Nonlinear Time Series Modeling Using Spline-based Nonparametric Models

Jun M. Liu
Georgia Southern University
Department of Finance & Quantitative Analysis
Statesboro, GA 30460
USA
jliu@georgiasouthern.edu

Abstract: In this paper we use the polynomial splines-based nonparametric transfer function method to study how river ¤ow is affected by multiple factors. The highly nonlinear relationship between river ¤ow and the independent variables (the transfer function) is modeled using polynomial spline, and the noise term assumed to follow a parametric Autoregressive (AR) model. The transfer function is modeled jointly with the AR parameters. Because of its ¤exibility, spline functions are ideal for modeling highly nonlinear relationships with unknown functional forms; by modeling the noise explicitly, the correlation in the data is removed so the transfer function can be estimated more ef£ciently. Additionally, the estimated AR parameters can be used to improve the forecasting performance. The proposed polynomial splines-based estimator is also highly computationally ef£cient. A comparison of the results show that the performance of this model is better than some widely accepted benchmark models.

Key–Words: Transfer function, splines, time series, hydrology

1 Introduction

In this paper we consider a new method to model relationships between 'output' and 'input' time series. This is a problem that plays an important role in our endeavor of forecasting and control. As a result, extensive research has been conducted in this area. For example, the linear transfer function model (Box and Jenkins, 1976) has been extensively used in practice and proven successful in many applications. However, in practice we often encounter nonlinear relationships that cannot be well approximated by linear models. Consequently, many nonlinear parametric models are introduced (Chen and Tsay, 1996; Tong, 1990; Haggan and Ozaki, 1981; Engel, 1982; Bollerslev, 1986). One problem with nonlinear parametric model is, beyond the linear domain there are in£nitely many candidate nonlinear functions, so it is usually dif£cult to justify the explicit parametric functional forms a priori. To avoid the subjectivity in selecting the parametric models, researchers adopt the principle of "letting the data speak for themselves" and use nonparametric smoothing methods to model nonlinear time series (Robinson, 1983; Auestad and Tjøstheim, 1990; Lewis and Stevens, 1991; Masry, 1996a&b; Fan and Gilbels, 1996; Smith, Wong, and Kohn, 1998). To overcome the 'curse of dimensionality', various specially structured nonparametric models have been proposed, including the functional-coef£cient autoregressive (FAR) model (Chen and Tsay, 1993a; Cai, Fan and Yao, 2000), the nonlinear additive autoregressive model (Chen and Tsay, 1993b), the adaptive functional-coef£cient model (Ichimura, 1993; Xia and Li, 1999; Fan, Yao and Cai, 2003), the single index model (e.g., Härdle, Hall, and Ichimura, 1993; Carroll, Fan, Gijbels, and Wand, 1997; Newey and Stoker, 1993; Heckman, Ichimura, Smith, and Todd, 1998; Xia, Tong, Li, and Zhu, 2002) and the partially linear models (Härdle, Liang and Gao, 2000). The literature about nonlinear and nonparametric time series analysis is extensive, reviews can be found in Tjøstheim (1994), Härdle, Lütkepohl and Chen (1997) and Fan and Yao (2003).

In this paper we consider the following relationship between two time series:

$$Y_t = f(X_t) + e_t, \qquad (1)$$

where $f(\ )$ is an unknown and smooth function, $\{X_t, e_t\}$ are jointly strictly stationary. Recently Xiao, Linton, Carroll and Mammen (2003), Su and Ullah (2006), and Liu, Chen and Yao (2007) developed methods to estimate the transfer function ef£ciently. In the studies local polynomial is used to model the transfer function $f(\ )$ and established that by modeling the serial correlation in the noise, $f(\ )$ can be estimated at the usual rate of convergence as if e_t is iid. The above methods differs mainly in the treatment of the noise e_t. Xiao, et al. (2003) assumes the noise

is a general linear process and approximates it by an AR process whose order is allowed to grow to infinity. Su and Ullah (2006) assumes the noise is a finite-order nonparametric AR process. Liu, Chen and Yao (2009) models the noise explicitly with an ARMA model. The above methods are all local polynomial-based therefore are computationally intensive. As a result they may be difficult to apply in certain practical situations, for example, it may require a very long time to generate multiple step ahead forecast by simulation. By modeling the transfer function $f(\ )$ nonparametrically, the model is flexible therefore can be used to model nonlinear relationship of unknown functional forms. By modeling $\{e_t\}$ explicitly, the autocorrelation in the data is removed so $f(\ )$ can be estimated more efficiently. Additionally, the explicit correlation structure can be used to improve the forecasting performance.

