for given $\varepsilon > 0$ by adjusting n. In the sequel we will show that one can also derive *Theorem 1* by utilizing the Newton form of interpolating polynomial property.

V. Introduction of Newton form of interpolating polynomial [23]

The Newton polynomial is sometimes called Newton's divided differences interpolation polynomial because the coefficients of the polynomial are calculated using divided differences.

Given a set of $n+1$ data points $\left(x(0), y(0)\right), \cdots, \left(x(n), y(n)\right)$, where no two $x(j)$ are the same, the Newton form of the first order interpolating polynomial $P_1(x)$ is depicted as

$$P_1(x) = P_0(x) + \Delta[y(0), y(1)]\left(x - x(0)\right), \qquad (32)$$

where initial $P_0(x) = y(0) = \Delta^0[y(0)]$, and $\Delta[y(0), y(1)] = \dfrac{y(1)-y(0)}{x(1)-x(0)}$ called the first order divided difference. The second order $P_2(x)$ is

$$\begin{aligned} P_2(x) &= P_1(x) + \Delta^2[y(0), y(1), y(2)]\left(x - x(0)\right)\left(x - x(1)\right) \\ &= \Delta^0[y(0)] + \Delta[y(0), y(1)]\left(y - y(1)\right) \qquad (33) \\ &\quad + \Delta^2[y(0), y(1), y(2)]\left(x - x(0)\right)\left(x - x(1)\right), \end{aligned}$$

where $\Delta^2[y(0), y(1), y(2)] = \dfrac{\Delta[y(1), y(2)] - \Delta[y(0), y(1)]}{x(2) - x(0)}$ is the second order divided difference. Recursively, the n^{th} order polynomial is

$$P_n(x) = \sum_{i=0}^{n} \Delta^i[y(0), \cdots, y(i)] \prod_{j=0}^{i-1}\left(x - x(j)\right), \qquad (34)$$

with defining $\prod_{j=0}^{-1}\left(x - x(j)\right) = 1$ and where

$$\Delta^i[y(0), \cdots, y(i)] = \frac{\Delta^{i-1}[y(1), \cdots, y(i)] - \Delta^{i-1}[y(0), \cdots, y(i-1)]}{x(i) - x(0)} \qquad (35)$$

called the i^{th} order divided difference.

Formula (34) is known as the Newton form of the interpolating polynomial. Formula (35) allows us to generate all the divided difference needed for the Newton formula in a simple manner by using a divided difference table, rather than using formula. A table for the case $n = 3$ is illustrated below.

TABLE 1 DIVIDED DIFFERENCE TABLE

$x(k)$	$y(k)$	$\Delta[\,]$	$\Delta^2[\,,\,]$	$\Delta^3[\,,\,,\,]$
$x(0)$	$y(0)$			
		$\Delta[y(0), y(1)]$		
$x(1)$	$y(1)$		$\Delta^2[y(0), y(1), y(2)]$	
		$\Delta[y(1), y(2)]$		$\Delta^3[y(0), y(1), y(2), y(3)]$
$x(2)$	$y(2)$		$\Delta^2[y(1), y(2), y(3)]$	
		$\Delta[y(2), y(3)]$		
$x(3)$	$y(3)$			

Some important properties of divided differences are [23]:

1) Divided differences are invariant under permutations of $x_0, x_1, x_2, \cdots$.
2) If $y(k)$ is from a polynomial of order n_1 and $n > n_1$, then the n-th divided difference of $y(k)$ vanish identically.
3) If $y(k)$ is from a polynomial of order $n_1 + n$, then the n-th divided difference is a polynomial in x of order n_1.

VI. Model-free predictor derived from Newton form of interpolating polynomial property

Consider the special case where $x(k) = kT$ and T is the sample period, and $y(k-n+1), \cdots, y(k), y(k+1)$ a given sequence. The divided difference table of the Newton form of the polynomial is depicted as in Table 2.

