Detecting and Modeling Nonlinearity in the Gas Furnace Data

Houston H. Stokes
Department of Economics (M/C 144)
University of Illinois at Chicago
601 South Morgan Street
Chicago, IL 60680
hhstokes@uic.edu

Melvin Hinich
3902 Cresthill Drive
Austin, TX 78731
hinich@mail.la.utexas.edu

Revised August 11, 1998

Note: Correspondence should be addressed to Houston H. Stokes, Department of Economics (M/C 144), University of Illinois at Chicago, 601 South Morgan Street, Chicago, IL 60607. E-mail address hhstokes@uic.edu

Abstract

The Hinich (1982) bispectrum test for nonlinearity and Gaussianity indicates that the residuals of the Tiao-Box (1981) constrained and unconstrained VAR models for the gas furnace data reject the assumption of Gaussianity and linearity over a grid of bandwidths for estimating the bispectrum. These findings call into question the specification of the linear VAR and VARMA models assumed by Tiao-Box (1981). Utilizing the Hinich (1996) nonlinearity test, the residuals of the VAR model were shown to exhibit episodic nonlinearity. The sensitivity of the results to outliers is investigated by estimating and testing the residuals of L1 and MINIMAX models from 1-6 lags. Building on the linear dynamic specification, a multivariate adaptive regression splines (MARS) model is estimated and shown to remove the nonlinearity in the residuals. Using MARS ANOVA analysis, a parsimonious four variable model was estimated and tested.

KEY WORDS: Bispectrum, VAR, Nonlinear, Episodic Nonlinearity, Gaussian, Skewness Function, MARS, Hinich Test, L1 estimation, MINIMAX estimation

1. **Introduction**

Extensions of the Hinich (1982, 1996) nonlinearity tests are used to test the classic gas furnace data studied by Box-Jenkins (1970, 1976, 1994). This data set was selected by Tiao-Box (1981) when they illustrated a multistep identification strategy for identifying a vector autoregression moving average (VARMA) model from an assumed class of linear models. Stokes (1991, 1997), building on unpublished results contained in Stokes-Hinich (1989), reported that the residuals of the constrained and unconstrained VAR model for the gas data suggested by Tiao-Box (1981) failed the Hinich (1982) Gaussianity and linearity test. Although multiple nonlinear models that could be linearized were tried, the nonlinearity persisted in the residuals. The present paper first shows that the nonlinearity in the residuals is eposodic, as measured by the Hinich (1996) test. The multivariate adaptive regression splines (MARS) model, first proposed by Friedman (1991), is shown to remove the nonlinearities in the gas furnace data model. After a brief discussion of the Hinich bispectrum test, the gas furnace residuals are tested. Next, the model is estimated, using L1 and MINIMAX procedures to determine how sensitive the results are to outliers. The Hinich (1996) test is used to determine if the measured nonlinearities are episodic. Finally, the MARS approach is outlined and MARS models are reported for both the gas input series (GASIN) and the gas output series (GASOUT).

2. **Var Model Setup and Overview of the Hinich Bispectrum Test**

In an influential paper on the identification of VAR and VARMA models, Tiao and Box (1981) used the gas furnace data discussed in Box and Jenkins (1970) as an example. Their model relating the gas input (GASIN) and CO_2 concentration (GASOUT) was

$$A(B)Z_t = u_t, \tag{1}$$

where Z'_t is a row vector of the t^{th} observation of the two series, u'_t is a row vector for period t of the estimated error vector of the model for the 2 series, and A(B) is the k by k autoregressive VAR matrix. Each element in A(B) is a polynomial in the lag operator B. Successful estimation of equation (1) assumes that the roots of A(B) are on or outside the unit circle (the invertibility condition), and that the expected value of the error vectors is zero and the error vectors are pure white noise. Tiao and Box (1981) only tested the error terms for significant autocorrelations and cross correlations. No attempt was made to test whether the linear specification was appropriate. Assuming linearity, Tiao and Box (1981) determined that an unconstrained model of order 6 would clean the residuals of any measurable autocorrelation. Next, they removed nonsignificant VAR coefficients and estimated the constrained model using conditional least squares. We have replicated both their constrained and unconstrained models and their results are reported in Table 1. We next applied the Hinich (1982) test to the residuals of these equations to test for nonlinearity. Before a detailed discussion of these results, a brief discussion of the Hinich (1982) test is given.

If $\{x(t)\}$ is a third-order stationary time series, the mean $\mu_x = E[x(t)]$, the covariance $c_{xx}(m) = E[x(t+m)x(t)] - [\mu_x]^2$, and the general third-order moments $c_{xxx}(s,r)= E[x(t+r)x(t+s)x(t)]$ are independent of t. If $c_{xx}(m) = 0$ for all m not zero, the series is white noise. Priestley (1981) and Hinich and Patterson (1985) stress that although a series may be white noise, if it is not Gaussian, x(n) and x(m) may not be independent. If the distribution of $\{x(n_1),...,x(n_N)\}$ is multivariate normal for all $n_1,...,n_N$, then in addition to being white noise, the series is Gaussian. Hinich and Patterson

(1985, p. 70) fault Box and Jenkins (1970, p. 8 vs. p. 46) and Jenkins and Watts (1968, p. 149 vs. p. 157) for blurring the definitions of whiteness and independence. Many researchers implicitly assume the errors of their models are Gaussian and test for white noise, using the covariance $c_{xx}(m)$, but ignore the information regarding possible nonlinear relationships, which are found in the third-order moments $c_{xxx}(s,r)$.