This paper is organized as follows. In section 2, the model is introduced and a short introduction of polynomial spline is included. The estimation method is introduced in section 3. The proposed procedures are applied to forecast river flow and the results are presented in section 4. Section 5 contains summary and discussion.

2 The model

In this paper we make the assumption that $\{e_t\}$ in model (1) follows a strictly stationary AR(p) process, $e_t = \sum_{i=1}^{p} \phi_i e_{t-i} + \varepsilon_t$. So model (1) can be rewritten as

$$Y_t = f(X_t) + \frac{\varepsilon_t}{1 - \sum_{i=1}^{p} \phi_i B^i} , \qquad (2)$$

where B is the back-shift operator, $B^i X_t = X_{t-i}$, $\{\varepsilon_t\}$ is a sequence of independent random variables with mean 0 and standard deviation σ. Note that the assumption of AR(p) noise is mainly for the convenience of discussion, the idea presented in this paper can be extended to more general structures of the noise, such as the ARMA(p, q) model. We also assume that $\{X_t\}$ and $\{\varepsilon_t\}$ are independent. Our interest is in estimating both $f(\ )$ and the AR parameters.

In this paper we use polynomials to model the transfer function f. Polynomial splines are piecewise polynomials defined on disjoint partitions of the support of X, with the pieces joining smoothly at a set of interior points (the *knots*). Precisely, a polynomial spline of degree $m \geq 0$ defined on an interval $\mathcal{X}$ with knot sequence $\boldsymbol{\lambda} = \{\tau_0, \tau_1, \ldots, \tau_{k+1}\}$ ($\tau_0 < \tau_1 < \cdots < \tau_{k+1}$) is a function consisting of pieces of polynomials of degree m on each of the intervals $[\tau_i, \tau_{i+1}), i = 0, \ldots, k$, and $[\tau_k, \tau_{k+1}]$, where τ_0

and τ_{k+1} are the end points of $\mathcal{X}$. Given knot sequence $\boldsymbol{\lambda}$ and degree m, the collection of spline functions form a function space spanned by basis functions. Commonly used basis functions include the well-known truncated power basis, which is the set of functions $\{1, x, \ldots, x^m, (x - \tau_1)_+^m, \ldots, (x - \tau_k)_+^m\}$, where $(x)_+^m \equiv (x_+)^m$, the dimension of the spline function space is given by $K = m + k + 1$. B-spline is often used to develop the asymptotic properties because of its nice theoretical properties (for details please see de Boor, 2001; Schumaker, 1981), but the result does not depend on the choice of the basis functions. Denote a set of basis functions as $\{B_j(\ )\}_{j=1}^{K}$, using the polynomial spline to approximate the transfer function $f(\ )$ in (2), $f(X_t) \approx \sum_{i=1}^{K} a_i B_i(X_t)$, after "pre-whitening" the noise e_t, we have the following regression model

$$Y_t \approx \sum_{i=1}^{p} \phi_i Y_{t-i} + \sum_{j=1}^{K} a_j \Big[B_j(X_t) - \sum_{i=1}^{p} \phi_i B_j(X_{t-i}) \Big] + \varepsilon_t ,$$
$$(3)$$

the estimation of the unknown parameters are estimated by solving the following optimization problem

$$\arg_{a_j, \phi} \min \sum_{t=1}^{n} \Big\{ Y_t - \sum_{i=1}^{p} \phi_i Y_{t-i} - \sum_{j=1}^{K} a_j \Big[B_j(X_t) - \sum_{i=1}^{p} \phi_i B_j(X_{t-i}) \Big] \Big\}^2 . \qquad (4)$$

3 Estimation methodology

The optimization of (4) can be carried out by standard nonlinear optimization methods. In this paper we consider the following iterative estimation algorithm, which is equivalent to the commonly used nonlinear optimization methods, such as the Gauss-Newton method. This algorithm allows us to investigate the estimators individually, which makes the discussion of the asymptotic properties more convenient.