Table II Divided difference table for number sequence

k	$y(k)$	$\Delta[\ ]$	$\ldots$	$\Delta^n[y(k-1+n),\cdots,y(k+1)]$
$(k-n+1)T$	$z^{-(n-1)}y(k)$			
		$z^{-(n-1)}(z-1)y(k)$		
$(k-n+2)T$	$z^{-(n-2)}y(k)$			
		$z^{-(n-2)}(z-1)y(k)$		
$(k-n+3)T$	$z^{-(n-3)}y(k)$			
$\vdots$			$\cdots$	$\dfrac{z^{-(n-1)}}{n!}(z-1)^n y(k)$
$(k-1)T$	$z^{-1}y(k)$			
		$z^{-1}(z-1)y(k)$		
kT	$y(k)$			
		$(z-1)y(k)$		
$(k+1)T$	$zy(k)$			

Assumption 2: $y(k)$ is generated from n_1-th order Newton form of interpolating polynomial.

From property 2) of Newton form of interpolating polynomial, when $n > n_1$, we have

$$\Delta^n[y(k-1+n), \cdots, y(k+1)] = \frac{1}{n!} z^{-(n-1)}(z-1)^n = 0\,, \qquad (36)$$

i.e.

$$z - \sum_{i=1}^{n} (-1)^{i-1} C_n^i z^{-(i-1)} = 0\,. \qquad (37)$$

Therefore,

$$y(k+1) = \left(\sum_{i=j}^{n} (-1)^{i-1} C_n^i z^{-(i-1)} \right) y(k)\,, \qquad (38)$$

is a one-step predictor based on the known information $y(k-n+1), \cdots, y(k)$ which is exactly the same as the (22).

For a sequence $y(k-n+1), \cdots, y(k), y(k+1), y(k+2)$, we have

$$\Delta^{n+1}[y(k-1+n), \cdots, y(k+2)] = z^{-(n-1)}(z-1)^{n+1} = 0 \qquad (39)$$

and

$$\begin{aligned} & z^{-(n-1)}(z-1)^{n+1} \\ & = z^2 - \sum_{i=2}^{n} (-1)^{i-1} C_n^i z^{-i+2} - C_n^1 \sum_{i=1}^{n} (-1)^{i-1} C_n^i z^{-i+1} \end{aligned} \qquad (40)$$

Therefore,

$$y(k+2) = \left(C_n^1 \sum_{i=1}^{n} (-1)^{i-1} C_n^i z^{-i+1} + \sum_{i=2}^{n} (-1)^{i-1} C_n^i z^{-i+2} \right) y(k) \qquad (41)$$

is a two-step predictor based on information $y(k-n+1), \cdots, y(k)$. Similarly, for a sequence $y(k-n+1), \cdots, y(k), \cdots, y(k+m)$, the final form *m*-step predictor is expected to be

$$z^m - f(z^{-1}) = 0\,, \qquad (42)$$

where $f(z^{-1}) = f_0 + f_1 z^{-1} + \cdots + f_{n-1} z^{-(n-1)}$, and $f_i, i = 0, \cdots, n-1$ are underdetermined coefficients.

Set $g(z)$ a monic polynomial of z, with order $m-1$. By multiplying $\Delta^n[y(k-1+n), \cdots, y(k+1)]$ by $g(z)$, from (36), we have

$$z^{-(n-1)}(z-1)^n g(z) = 0\,. \qquad (43)$$

From (42) and (43), we have

$$z^{-(n-1)}(z-1)^n g(z) = z^m - f(z^{-1}) = 0\,. \qquad (44)$$

That is

$$(z-1)^n g(z) = z^{n+m-1} - z^{n-1} f(z^{-1}) = 0\,, \qquad (45)$$

which is exactly same as (16). So from property 2) of Newton form of interpolating polynomial, we arrived at the same model-free predictor (21) which satisfy *theorem 1*.

VII. Simulation

Example: For our simulation, the daily closing values of the Dow 10 from 2005-6-2 to 2007-5-31, were used. The Dow 10 is a market index constructed as a subset of the Dow Jones Industrial Average. The stock market series consists of 500 daily closing prices obtained from the Dow Jones web site, and is shown in Fig. 1. Note that there are no trading on Saturday and Sunday.