Define the skewness function $\Gamma(f_1,f_2)$ at frequency pairs (f_1,f_2) in terms of the bispectrum $B_{xxx}(f_1,f_2)$ as

$$\Gamma^2(f_1,f_2) = \quad (B_{xxx}(f_1,f_2))^2 / S_x(f_1)S_x(f_2)S_x(f_1+f_2), \tag{2}$$

where $S_x(f)$ is the power spectrum of $x(t)$ at frequency f. The Hinich (1982) test is based on the Brillinger (1965) proof that the skewness function $\Gamma(f_1,f_2)$ is constant over all frequencies f_1, f_2 in Ω if $\{x(t)\}$ is linear, and zero over all frequencies if $\{x(t)\}$ is Gaussian. This key proof suggests that once a consistent estimator of the bispectrum is calculated, nonlinearity and Gaussianity tests can be performed. For the sample $\{x(0),x(1),...,x(N-1)\}$, define $F_{xxx}(j,k)$ as an estimate of the bispectrum of $\{x(t)\}$ at frequency pair (f_j,f_k) with $f_k = k/N$ for each integer k.

$$F_{xxx}(j,k) = X(f_j)X(f_k)X^*(f_{j+k})/N \tag{3}$$

$X(f_j) = \Sigma_{t=0}^{N-1}x(t)\exp(-i\omega_j t)$, where $\omega_j = 2\pi n/N$ for n=0,1,...,N-1. Hinich (1982) suggests smoothing $F_{xxx}(j,k)$ to form a consistent estimator. Define $<B_{xxx}(m,n)>$ as a estimator of $B_{xxx}(m,n)$, which is calculated by averaging over adjacent frequency pairs of $F_{xxx}(j,k)$.

$$<B_{xxx}(m,n)> = M^{-2}\Sigma^{mM-1}_{j=(m-1)M}\Sigma^{nM-1}_{j=(n-1)M} F_{xxx}(j,k) \tag{4}$$

Equation (4) illustrates that the estimator of the bispectrum is an average value of a square of M points. The larger (smaller) M, the smaller (larger) the finite sample variance and the larger (smaller) the sample bias. Because of this tradeoff, there is no one unique M that is appropriate to use for performing nonlinearity and Gaussianity tests. Hinich (1982) has suggested that a good value for M is the square root of the number of observations. A lower M would be $(N/3)^{.5}$. In the models fit to the Box-Jenkins (1970, 1976, 1994) gas furnace data, where N=296 for the complete sample, we have reported M values from 9 to 18 to insure that our findings are not sensitive to the M value selected. To test the null hypothesis that $\{x(t)\}$ is Gaussian and thus B_{xxx} is identically zero, Hinich suggests the G statistic, which is normally distributed. To test whether the series is linear, or the skewness function $\Gamma(f_1,f_2)$ is constant, Hinich suggests the normally distributed L statistic. For details of this test, see Hinich (1982) , Hinich and Patterson (1985), and Ashley, Patterson and Hinich (1986). In this paper we report normal approximations of the Gaussianity and linearity tests (G and L) for a range of values of M. Since there is no one optimum blocksize, mean values for G and L over all admissible values of M are also reported. Lemos-Stokes (1998) developed empirical critical values for the mean L statistic for 3000 AR(1) models having 2000 observations for Φ_1 = -.9, -.6, -.3, 0.0, .3, .6, and .9 of .121 .672, .789, .874, .811, .636 and -.096 respectively, well below the asymptotic one tail value of 1.64. These results suggest if the asymptotic critical value of 1.64 is used it will be conservative.

Ashley, Patterson and Hinich (1986 p. 174) presented an equivalence theorem which proved that the Hinich bispectal linearity test statistic is invariant to linear filtering of the data. This important result proves that the linearity test can be either applied to the raw series, or the residuals of a linear model. An additional important implication of the theorem is that if X(t) is found to be nonlinear, then the residuals of a linear model of the form Y(t) = f(X(t)) will be nonlinear, since the nonlinearity in X(t) will pass through any linear filter. The above paper also reported tables on the power of the Hinich linearity test for detecting violations of the linearity assumption for a variety of common nonlinear models appearing in the literature and a table of the power of the linearity and Gaussianity tests for a number of sample sizes and M values. The table indicates substantial power for both tests, even when N is a small value, such as 256, if the value of M used is between 12 and 17. For this sample size, as M increases, the power of the test falls off. This is later illustrated in our test results.

Hinich (1996) proposed another testing strategy that could be applied to two series within the sample that was based on the sample cross-correlation at lags r and s, $C_{xy}(r,s)$, and the sample cross bicorrelation, $C_{xxy}(r,s)$. Define m = max(r,s)

$$C_{xxy}(r,s) = (N-m)^{-1}\Sigma_{t=1}^{N-m}x(t)x(t+r)y(t+s) \tag{5}$$

Let $L=N^c$, where 0 < c < .5. Test statistics for non-zero cross correlations and cross bicorrelations are

$$H_{xy}(N) = \Sigma_{r=1}^{L} (N-r)C_{xy} (r) \tag{6}$$

$$H_{xxy}(N) = \Sigma_{s=-L}^{L} \Sigma_{r=1}^{L}(N-m) C_{xxy} (r,s), \tag{7}$$

where $s \neq 0$.

The $H_{xy}(N)$ and $H_{xxy}(N)$ are asymptotically chi-squared with L(2L+1) degrees of freedom but for the purposes of this paper have been transformed to U(0,1) under the null. The Hinich (1996) test has a number of advantages that include being able to test for nonlinearity mapping from one series to another and being relatively quick to calculate. This latter advantage allows the test to be performed within the sample to test for episodic nonlinearity. In the empirical section the probabilities of $H_{xy}(N)$

and $H_{xxy}(N)$ are given. For testing one series, say x, Hinich (1996, eq. 3.1) recommends using a varianiant of (5). Define G(r,s) as the r,s sample bicorrelation multiplied by $(N\text{-}s)^{-.5}$ to standardize its variance. The statistic H_x which tests the series x for nonlinearity, is defined as

$$H_x = N^{-c}\Sigma^{L}_{s=2}\Sigma^{s-1}_{r=1}[G\ (r,s)-1] \qquad (8)$$

where 0<c<.5, is N(0,1) which can be converted to a probability of rejection of the assumption of linearity.

3. **Results**

In Table 2 the Hinich bispectrum tests are performed on a grid of M values, going from 9 to 18, to test the residuals from the two equations implicit in (1) for both Gaussianity (G test) and linearity (L test). The G test values are all above 4.99 for GASIN and 10.92 for GASOUT, indicating that <u>both</u> residual series, for both constrained and unconstrained models, reject the assumption of Gaussianity (G test) at a very high level of significance. For virtually all values of M, the <u>assumption of linearity is also rejected</u> (L test). The lower L scores were found only with the higher M scores (17 and 18), which have a large bandwidth. For purposes of comparison, e e for the GASIN and GASOUT residuals was 9.8847 and 16.139, respectively. As was mentioned earlier, Ashley, Patterson and Hinich (1986 Table I and II) investigated the size (number of observations) needed for the Hinich linearity and Gaussianity tests and the power of such tests for various values of M and number of observations. Their findings indicate that both tests give satisfactory convergence and that both tests detect nonlinearity with considerable frequency, even in cases in which N=256. These simulation results suggest that it is appropriate to use the Hinich tests in the present case in which N=290.