1. Obtain preliminary estimates $\widetilde{a}_i, i = 1, \ldots, K$, which are the solutions of

$$\arg_{a_i} \min \sum_{t=1}^{n} \Big\{ Y_t - \sum_{i=1}^{K} a_i B_i(X_t) \Big\}^2 ,$$

   a preliminary estimate of f is given by $\widetilde{f}(x) = \sum_{i=1}^{K} \widetilde{a}_i B_i(x)$.

2. For given $a_j, j = 1, \ldots, K$, obtain $\widehat{\phi}_1, \ldots, \widehat{\phi}_p$ by solving

$$\arg_{\phi_i} \min \sum_{t=1}^{n} \Big\{ Y_t - \sum_{i=1}^{p} \phi_i Y_{t-i}$$

$$\sum_{j=1}^{K} a_j \Big[B_j(X_t) \quad \sum_{i=1}^{p} \phi_i B_j(X_{t-i}) \Big] \Big\}^2$$

3. For given $\phi_1, \quad , \phi_p$, obtain $\widehat{a_j}, \ j = 1, \quad , K$ by solving

$$\arg_{a_j} \min \sum_{t=1}^{n} \Big\{ Y_t \quad \sum_{i=1}^{p} \phi_i Y_{t-i}$$

$$\sum_{j=1}^{K} a_j \Big[B_j(X_t) \quad \sum_{i=1}^{p} \phi_i B_j(X_{t-i}) \Big] \Big\}^2$$

It can be easily seen that the above algorithm is guaranteed to converge. The terminating values of $\widehat{\phi}$ and $\widehat{a_i}$ provide the £nal estimates, speci£cally, $\widehat{f}(x) = \sum_{j=1}^{K} \widehat{a_j} B_j(x)$ is the £nal estimate of $f(x)$. e_t is assumed to follow a stationary AR process, so it is a mixing process, and we can expect that the preliminary estimate $\widetilde{f}$ to be consistent. The AR parameters ϕ can be estimated with the parametric rate. By modeling the serial correlation in e_t, the transfer function f can be estimated as if e_t is iid and the asymptotic results established in Huang (2003) continue to hold.

4 Modeling river ¤ow

River ¤ow forecasting is an important issue in many practical areas. For example, accurate and ef£cient forecast of river ¤ow is important to the safety of river transportation and river structures such as river dams and bridges, it also enables us to manage water resources more ef£ciently and helps protect the environment. A typical dif£culty in river ¤ow modeling is that many factors affect the river ¤ow, usually in a highly nonlinear fashion. Non-linear parametric models have been developed to address this issue (some highly-regarded models include Tong, Thanoon and Gudmundsson, 1985; Chen and Tsay, 1993b). A dif£culty in nonlinear parametric models is, with the inde£nitely many functions in the nonlinear domain, it is very dif£cult to identify the best candidate function a priori, therefore there is usually more subjectivity in the model selection. The subjectivity can be avoided by using nonparametric methods, because of the "let the data speak for themselves" property.

In this section we use the proposed nonparametric transfer function model to study the effects of temperature and precipitation on the ¤ow rate of River Jökulsá eystri in Iceland. The data consists of daily records of river ¤ow Y_t (in m^3/sec), temperature X_t (in $^\circ$C), and precipitation Z_t (in mm) from January 1, 1972 to December 31, 1974. A interesting feature

of this river is that there is a glacier in its drainage area, so we expect that the temperature effect on river ¤ow is more than melting snow. Tong et. al. (1985) used the threshold autoregressive model (TAR) to analyze this data set. Chen and Tsay (1993b) used it as an example of the nonlinear additive ARX model (NAARX). For more detailed information of the data, see Tong, et.al. (1985).