Su, *et al.* [14] proposed a prediction approach termed the Markov–Fourier Gray Model (MFGM). In this paper we compare the MFGM approach with the MFP approach. The MFGM can provide improved performance when compared to the existing prediction schemes, such as neural networks, neural fuzzy networks, and first order polynomial fitting. The MFGM has been demonstrated by predicting the Dow Jones Industrial Average stocks during the period 1999 to 2001. 1000 data points were used, in which, the first 700 were used as the training data set and the last 300 are the test data set. The influence window size was 5. Since the stock index is very noisy, pre-processing using a 5-day average (a kind of LP filtering) of the series was conducted. Both the mean square error (MSE) and the absolute mean error (MAE) were computed and recorded as

$$\text{MSE} = \frac{1}{N} \sum_{k=1}^{N} \left(y(k+1) - \hat{y}(k+1) \right)^2 = 3.18 \times 10^3\,, \qquad (46)$$

$$\text{ASE} = \frac{1}{N} \sum_{k=1}^{N} \left| \hat{y}(k+1) - y(k+1) \right| = 42.63\,, \qquad (47)$$

where, N is the number of test data used for prediction.

For the proposed MFP approach, we pre-processed the data by filtering the data using a combination of lower pass Butterworth IIR digital forward-backward filtering which is implemented by first filtering in the forward direction, then reversing the filtered sequence and running it back through the filter [21, 22]. The resulting sequence has zero-phase distortion and double the filter order. The order of the filter was chosen as 2 and the normalized frequency as w_n=0.35.

To determine the information window, we used the False Nearest Neighbor (FNN) method [20], one of the most popular techniques to find an optimal value for *n*. The optimal value for *n* is where the number of these FNNs reduces to 0. To ensure that the FNN converge, we used 2252 daily original stock data samples from 1999/12/31 to

2007/12/10. The embedding-dimension convergence diagram is shown in Fig. 2 for the stock time series. As can be seen from the figure, the FNN reduced to 0 beyond information window of 4, and hence we chose n=4.

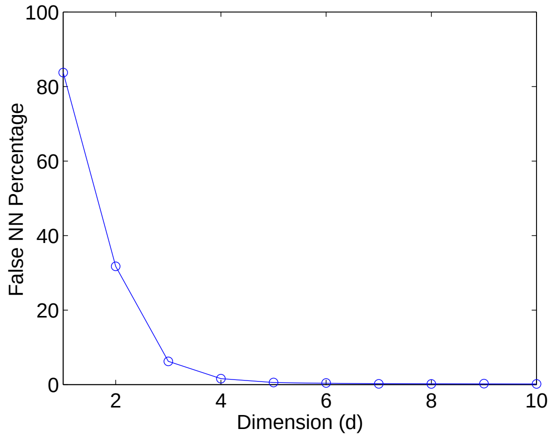


Fig. 2. The FNN plot of Dow 10 indexes for 2252 daily stock data from 1999/12/31 to 2007/12/10.

After pre-processing the stock data and the determining the information window, we used the proposed MFP (22) with $n = 4$ to produce a one-step prediction of the stock. The MFP performed excellently as shown in Fig. 3. The costs, MSE=5.6263 and MAE=1.8444, are small compared with eqs. (46) and (47) using the MFGM, indicating superior prediction precision.

To be shown more clearly, the zoom between day 100 to day 200 is shown in Fig. 4. As can be seen from Figs. 3 and 4, the time series of the index looks quite irregular, with the indexes exhibiting sudden peaks. The proposed MFP does not only achieve good prediction of the stock time series but also accurately predicts the peaks in the stocks. The efficient prediction of such peaks is important for investors. The lack of adequate tools to predict stock indices, especially peaks, might result in economic loss.