Our findings of nonlinearity are invariant as to whether the estimated form of equation (1) is unconstrained or constrained. In results not reported, we experimented with increasing the lag length of the unconstrained VAR model from 6 to 12. The results were similar. We conclude that even though the distribution of the Hinich tests is known only asymptotically, the magnitude of the Z scores indicates (at a high confidence level) that both the input series and the output series fail the null hypotheses of Gaussianity and linearity, which were assumed by Tiao and Box for the gas furnace data[2]. If we had found that only the output series residuals were nonlinear, it might be possible to select a linear model that would both improve the sum of squares of the residuals and reduce the measured nonlinearity in the output series residuals. In the present case both input and output residuals failed the nonlinearity test. In this situation, it is impossible to identify a linear model that would transform the output series residuals so that they no longer fail the linearity tests, because the nonlinearity in the input series residual will pass through any linear filter to the output series. This statement is based on the powerful equivalence theorem tested and proved using simulations by Ashley, Patterson and Hinich (1986). Their theorem proved that the Hinich linearity test could be applied to either the raw series or the residuals of a linear model, since the nonlinearity would pass through any linear model.

In an attempt to remove the indications of nonlinearity in the test statistics, an exhaustive search of alternative nonlinear models that could be linearized was attempted without success. All estimations were done, assuming maximum likelihood or that the sum of the squared residuals were minimized. To investigate the possibility that outliers were giving the illusion of nonlinearity, L1 and MINIMAX models were tried. L1 models minimize $\Sigma \; e_t$, and are less sensitive to outliers than OLS models. MINIMAX models minimize max e_t , and are more sensitive to outliers than OLS models. Using all three estimation methods, six alternative models of GASOUT of the form

$$GASOUT_t = \alpha + \beta_0 GASIN_t + \Sigma^{k}_{i=1} GASIN_i + \Sigma^{k}_{i=1} GASOUT_i \tag{9}$$

were tried for lags of k= {1,...,6} and the results are reported in Table 3. Gaussianity is rejected for lags 1-6 for all estimators. Linearity is rejected for lags 2-6 for L1 and OLS. The MINIMAX model rejects linearity for models with 3 lags and 4 lags. In the MINIMAX model, the cost of reducing the maximum e_t is reflected by e e values that are two or more times bigger than their OLS and L1 counterparts. MINIMAX models with 5 lags and 6 lags have no indication of nonlinearity, but have e e values that are two times their OLS counterparts (32.36 vs 16.44 for lag 5 and 30.41 vs 16.09 for lag 6). Since L1 models are less sensitive than OLS models to outliers, the poor performance of L1 in removing the nonlinearity suggests that outliers are not tripping the Hinich (1982) test.

The next experiment was to test whether the measured nonlinearity varies over time. Table 4 uses the Hinich (1996) test to investigate the within-sample properties of the residuals of the VAR(6) model of the gas furnace data series given in Table 1. Two window sizes of 20 and 30 were used. For notational simplicity define x=the residuals of GASIN and y=the residuals of GASOUT. Hx and Hy measure the probability of nonlinearity remaining in the residuals of the GASIN and GASOUT models, respectively. Hxy and Hyx measure the probability of there being a nonlinear relationship between the residuals of GASIN to the residuals of GASOUT or the residuals of GASOUT to the residuals of GASIN, respectively. Px, Py and Pxy measure the probability of autocorrelation in the GASIN residuals, the GASOUT residuals and between the GASIN and GASOUT residuals, respectively. Inspection of Hx and Hy for both windows shows periods of nonlinearity in each series. Using a window size of 20, Hx was .92 and .90 for windows 1 and 2, respectively, which included observations 1-40. Using a window size of 30, Hx was .94 in window 1, covering observations 1-30. For the 20-observations window, Hx and Hy were, respectively, .94 and .91 in window 6, covering observations 101-120. For the same window Hxy and Hyx were .9995 and .9837, respectively. Using the 30-observations window, Hx and Hxy had values of .98 and .94, respectively, for window 4, covering observation 91-120, indicating that there was a relationship between the nonlinearity in the GASIN series residual and GASOUT series residual. The autocorrelations and cross correlations of the residual series for the complete sample are flat. However, for the subsamples significant values show up. For the 20- observations window, Px is significant for windows 1 (.93), 9 (.93) and 13 (.97). Py is significant for windows 5 (.97), 7 (.99) and 14 (.99). For the 30-observations window, Px was significant for window 9 (.96) and Py was significant for window 3 (.95), 5 (.99), and 6 (.99). Pxy was significant for windows 2 (.96), 8 (.97) and 9 (.98). The results of the subsample estimation suggest that there are episodic periods of both nonlinearity and linear memory in the model. These results suggest that it might be promising to attempt a modeling strategy that includes a quite general class of threshold models to remove the nonlinearity. The technique chosen was the MARS model, a general data-driven, nonparametric approach, which has had success in other recent applications.[3] These results are discussed next.

4. **MARS Results**

The results in the preceding section indicate that the residuals in both the constrained and unconstrained models of the gas furnace data suggested by Tiao-Box (1981) fail the Hinich (1982) tests for nonlinearity and Gaussianity. The residuals of the unconstrained VAR model for GASOUT and GASIN are graphed in Figures 1 and 2. Looking first at the GASOUT series shown on the bottom of Figure 1, note the relatively homogenous pattern for the first 60% of the series, with larger spikes at observation 193 of .975, observation 230 of -.7526 and observation 259 of 1.431. The GASIN series graphed in the bottom of Figure 2 shows a number of outliers at observation 37 of 1.016 and observation 49 of -.8851. The task is to select an estimation method having VAR as a special case, but allowing level-specific function changes that might be able both to reduce the variance of the residual and model some of these outliers. The MARS approach that we have selected is discussed next.