To apply the proposed approach, we £rst check the stationarity of the data. The sample ACF and PACF of Y_t show indications of non-stationarity, (Figure 1), however the result of Augmented Dicky-Fuller (ADF) test rejects the hypothesis that a unit root exists. Similar analysis shows that the exogenous variables X_t and Z_t are stationary. The details of the ADF tests are omitted to save space. To obtain
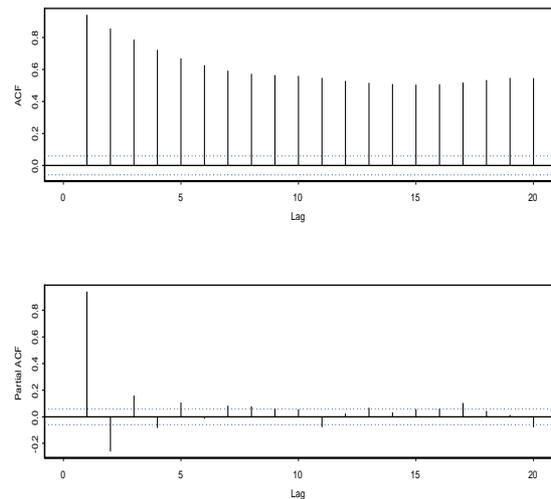

Figure 1: Upper panel: The sample ACF of Y_t. Lower panel: The sample PACF plot of Y_t.

some rough idea about the candidate variables of the model, as an initial step of model identi£cation, we estimated the linear transfer function weights using the linear transfer function method (Liu, 1982). The estimated linear transfer function weights are plotted against the lags in Figure 2, with the 95% con£dence band plotted in the dashed lines. The estimated transfer function weights suggest the following variables are good candidates of the model: $\{X_t, X_{t-1}, X_{t-2}, X_{t-3}, X_{t-4}, Z_t, Z_{t-1}, Z_{t-6}\}$. However, without any restriction, a nonparametric model containing all the above variables is dif£cult to estimate because of the *curse of dimensionality*. To overcome this problem, we consider an additive
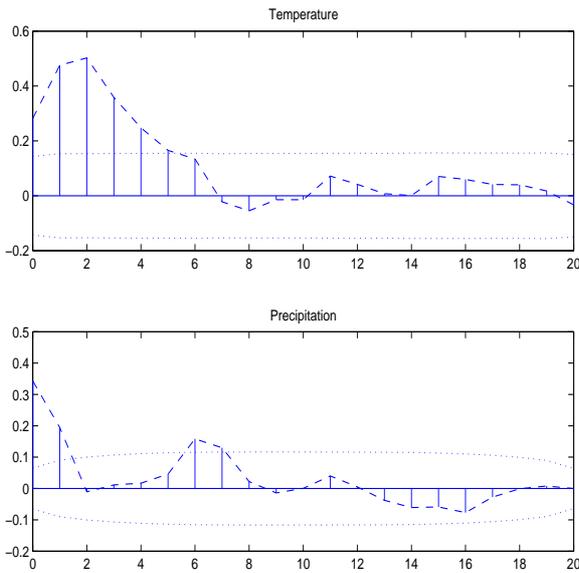
Figure 2: The estimated linear transfer function weights

model

$$Y_t = \sum_{i=0}^{4} f_i(X_{t-i}) + g_0(Z_t) + g_1(Z_{t-1})$$
$$+ g_6(Z_{t-6}) + e_t,$$

where each additive component is approximated by regression spline. The truncated power basis is used in the spline approximation. To simplify discussion, we assume that the orders and the number of knots of the spline functions are the same, thus the model can be written as

$$Y_t = a_0 + \sum_{i=0}^{4} \Big[\sum_{j=1}^{k} a_{ij} X_{t-i}^j + \sum_{r=1}^{m} a_{ir}(X_{t-i} \quad _r)_+^k \Big]$$
$$+ \sum_{i=0}^{1} \Big[\sum_{j=1}^{k} b_{ij} Z_{t-i}^j + \sum_{r=1}^{m} b_{ir}(Z_{t-i} \quad _r)_+^k \Big] + e_t.$$