In this method, 1) the MFP is independent of the model of the stock time series; 2) the dynamic model of the MFP is constant once the predictive step m and the embedding dimension n are fixed, and the prediction changes as newly available data change; 3) the model training process is avoided. Without model identification or fitting, the whole 500 daily stock values were used as the test set. 4) it does not need to online adapt to changing parameters or nonlinearities and 5) it has two parameters and simple calculation using eq. (21) therefore having a low computational cost. It is very convenient to apply to practical applications since only the prediction step m and information window n need to be determined.

We also performed two-day prediction with $m = 2$ and $n = 4$ by applying (23). The performance achieved is shown in Fig. 5 with the costs

$$\text{MSE} = \frac{1}{N}\sum_{k=1}^{N}\left(y(k+2) - \hat{y}(k+2)\right)^2 = 20.2150,$$

$$\text{ASE} = \frac{1}{N}\sum_{k=1}^{N}\left|\hat{y}(k+2) - y(k+2)\right| = 3.5193,$$

indicating good performance.

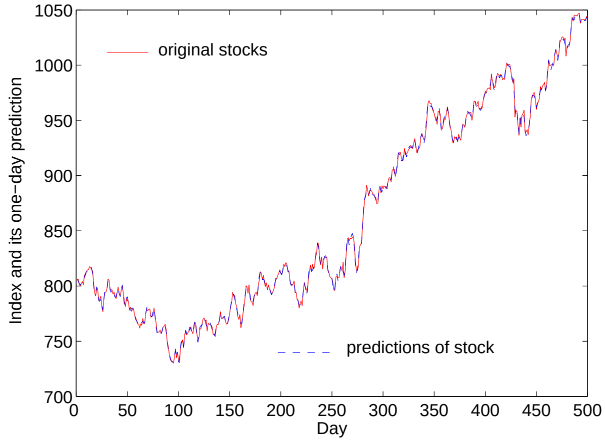


Fig. 3. The comparison of Dow 10 indexes and their one-day predictions through the MFP for 500 days from 2005-6-2 to 2007-5-31.

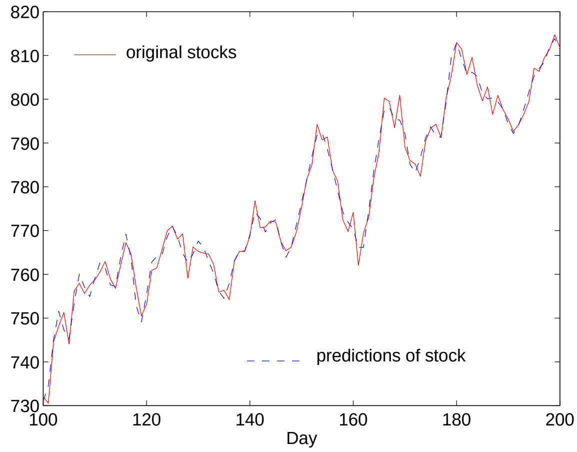


Fig. 4. The comparison of Dow 10 indexes and their one-day predictions through the MFP for 100 days.

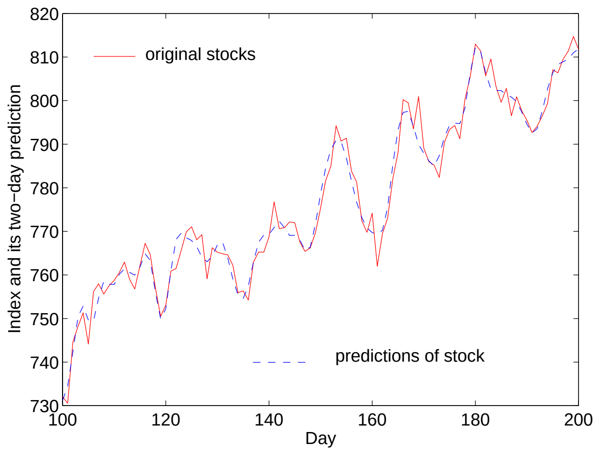


Fig. 5. The comparisons of Dow 10 indexes and their two-day predictions through the MFP for 100 days.