The MARS technique assumes a nonlinear model of the form

$$y = f(x_1,...,x_m) + e, \tag{10}$$

involving N observations on m right-hand-side variables, $x_1,...,x_m$, which are column vectors in the N by m matrix X. The function f(X) is approximated by

$$\hat{f}(X) = \Sigma_{j=1}^{s} c_j k_j(X), \tag{11}$$

where $\hat{f}(X)$ is an additive function of the product basis functions $\{K_j(X)\}_{j=1}^{s}$ associated with the s disjoint subregions $\{R_j\}_{j=1}^{s}$ of D ($D = \cup_{j=1}^{s} R_j$) and c_j is the coefficient for the jth product basis function. OLS is a special case of MARS if all subregions include the <u>complete range</u> of each of the right-hand-side variables. In this situation, the coefficients $\{c_j\}_{j=1}^{s}$ can be interpreted as OLS coefficients of variables. The MARS approach identifies the subregions under which the coefficients are stable and detects any possible interactions up to a maximum number of possible interactions controllable by the user. As an illustration, assume the model

$$\begin{aligned} y &= \alpha + \beta_1 x + e \quad \text{for } x > 100 \\ &= \alpha' + \beta_2 x + e \quad \text{for } x < 100. \end{aligned} \tag{12}$$

The MARS approach writes this model as

$$y = \alpha'' + c_1(x-\tau^*)_+ - c_2(\tau^*-x)_+ + e, \tag{13}$$

where $\tau^* = 100$ and $(\ )_+$ is the right (+), truncated spline function, which takes on the value 0 if the expression inside $(\ )_+$ is negative and its actual value if the expression inside $(\ )_+$ is ≥ 0. Here $c_1 \equiv \beta_1$ and $c_2 \equiv \beta_2$, $\alpha \equiv \alpha''-c_1\tau^*$, $\alpha' \equiv \alpha''-c_2\tau^*$. In terms of equation (2), $K_1(X) \equiv (x-\tau^*)_+$ and $K_2(X) \equiv (\tau^*-x)_+$ and $K_3(x) \equiv 1$. In contrast to other spline approaches that require the user to specify the knots of τ^*, the MARS algorithm produces an estimate of the knot. If all knots are found to be at the minimum of the x variable, then the MARS algorithm has signalled that OLS is the correct estimation procedure. Thus, the VAR model is a special case of MARS.

The derivative of the spline function is not defined for values of x at the knot value of 100. Friedman (1991) suggests using either a linear or cubic approximation to determine the exact y value. In the results reported later, both evaluation techniques have been tested and the one with the lowest sum of squares of the residual has been selected. The MARS user selects the maximum number of knots to consider and the highest-order interaction to investigate. Alternatively, the minimum number of observations between knots can be set. An example of a more complex model for $y = f(x,z)$ is given next.

$$y = \alpha + c_1(x-\tau_1^*)_+ + c_2(\tau_1^*-x)_+ + c_3(x-\tau_1^*)_+(z-\tau_2^*)_+ + e \tag{14}$$

implies that

$$\begin{aligned} y &= \alpha + c_1x - c_1\tau_1^* + e \quad &&\text{for } x > \tau_1^* \text{ and } z \leq \tau_2^* \\ &= \alpha - c_2x + c_2\tau_1^* + e \quad &&\text{for } x \leq \tau_1^* \\ &= \alpha + c_1x - c_1\tau_1^* + c_3(xz-\tau_1^*z-\tau_2^*x+\tau_1^*\tau_2^*) + e \end{aligned}$$

$$\text{for } x > \tau_1^* \text{ and } z > \tau_2^*. \tag{15}$$

As an aid in determining the degree of model complexity, Friedman (1991) suggests using a modified form of the generalized cross validation criterion (MGCV).

$$MGCV = [(1/N)\Sigma_{i=1}^{N}(y_i - \hat{f}(X_i)) \;] / [1 - [C(M)^*/N] \;], \tag{16}$$

where there are N observations, $\hat{f}(X_i) \equiv \hat{y}_i$ and $C(M)^*$ is a complexity penalty. The default is to set $C(M)^*$ equal to a function of the effective number of parameters. The formula used is

$$C(M)^* = C(M) + dM. \tag{17}$$

The parameter d, which is user-controlled, has been set to the default value of 3 as suggested by Friedman (1991). C(M) is the number of parameters being fit and M the number of nonconstant basis functions in the model. The MARS approach starts by choosing where to place the knots for a noninteraction model. Next, more complex interactions are chosen up to a user-controlled maximum number of interactions and maximum number of parameters in the model. Once the forward selection is completed, the MGCV statistic is used to eliminate parameters that improve the model only slightly. The MGCV value controls how many parameters finally remain in the model and can be used to form an estimate of the relative importance of each x_i variable in the model.[4]

The MARS technique requires that the user select the variables $x_1,...x_m$ to use in (10). Since the gas furnace data model involves lags, an immediate concern is how to select the appropriate lags of GASIN and GASOUT to place in the $x_1,...,x_m$ vector. The technique proposed in this paper is first to use a VAR model, such as proposed by Tiao-Box (1981), to determine the maximum number and placement of the lags of x and y to estimate a VAR model of the series in the linear domain. If the resulting residuals show evidence of nonlinearity, as measured by the Hinich (1982) test, then these would first be used on the right-hand side of the MARS model equation. In the second MARS stage, a simplified model containing only those variables found important with the first-round MARS estimation would be used. It must be emphasized that such a procedure would not be strictly appropriate if evidence of feedback were found in the VAR or VARMA step of the model. In a VAR model, since contemporaneous effects (instantaneous causality) are seen in the off-diagonal elements of the covariance matrix, contemporaneous values of some X variables must be included on the right-hand side of the MARS equation if the VAR model has identified instantaneous causality.