We further assume that the knots are placed on the percentile points so that there are equal number of observations between any two adjacent knots. A grid search is conducted to determine the number of interior knots k and spline degree m. In the grid search k is in the interval $[1, \text{int}(5n^{1/(2m+3)})]$, where the upper limit is a multiple of the theoretical optimum number of knots (Huang, 2003). We consider linear, quadratic and cubic splines, ($m = 1, 2, 3$, respectively). The values k and m that minimizes BIC=$\log(MSE) + \log(n)[1 + (k+m)(d_x + d_z)]/n$ is the optimal number of knots, where d_x and d_z are the number of lags in X_t and Z_t, respectively. The results suggest a linear spline model with $k = 1$. Similar to the case of Y_t, the sample ACF and PACF of

the partial residual $\widetilde{e}_t$ show indications of nonstationarity, however the ADF test results again reject the existence of a unit root. As a result, an AR(4) model is selected using the BIC criterion, the sample PACF (Figure 3) also suggests such a model. Based on the
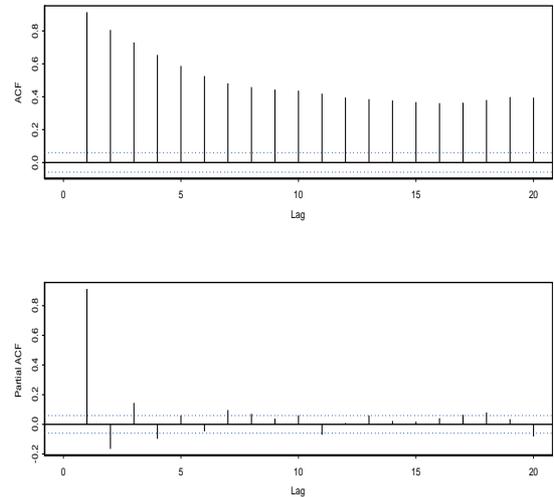

Figure 3: Sample ACF and PACF of the preliminary residual $\widetilde{e}_t$

above preliminary information about the underlying model, we re£ne the model by selecting the knots locations to minimize the residual sum of squares. The results show that Z_t and Z_{t-1} have such large knots that beyond these knots there are only a few observations, this indicates that their effects are essentially linear, also, X_{t-4} and Z_{t-6} are found to be not significant. As a result, the model simpli£es further to

$$Y_t = c + \sum_{i=0}^{3} \Big[a_{i1} X_{t-i} + a_{i2}(X_{t-i} \quad _i)_+ \Big]$$
$$+ \sum_{j=0}^{1} b_j Z_{t-j} + \frac{\varepsilon_t}{1 \quad \sum_{i=1}^{4} \phi_i B^i}. \quad (5)$$

The optimized knots are -1.3, 0.5, 0.2, and -0.2 for X_t, X_{t-1}, X_{t-2}, and X_{t-3}, respectively, and the estimated parameters and their stand deviations are given in Table 1. The sample ACF show that the residual series is roughly "white", the plots are omitted. To put the performance of the proposed NPTF model in perspective, we consider two models that were used to analyze the same data set. The £rst model is the TAR model used by Tong, et.al.(1985, page 658),

$$Y_t = c_1 + (1, \quad , 6)Y_t + (0, \quad , 5)Z_t$$
$$+ (0, \quad , 3)X_t + e_{1t}, \text{ if } X_t \leq \quad 2$$
$$= c_2 + (1, \quad , 8)Y_t + (0, \quad , 7)Z_t$$

Table 1: Parameter Estimates

Parameter	c	a_0	a_{01}	a_{11}	a_{12}
Estimate	29.67	.03	.47	-.11	2.44
StdErr.	.08	.18	.08	.17	.03
Parameter	a_{21}	a_{22}	a_{31}	a_{32}	b_0
Estimate	.05	1.42	.03	.61	.32
StdErr.	2.24	.08	.16	.08	.18
Parameter	b_1	ϕ_1	ϕ_2	ϕ_3	ϕ_4
Estimate	.17	1.15	-.40	.25	-.09
StdErr	.03	.03	.05	.05	.03