Example 2: The Mackey–Glass [24] differential delay equation

$$\frac{dx(t)}{dt} = -0.1x(t) + \frac{0.2x(t-\Delta)}{1 + x(t-\Delta)^{10}} \tag{48}$$

is a classic model to test model reconstruction and prediction [14, 25]. Here, we used $\Delta = 17$, initial conditions $x(0) = 0.9$, and sample time $T = 1$, Lyapunov dimension $D = 2.1$ and Lyapunov exponents $l_1 = 0.0086, l_2 = 0.001, l_3 = -0.0395, \quad l_4 = -0.0504$ [22].

During the last two decades, many forecasting algorithms have been developed to predict chaotic time series based on the theory of dynamical reconstruction from scalar time series [25]. These algorithms for predicating chaotic time series are almost all based on reconstructed attractors [26]. Recently, adaptive second-order Volterra filters have successfully been used to predict some low dimensional chaotic time series [27]. More recently, an adaptive higher-order nonlinear finite impulse response (HONFIR) based on the high-order Volterra expansions was proposed to predict high dimensional hyper-chaotic time series [25]. As is well-known, hyper-chaos is a chaotic behaviour in which more than one positive Lyapunov exponents exist in the high dimensional system; the systematic dynamics is more complex and more sensitively dependent on initial conditions than chaotic ones. However, we must point out that although the hyperchaos or the chaos is very disordered, it is pseudo-noise, since it is derived from a determinative dynamic function as opposed to a stochastic one like noise or stock, etc. The values of the positive Lyapunov exponents determine the degree of disorder of chaos or hyperchaos.

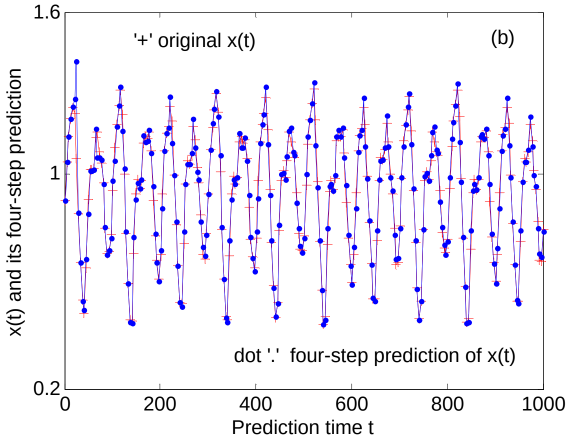
We predicted the Mackey–Glass equation using the proposed MFP. Since the two positive Lyapunov exponents of the hyperchaotic signal are very small, the hyperchaos is not extremely noise like, as shown in Fig. 6. Therefore, it is not necessary to filter the sequence as we've done for the stocks, i.e. pre-processing the sequence was ignored. The proposed MFP (21) was used to obtain one-step and four-step predictions, respectively. The MFP performed excellently as shown in Fig. 6 where MSE=2.1434 × 10^{-6} (one-step) and MSE=1.1 × 10^{-3} (four-step), respectively. Moreover, since noise-free data for the Mackey-Glass equation is unrealistic, we have also added noise to the data set as was done in [14]. A normal distribution with zero mean and 1% standard deviation was used as the noise distribution function. With additional noise added, we pre-processed the data using the 'Butterworth' filter with order n=2, normalized frequency w_n=0.35, and zero-phase filter. The MSE=5.1844×10^{-4} obtained is almost the same as the MSE=6.724×10^{-4} in [14], but the proposed MFP is much easier and simpler to implement than the MFGM [14].

Zhang and Xiao [25] also conducted the example without adding any noise using the adaptive HONFIR and compared with other methods such as the fuzzy neural networks, the conventional neural networks using the root-mean-square-error