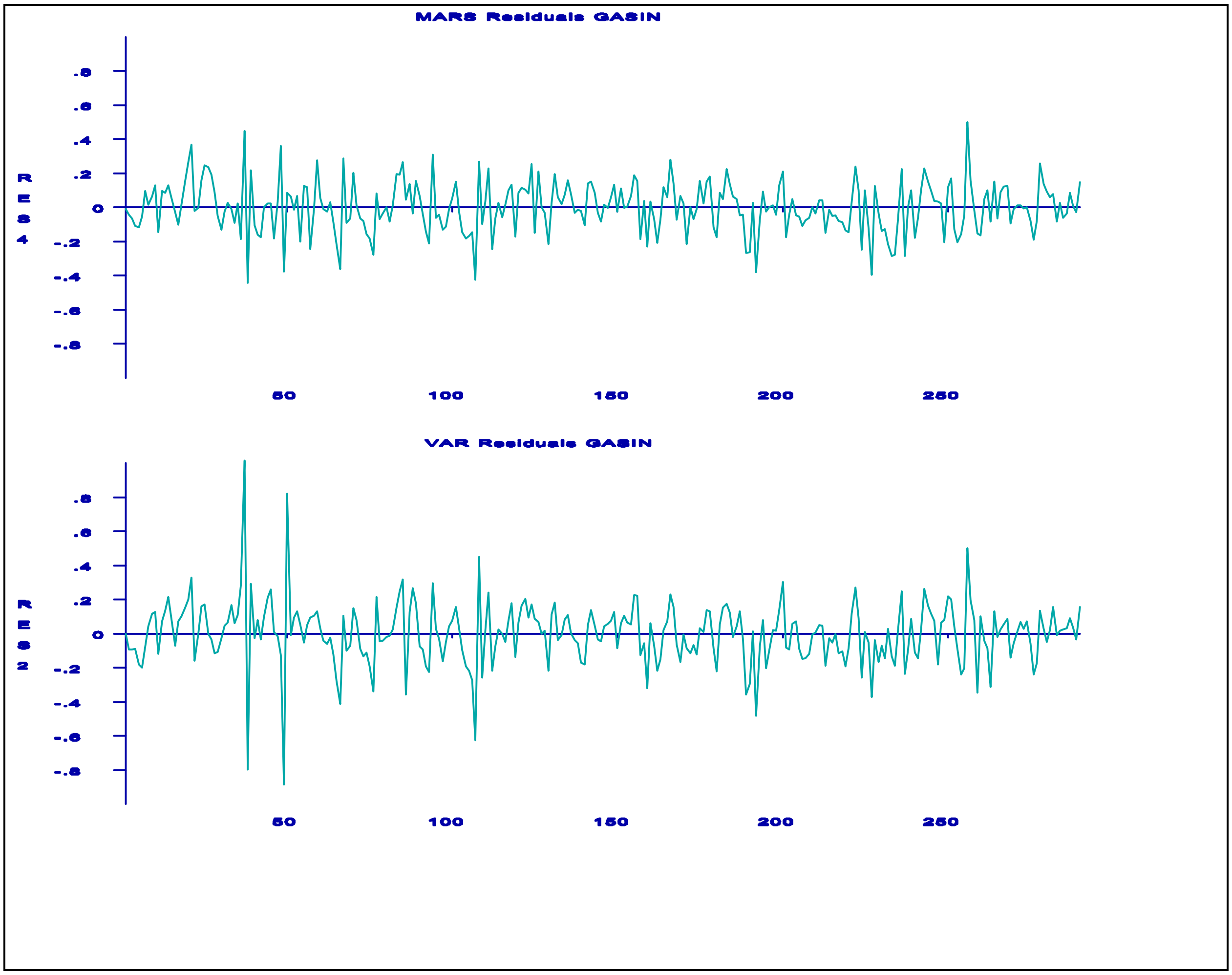

Figure 1 VAR and MARS GASIN Residuals

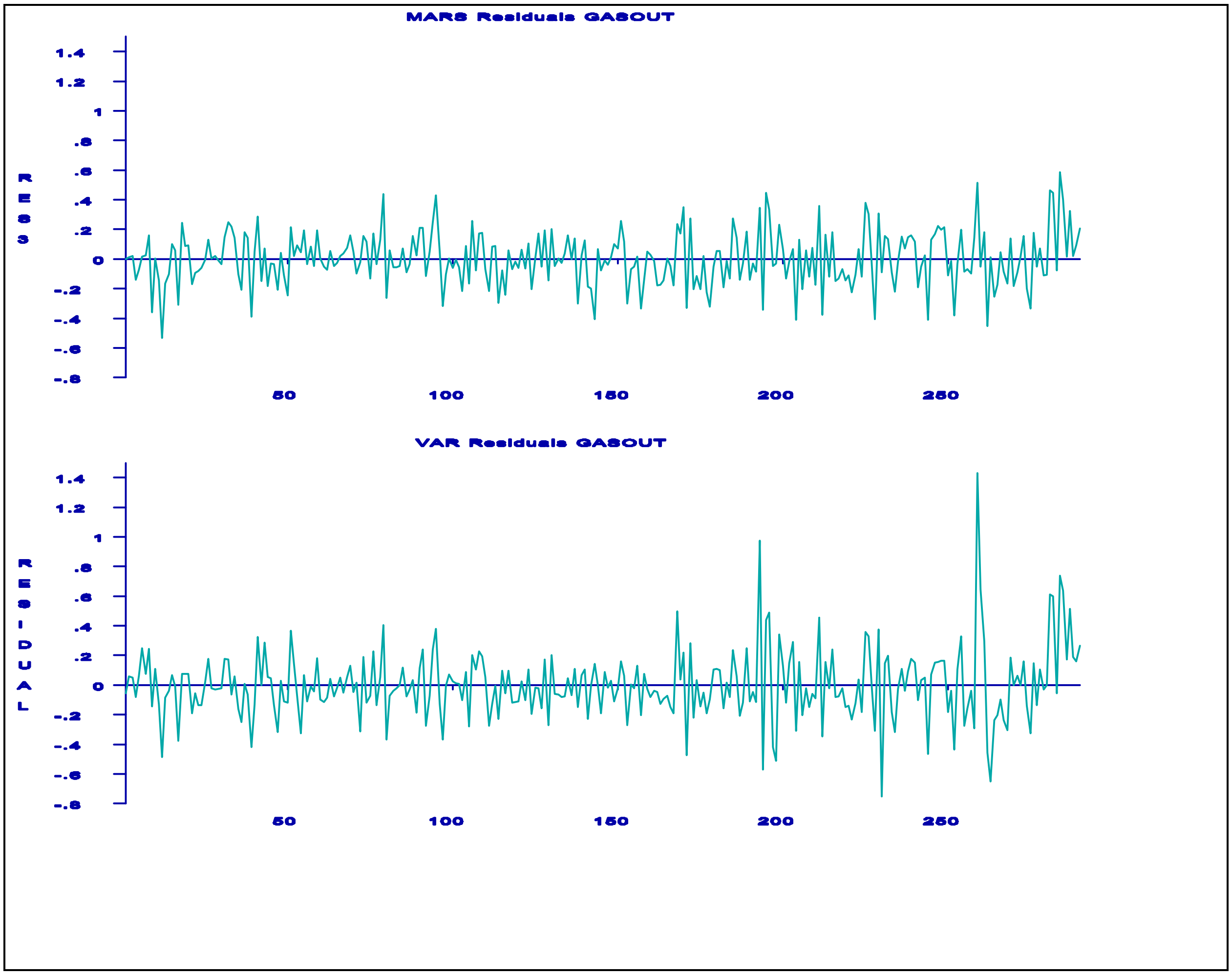

Figure 2 VAR and MARS GASOUT Residuals

A MARS model with 40 knots and a maximum level of four interactions was estimated for GASOUT and GASIN and the results are shown in Tables 5 and 6. Note that the mean G and L values for GASOUT was now .2740 and -.2253, respectively, indicating that the residuals of the MARS procedure no longer fail the Gaussianity and nonlinearity test. For GASIN the G and L values were 1.815 and 1.120, respectively. Plots of the residuals of the VAR and MARS models and GASOUT and GASIN series are shown in Figures 1 and 2, respectively. Inspection of the plots indicates how the MARS estimation method has both reduced the variance of the residuals and reduced some of the extreme values. e e values for the GASIN and GASOUT models are now 6.25 and 9.65, respectively, less than the unconstrained VAR modelvalues of9.88 and 16.14. The exact model has not been listed to same space but the effect of various variables and interactions of variables on the MGCV value are shown. Looking first at GASOUT model, we note that the four most important variables, in order, are $GASOUT_{t-1}$ (.4715), $GASOUT_{t-2}$ (.1550), $GASIN_{t-3}$ (.1445) and $GASIN_{t-6}$ (.0565), where the value in parentheses is the change in the MGCV if the variable was removed. For GASIN the top two variables were $GASIN_{t-1}$ (.3119) and $GASIN_{t-2}$ (.0779). The MARS procedure has confirmed that the GASIN series is not affected by the GASOUT series. Tables 5 and 6 also give ANOVA analysis of the effect of interactions on the MGCV. For GASOUT the most important two-way interactions are $GASOUT_{t-1}$ and $GASOUT_{t-2}$.

A simplified GASOUT model containing only the four most important variables estimates GASOUT = f($GASOUT_{t-1}$, $GASOUT_{t-2}$, $GASIN_{t-3}$, $GASIN_{t-6}$) and is reported in Table 7. Here the e e has increased from 9.648 to 12.10. The mean G indicates lack of Gausianity but the mean L of .97 indicates that there is no nonlinearity in the residual. The e e value of 12.10 is substantially less than the OLS VAR e e of 16.139 in Table 1. These findings suggest that a parsimonious MARS model that is able to model a linear and nonlinear process can outperform a VAR model with many variables. The Hinich (1982) test allows us to gage whether the MARS model has been successful.

5. **Conclusion**