Table 2: MSE of the post-sample forecasts

Lead Time	NPTF	NAARX	TAR
1	68.38	70.06	72.84
2	164.38	169.49	183.35
3	229.10	242.17	262.32
4	291.03	303.28	336.04
5	337.66	349.77	390.82
6	366.83	382.34	421.13
7	386.43	406.94	443.02
8	391.70	418.05	457.87
9	400.98	431.81	476.08
10	408.22	439.77	490.63
11	419.27	453.05	506.71
12	440.58	477.45	536.99
In-sample	31.65	33.15	31.74

$$+(0, \quad ,7)X_t + e_{2t}, \text{ if } X_t > \quad 2, \quad (6)$$

where $(1, \quad ,6)Y_t$ means that $Y_{t-1}, \quad , Y_{t-6}$ are included in the model, other terms are similarly de£ned. The second model is the NARRX model used by Chen and Tsay (1993b, page 963),

$$
\begin{aligned}
Y_t &= c + \phi_{1,1}Y_{t-1} + \phi_{1,2}Y_{t-1}I(Y_{t-1} \quad c_1) \\
&+ \phi_{1,3}Y_{t-1}I(Y_{t-1} \quad c_2) + \phi_2 Y_{t-2} + \phi_3 Y_{t-3} \\
&+ \phi_4 Y_{t-4} + \beta_1 Z_t + \beta_2 Z_{t-1} + \omega_{1,1}X_{t-1} \\
&+ \omega_{1,2}X_{t-1}I(X_{t-1} \quad c_3) + \omega_{3,1}X_{t-3} \\
&+ \omega_{3,2}X_{t-3}I(X_{t-3} \quad c_4) + \varepsilon_t \qquad (7)
\end{aligned}
$$

The residual variances of the NPTF model, together with the residual variances of the NAARX and the TAR (Chen and Tsay 1993b) are shown in the last row of Table 2. We can see while the within-sample performance of NPTF, NAARX and TAR are similar, the NPTF model has the smallest residual variance in the three models. Although the NPTF model uses two more parameters than the NAARX model, it is still preferred by the AIC criterion. It is interesting to see that the NPTF model (5) and the NAARX model reveals similar features of the underlying process, for example, in both models piecewise linear functions are found to well describe the relationship between temperature and river ¤ow; in both models the precipitation effect is linear. The main difference is that in the NPTF model (5) an AR model is used to model the noise to account for the serial correlation, while in the NAARX model lags of Y_t is used.

To study the forecasting performance of the proposed model, we consider the following post-sample forecast scheme: model (5) is re-estimated using the £rst two years of data, one-step to 12-step ahead post-sample forecasts are conducted using the data of the third year. This forecasting scheme is similar to the one used in Chen and Tsay (1993b), the main difference is that in Chen and Tsay (1993b), actual observations of X_t, Z_t and their lags are used in the forecasts,

while here the forecast values are used. Two simple AR(1) models are found appropriate for this purpose: $X_t = \phi_x X_{t-1} + a_{1t}$ and $Z_t = c_0 + \phi_z Z_{t-1} + a_{2t}$. The mean squared errors (MSE) are calculated and shown in Table 2 under "NPTF". For the purpose of comparison we produce the forecasts using the aforementioned NAARX model (7) and the TAR model (6) under the same setting and report the MSE in Table 2.

The results in Table 2 show that the forecast MSE of the nonparametric transfer function (NPTF) model are consistently smaller than those of the NARRX model and the TAR model. In this example, the proposed NPTF model performances well in both within-sample and post-sample, this shows the good potential of the NPTF model in analyzing nonlinear time series data.

5 Summary and Discussion

In this paper we introduce the regression spline-based nonparametric transfer function model. This method is ¤exible and ideal for modeling highly nonlinear relationships between time series. Ef£cient estimation of the transfer function model is achieved by incorporating the serial correlation in the noise. Compared with the local polynomial-based methods, the explicit functional form of polynomial splines makes estimation much less intensive computationally. This model is used to model river ¤ow based on temperature and precipitation and found successful when compared with widely-accepted nonlinear parametric models.

References

Auestad, B. and D. Tjöstheim, 1990, Identification of nonlinear time series: First order characterization and order estimation. *Biometrika* **77**, 669-687.

Box, G.E.P. and G.M. Jenkins, 1976, *Time Series Analysis: Forecasting and Control*. Holden-Day, San Francisco, 1st ed.

Cai, Z., J. Fan and Q. Yao, 2000, Functional-coefficient regression models for nonlinear time series. *Journal of the American Statistical Association* **95**, 941–956.

Carroll, R.J., J. Fan, I. Gijbels and M.P. Wand, 1997, Generalized partially linear single-index models. *Journal of the American Statistical Association* **92**, 477–489.