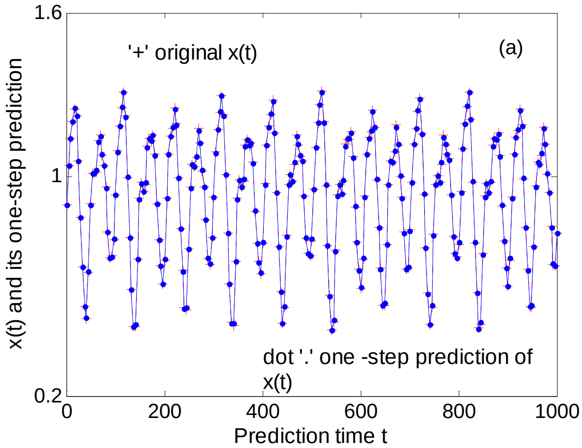
$$E_{rms} = \frac{1}{N}\sum_{k=1}^{N} \frac{\left\langle \left(\hat{y}(k+m) - y(k+m)\right)^2 \right\rangle^{1/2}}{\left(y(k+m)\right)^{1/2}}. \quad (49)$$

Table 3 lists the simulation results for one-step, two-step and four-step predictions [25]. It is obvious that the adaptive HONFIR filter provides better prediction performance than the rest, excluding the MFP.

The proposed MFP results are listed in the last row in table 3. Clearly, the proposed MFP demonstrates a superior prediction performance to the methods in the table.



(a) one-step prediction



(b) four-step prediction

Fig. 6. The comparison between Mackey-Glass data and their prediction. To show clearly the prediction for making future predictions of chaotic series, only one is plotted for each four points compared with the actual values.

TABLE III PERFORMANCES OF VARIOUS PREDICTION SCHEMES FOR MACKEY-GLASS SERIES

Prediction scheme	E_{rms} (1-tep)	E_{rms}(2-step)	E_{rms}(4-step)
Conventional Neural network	1.0×10^{-2}	1.9×10^{-2}	4.9×10^{-2}
Fuzzy neural network	9×10^{-3}	1.8×10^{-2}	4.7×10^{-2}
HONFIR	8.5285×10^{-5}	2.9410×10^{-4}	1.4×10^{-3}
MFP	1.12×10^{-6}	5.4624×10^{-6}	1.9933×10^{-5}

VIII. CONCLUSIONS

In this paper, a model-free predictor for nonlinear time series was introduced based on a zero tracking error methodology. An identical derivation of the MFP using the Newton form of the interpolating polynomial was also proposed to verify the reasoning and the reliability of the MFP. The underlying model-free prediction approach is very different from existing methodologies. The MFP performed well as was demonstrated using an example. It is desirable in engineering applications due to its simplicity, linearity, few adjustable parameters and low computational load. More theoretical analysis and more application verifications such as control, image motion prediction, oil price and weather forecasting, etc. will be conducted.

ACKNOWLEDGEMENT

This work was supported by the grants: National Research Foundation of South Africa (No. IFR2009090800049), Eskom - Trusts /Tertiary Education Support Programme of South Africa.

REFERENCES

[1] D. J. Trudnowski, W.L. McReynolds, J.M.Johnson, "Real-time very short-term load prediction for power-system automatic generation control," *IEEE Trans. Control Systems Technology* vol 9 pp: 254 – 260, 2001.

[2] M. H. Ali, T. Murata, J. Tamura, "Influence of Communication Delay on the Performance of Fuzzy Logic-Controlled Braking Resistor Against Transient Stability," *IEEE Trans Control System and Technology*, vol 16, pp: 1232 – 1241, 2008.

[3] A.S. Sharma, D.J.N. Limebeer, I.M. Jaimoukha, J.B. Lister, "Modeling and control of TCV", *IEEE Trans. Control Systems Technology*, vol 13, pp: 356 -369, 2005

[4] B. L. Bowerman and R. T. O'Connell, *Time Series Forecasting*. New York: PWS, 1987.

[5] J. T. Spooner, M. Maggiore, R. Ordóñ, and K. M. *Passino, Adaptive Control and Estimation for Nonlinear Systems: Neural and Fuzzy Approximator Techniques*. New York: Wiley, 2002.

[6] G. C. Goodwin and K. S. Sin, *Adaptive Filtering, Prediction, and Control*. Englewood Cliffs, NJ: Prentice-Hall, 1984.

[7] T. Matsumoto, Y. Nakajima, M. Saito, J. Sugi and H. Hamagishi, "Reconstructions and predictions of nonlinear dynamical systems: a hierarchical bayesian Aapproach," *IEEE Trans. Signal Proc*., vol. 49, pp. 2138 – 2155, 2001.