The Hinich (1982) test was used to test the adequacy of the linearity assumptions in the classic Box-Jenkins (1970, 1976, 1994) gas furnace data. After finding evidence of nonlinearity, various linearizable, nonlinear models were tried with out success. L1 and MINIMAX estimation models were used to determine if the measured nonlinearity was sensitive to outliers where the L1 (MINIMAX) model is less (more) sensitive to outliers than OLS models. The L1 models of GASOUT were found not to remove the measured nonlinearity in the residuals. In four out of six MINIMAX models, measured nonlinearity was removed at the cost of relatively large e e values. The Hinich (1996) test, applied to subsamples of residuals, indicated that the nonlinearity was episodic. A MARS model was shown to remove the measured nonlinearity in the model residuals for GASIN and GASOUT and produce a close fit. An ANOVA decomposition of the model was used to focus on the relative importance of the right-hand-side variable and various interactions. Based on this analysis, a parsimonious MARS model for GASOUT was estimated, using only four variables, and was found both to remove the nonlinearity in the GASOUT residual and have a e e value that was substantially less than the best VAR representation

Table 1 Estimated Coefficients for Unconstrained and Constrained Models for the Tiao-Box Gas Furnace Data

	Unconstrained Model				**Constrained Model**			
Lag	$A_{1,1}$	$A_{1,2}$	$A_{2,1}$	$A_{2,2}$	$A_{1,1}$	$A_{1,2}$	$A_{2,1}$	$A_{2,2}$
1	1.93*	-.0508	.0632	1.55*	1.982*			1.522*
	(.0581)	(.0457)	(.0743)	(.0585)	(.055)			(.0571)
2	-1.20*	.0999	-.133	-.593*	-1.387			-.568
	(.126)	(.0843)	(.161)	(.108)	(.0998)			(.1063)
3	.17	-.0796	-.441	-.171	.349*		-.530	-.159
	(.144)	(.0881)	(.184)	(.113)	(.0551)		(.0741)	(.0997)
4	- .16	.0269	.152	.132			.1180	.1312
	(.145)	(.0877)	(.186)	(.112)			(.1631)	(.0431)
5	.38*	-.0414	-.120	.0569			-.0451	
	(.137)	(.0771)	(.175)	(.0985)			(.1734)	
6	-.214*	.0305	.249*	-.0421			.2091	
	(.0839)	(.0328)	(.107)	(.0419)			(.1056)	
S	.03408				.03593			
	-.00229	.0557			-.00290	.056143		

Estimation was done using conditional least squares, using a variant of the WMTS-1 program of Tiao, Box, et. al (1979) contained in B34S Stokes(1997). Standard errors are listed parenthetically. Constants were estimated for the constrained problem and were -.004138 and 3.9992, with standard errors .01115 and .8335 for equations 1 and 2, respectively. S = residual error covariance matrix. e e for the unconstrained GASIN and GASOUT model was 9.88 and 16.139, respectively.