Chen, R. and R.S. Tsay, 1996, Nonlinear transfer functions. *Journal of Nonparametric Statistics* **66**, 193-204.

Chen, R. and R.S. Tsay, 1993a, Functional-coefficient autoregressive models. *Journal of the American Statistical Association* **88**, 298-308.

Chen, R. and R.S. Tsay, 1993b, Nonlinear additive ARX models, *Journal of the American Statistical Association* **88**, 955-967.

de Boor, C., 2001. A Practical Guide to Splines. Springer, New York.

Fan, J. and I. Gilbels, 1996, *Local Polynomial Modeling and Its Applications*. Chapman and Hall, Suffolk.

Fan, J. and Q. Yao, 2003, *Nonlinear Time Series: Nonparametric and Parametric Methods*. Springer, New York.

Fan, J., Q. Yao and Z. Cai, 2003, Adaptive varying-coefficient linear models. *Journal of the Royal Statistical Society, Series B* **65**, 57–80.

Hädle, W., H. Lütkepohl, and R. Chen, 1997, A review of nonparametric time series analysis. *International Statistical Review* **65**, 49-72.

Härdle, W., P. Hall, and H. Ichimura, 1993, Optimal smoothing in single-index models. *The Annals of Statistics* **21**, 157–178.

Härdle, W., H. Liang, and J. Gao, 2000, *Partially Linear Models*. Physica-Verlag, Heidelberg.

Haggan V. and T. Ozaki, 1981, Modeling nonlinear vibrations using an amplitude-dependent autoregressive time series model. *Biometrika* **68**, 189196.

Heckman, J., H. Ichimura, J. Smith, and P. Todd, 1998, Characterizing selection bias using experimental data. *Econometrica* **66**, 1017–1098.

Huang, J.Z., 2003, Local Asymptotics For Polynomial Spline Regression. *Ann. Statist.*, **31**, 1600-1635

Ichiruma, H., 1993, Semiparametric least-squares (SLS) and weighted SLS estimation of single-index models. *Journal of Econometrics* **58**, 71–120.

Liu, J.M., R. Chen and Q. Yao, 2007, Nonparametric Transfer Function Models. *Technical report, Georgia Southern University*.

Masry, E., 1996a, Multivariate local polynomial regression for time series: Uniform consistency and rates. *Journal of Time Series Analysis* **17**, 571-599.

Masry, E., 1996b, Multivariate regression estimation: Local polynomial fitting for time series. *Stochastic Processes and Their Applications* **65**, 81-101.

Newey, W.K., and T.M. Stoker, 1993, Efficiency of weighted average derivative estimators and index models. *Econometrica* **61**, 1199-1223.

Ozaki, T., 1985, Statistical identification of storage models with application to stochastic hydrology. *Water Resources Bulletin* **21**, 663-675.

Robinson, P.M., 1983, Nonparametric estimators for time series. *Journal of Time Series Analysis* **4**, 185-207.

Ruckstuhl, A., A.H. Welsh, and R.J. Carroll, 2000, Nonparametric function estimation of the relationship between two repeatedly measured variables. *Statistica Sinica* **10**, 51-71.

Schumaker, L.L., 1981, Spline Functions: Basic Theory. Wiley, New York.

Smith, M., C.M. Wong, and R. Kohn, 1998, Additive nonparametric regression with autocorrelated errors. *Journal of the Royal Statistical Society* **60**, 311-331.

Su, L. and A. Ullah, 2006, More ef£cient estimation in nonparametric regression with nonparametric autocorrelated errors. *Econometric Theory* **22**, 98-126.

Tjøstheim, D., 1994, Nonlinear time series: a selective review. *Scandinavian Journal of Statistics* **21**, 97-130.

Tong, H., 1990, *Non-linear Time Series: A Dynamical System Approach*. Oxford University Press, New York.

Tong, H., B. Thanoon, and G. Gudmundsson, 1985, Threshold time series modeling of two Icelandic river¤ow systems. *Water Resources Bulletin* **21**, 651–661.

Xiao, Z., O.B. Linton, R.J. Carroll, and E. Mammen, 2003, Model ef£cient local polynomial estimation in nonparametric regression with autocorrelated errors. *Journal of the American Statistical Association* **98**, 980-992.