[8] A. M. González, A. M. S. Roque and J. García-González. "Modeling and forecasting electricity prices with input/output hidden markov models," *IEEE Trans. Power Syst*., vol. 20, pp. 13- 24, 2005.

[9] V. J. Mathews, "Adaptive polynomial filters", *IEEE Signal Processing*, vol. 8, pp. 10-26, 1991.

[10] J. T. Conner, R. D. Martin, and L. Atlas, "Recurrent neural networks and robust time series prediction," *IEEE Trans. Neural Networks*., vol. 5, pp. 240–254, 1994.

[11] O. Raúl, T. J. Spooner and M. Kevin, "Passino. Experimental studies in nonlinear discrete-time Adaptive prediction and control," *IEEE Trans. on Fuzzy Syst*., vol. 14, pp. 275-286, 2006.

[12] C. V. Altrock, Fuzzy Logic & NeuroFuzzy *Applications Explained*, Englewood Cliffs, NJ: Prentice Hall, 1995.

[13] V. Varadan, H. Leung, and É. Bossé, "Dynamical model reconstruction and accurate prediction of power-pool time series." *IEEE Trans. Instru. Measurement*, vol. 55, pp. 327-336, 2006.

[14] S. F. Su, C. B. Lin and Y. T. Hsu, "A high precision global prediction approach based on local prediction approaches." *IEEE Trans. System Man and Cybernetic —Part C: Aplications and Reviews*, vol. 32, pp. 416-525, 2002.

[15] G. Qi, Z. Chen and Z. Yuan, "Model free control of affine chaotic system," *Phys. Lett. A*., vol. 344, pp. 189-202, 2005.

[16] [16] G. Qi, Z. Chen and Z. Yuan. "Adaptive high order differential feedback control for affine nonlinear system," *Chaos, Solitons & Fractals*, vol. 37, pp. 308-315, 2008.

[17] G. Qi, M.A. van Wyk and B. J. van Wyk, "Model-free differential states observer for nonlinear affine system," *The 7th International Federation of Automation Control (IFAC) Symposium on Nonlinear Control systems*, pp. 984-989, 2007.

[18] G. Qi, Z. Chen, Z. Yuan, Stable high orders differentiator design and its application in observer and controller. *The Chinese Journal of Electronics*,vol. 14, pp. 644-648, 2005.

[19] F. Takens, Detecting strange attractors in turbulence, in: D. A. Rand, L. S. Young (Eds.), *Dynamical Systems and Turbulence*, vol. 898, Springer, Berlin, 365–381, 1981.

[20] M. B. Kennel, R. Brown, and H. D. I. "Abarbanel. Determining embedding dimension for phase-space reconstruction using a geometrical construction," *Phys. Rev. A, Gen. Phys*., vol. 45, pp. 3403–3411, 1992.

[21] F. Gustafsson, "Determining the initial states in forward-backward filtering," *IEEE Trans. Signal Processing*, vol. 44, pp. 988—992, 1996.

[22] S. K. Mitra, *Digital Signal Processing*, 2nd ed., McGraw-Hill, Sections 4.4.2 and 8.2.5. 2001.

[23] http://planetmath.org/encyclopedia/DividedDifference.html

[24] M. C. Mackey and L. Glass, "Oscillation and chaos in physiological control systems," *Science*, vol. 197, pp. 287–289, 1977.

[25] J. Zhang, X. Xiao, Predicting hyper-chaotic time series using adaptive higher-order nonlinear filter. *Chin. Phys. Lett*. vol. 18, pp. 337-340, 2001.

[26] V. J. Mathews, "Adaptive polynomial filters", *IEEE Signal Processing*, vol. 8, pp. 10-26, 1991.

[27] F. Rauf and H. M. Ahmed, New Nonlinear Adaptive Filters with Applications to Chaos, *Int. J. Bifurcation Chaos*, vol. 7, pp. 1791-1809, 1997