Table 2 Z scores for Gaussianity and Linearity tests for Unconstrained and Constrained 6th-Order VAR models for Gas Data

	Equation 1				**Equation 2**			
	Unconstrained		**Constrained**		**Unconstrained**		**Constrained**	
M	G	L	G	L	G	L	G	L
9	10.86	5.80	11.76	7.75	11.27	5.53	11.63	5.81
10	12.07	6.30	12.53	5.91	11.86	4.52	12.05	6.34
11	7.05	6.77	8.62	7.30	11.24	5.29	11.35	6.00
12	12.75	3.08	12.77	7.01	12.22	6.27	12.21	4.87
13	5.99	2.63	6.94	2.74	11.19	4.00	11.41	5.81
14	7.51	1.45	8.07	1.37	10.92	8.45	11.36	4.75
15	4.99	4.04	4.98	1.98	11.45	3.27	11.58	6.21
16	6.47	3.40	6.95	8.19	12.91	3.46	12.91	2.49
17	7.63	1.11	9.30	6.59	12.05	.16	12.39	.99
18	6.48	.60	6.90	.63	12.46	4.23	12.67	4.06
Mean	8.18	3.52	8.88	4.95	11.76	4.52	11.96	4.73

G = Z score for normal approximation for Gaussianity test. L = Z score for linearity test. M = square root of the number of terms used to estimate the bispectrum at the center of the square. The number of residuals was 290. Equation 1 is for the gas input data (Box-Jenkins (1970)). Equation 2 is for the gas output data (Box-Jenkins (1970)). Coefficients for the unconstrained (Model 1) and constrained (Model 2) are given in Table 1. Estimated coefficients are consistent with those of Tiao-Box (1981).

Table 3 Alternative Estimators of the Linear Gas Furnace Data

	L1		**OLS**		**MINIMAX**	
Lag	**G**	**L**	**G**	**L**	**G**	**L**
1	16.10	-.71	16.15	-.18	8.00	-3.39
2	21.44	10.06	17.01	8.89	19.23	1.48
3	26.95	7.31	16.15	4.26	36.48	2.33
4	16.41	6.61	13.61	5.17	7.96	2.89
5	15.68	7.35	11.54	4.68	9.92	.28
6	14.56	5.10	11.51	4.38	17.03	.13

	e e	**max|e|**	**Σ |e|**	**e e**	**max|e|**	**Σ|e|**	**e e**	**max |e|**	**Σ|e|**
1	76.68	2.47	111.37	76.37	2.47	111.72	158.99	1.73	176.56
2	20.25	1.53	51.48	19.60	1.54	52.54	52.69	.91	105.01
3	19.43	1.47	50.23	18.02	1.50	51.32	48.59	.84	104.32
4	17.26	1.54	47.83	16.71	1.49	48.76	37.61	.79	85.09
5	17.57	1.55	47.53	16.44	1.40	48.89	32.36	.73	77.66
6	16.82	1.53	47.06	16.09	1.41	48.28	30.41	.73	74.78

The functional form of all models included a contemporaneous GASIN value and from 1 - 6 lags of GASIN and GASOUT mapping to GASOUT. The L1 estimator minimizes the absolute error. The MINAMAX estimator minimizes the maximum absolute error.

Table 4. Episodic Nonlinearity Tests on Residuals of VAR(6) Model of the Gas Furnace Data

20-Observations Window

Window	Obs Begin	Obs End	Hx	Hy	Hxy	Hyx	Px	Py	Pxy
1	1.000	20.00	.9161	.2585	.8313	.1657	.9319	.0298	.8673
2	21.00	40.00	.9023	.2420	.9977	.2821	.6265	.5210	.7833
3	41.00	60.00	.5499	.1330	.9882	.9550	.4752	.0829	.7788
4	61.00	80.00	.3074	.3405	.4314	.7664	.0388	.9000	.7628
5	81.00	100.0	.6689	.5759	.1725	.8941	.5412	.9743	.4024
6	101.0	120.0	.9445	.9145	.9995	.9837	.4257	.4402	.9271
7	121.0	140.0	.1372	.1906	.03957	.0240	.6553	.9899	.1687
8	141.0	160.0	.7008	.5053	.4304	.0192	.5055	.7427	.0988
9	161.0	180.0	.5193	.8839	.4963	.8676	.9271	.9729	.7095
10	181.0	200.0	.7077	.9568	.1443	.8591	.8089	.7413	.7274
11	201.0	220.0	.2225	.3257	.5370	.9212	.8503	.9087	.6343
12	221.0	240.0	.0353	.4755	.0503	.2564	.4339	.3199	.9025
13	241.0	260.0	.0461	.1492	.7665	.8273	.9695	.1152	.9379
14	261.0	290.0	.3780	.4675	.0046	.1469	.1768	.9986	.5782

30-Observations Window

Window	Obs Begin	Obs End	Hx	Hy	Hxy	Hyx	Px	Py	Pxy
1	1.000	30.00	.9401	.5035	.5274	.5016	.8710	.1031	.9428
2	31.00	60.00	.1143	.2178	.1715	.8537	.8200	.1457	.9589
3	61.00	90.00	.6934	.5751	.3824	.7185	.1962	.9526	.1077
4	91.00	120.0	.9829	.8760	.9427	.7067	.2149	.7155	.8948
5	121.0	150.0	.4982	.3916	.2873	.5187	.6239	.9851	.1880
6	151.0	180.0	.1118	.9450	.5312	.7611	.8083	.9856	.4810
7	181.0	210.0	.9247	.9925	.4945	.9901	.8557	.8896	.8737
8	211.0	240.0	.2800	.5237	.2165	.2548	.4533	.4997	.9695
9	241.0	290.0	.6338	.8891	.9999	1.000	.9579	.8227	.9774

Hx and Hy measure the probability on nonlinearity in x and y, respectively. Hxy measures the probability of nonlinearity in x being reflected in y, while Hyx measures the probability of nonlinearity in y being reflected in x. Px and Py measure the probability of autocorrelation in x and y, respectively while Pxy measures the probability of cross correlation between x and y.

Table 5. MARS Model for GASOUT Series

$GASOUT_t$ = f(GASIN{1 to 6} GASOUT{1 to 6})

MGCV calculated, using linear method = 44.1.

-MGCV removing each variable (relative importance listed under value)

$GASIN_{t-1}$	$GASIN_{t-2}$	$GASIN_{t-3}$	$GASIN_{t-4}$	$GASIN_{t-5}$	$GASIN_{t-6}$
.0478	.0463	.1445	.0590	.0463	.0565
5.94%	0.0%	48.06%	17.30%	0.00%	15.51%

$GASOUT_{t-1}$	$GASOUT_{t-2}$	$GASOUT_{t-3}$	$GASOUT_{t-4}$	$GASOUT_{t-5}$	$GASOUT_{t-6}$
.4715	.1550	.0615	.0590	.0508	.0473
100.0%	50.56%	18.90%	17.30%	10.33%	4.99%

ANOVA decomposition of 16 Basis Functions

Function	MGCV	#	Variables			
1	.5092	1	$GASOUT_{t-1}$			
2	.0969	1	$GASOUT_{t-2}$			
3	.1295	2	$GASIN_{t-3}$			
4	.0540	1	$GASIN_{t-6}$			
5	.0539	2	$GASOUT_{t-1}$,	$GASOUT_{t-2}$		
6	.0508	1	$GASIN_{t-6}$,	$GASOUT_{t-2}$		
7	.0503	1	$GASOUT_{t-1}$,	$GASOUT_{t-3}$,	$GASOUT_{t-5}$	
8	.0473	1	$GASOUT_{t-1}$,	$GASOUT_{t-5}$,	$GASOUT_{t-6}$	
9	.0494	3	$GASIN_{t-3}$,	$GASIN_{t-4}$,	$GASOUT_{t-2}$	
10	.0478	1	$GASIN_{t-1}$,	$GASIN_{t-3}$,	$GASIN_{t-4}$,	$GASOUT_{t-4}$
11	.0578	1	$GASIN_{t-3}$,	$GASIN_{t-4}$,	$GASOUT_{t-2,}$	$GASOUT_{t-3}$
12	.0580	1	$GASIN_{t-3}$,	$GASIN_{t-4}$,	$GASOUT_{t-2,}$	$GASOUT_{t-4}$

e e using piecewise-linear MARS model = 9.648
Mean for G = .2740
Mean for L = -.2253

MARS model was estimated using a maximum number of knots of 40 and a maximum level of interactions of four. For variable descriptions, see text.

Table 6. MARS Model for GASIN Series

$GASIN_t$ = f(GASIN{1 to 6}, GASOUT{1 to 6})

MGCV calculated using linear method = .03155

-MGCV removing each variable (relative importance listed under value)

$GASIN_{t-1}$	$GASIN_{t-2}$	$GASIN_{t-3}$	$GASIN_{t-4}$	$GASIN_{t-5}$	$GASIN_{t-6}$
.3119	.07787	.04012	.03295	.03481	.03450
100%	40.08%	17.24%	6.964%	10.64%	10.11%

$GASOUT_{t-1}$	$GASOUT_{t-2}$	$GASOUT_{t-3}$	$GASOUT_{t-4}$	$GASOUT_{t-5}$	$GASOUT_{t-6}$
.03155	.03155	.03506	.03155	.03657	.03532
0.0%	0.0%	11.03%	0.0%	13.19%	11.43%

ANOVA decomposition of 18 Basis Functions

Function	MGCV	#	Variables
1	.2094	1	$GASIN_{t-1}$
2	.05987	2	$GASIN_{t-2}$
3	.03204	1	$GASIN_{t-5}$, $GASOUT_{t-6}$
4	.03295	1	$GASIN_{t-4}$, $GASIN_{t-5}$
5	.03759	3	$GASIN_{t-1}$, $GASIN_{t-2}$
6	.04091	2	$GASIN_{t-1}$, $GASIN_{t-3}$
7	.03741	1	$GASIN_{t-2}$, $GASIN_{t-3}$
8	.03358	1	$GASIN_{t-5}$, $GASOUT_{t-5}$, $GASOUT_{t-6}$
9	.03321	1	$GASIN_{t-3}$, $GASIN_{t-5}$, $GASOUT_{t-6}$
10	.03226	1	$GASIN_{t-1}$, $GASIN_{t-5}$, $GASOUT_{t-6}$
11	.03506	1	$GASIN_{t-1}$, $GASIN_{t-2}$, $GASOUT_{t-3}$
12	.03450	1	$GASIN_{t-1}$, $GASIN_{t-6}$, $GASOUT_{t-5}$
13	.03227	2	$GASIN_{t-2}$, $GASIN_{t-3}$, $GASIN_{t-5}$, $GASOUT_{t-6}$

e e using piecewise-linear MARS model = 6.2525
Mean for G = 1.815
Mean for L = 1.120

Mars model was estimated using a maximum number of knots of 40 and a maximum level of interactions of 4. For variable descriptions, see text. The minimum for $GASIN_{t-i}$ and $GASOUT_{t-i}$ for i=1,...,6 was -2.716 and 45.60, respectively. G=1.815, L=1.120.

Table 7 Simplified MARS GASOUT Model

GASOUT = f($GASIN_{t-3}$, $GASIN_{t-6}$, $GASOUT_{t-1}$ $GASOUT_{t-2}$)

MGCV calculated using linear method = 0.5977E-01

-MGCV removing each variable (relative importance listed under value)

$GASIN_{t-3}$	$GASIN_{t-6}$	$GASOUT_{t-1}$	$GASOUT_{t-2}$
0.1733	0.07266	0.5809	0.3167
46.67%	15.73%	100.0%	70.22%

ANOVA decomposition on 16 basis functions:

Function	MGCV	#	Variable(s)		
1	0.2166	1	$GASOUT_{t-1}$		
2	0.07411	1	$GASOUT_{t-2}$		
3	0.1124	1	$GASIN_{t-3}$		
4	0.06673	3	$GASIN_{t-6}$,	$GASOUT_{t-2}$	
5	0.06025	1	$GASOUT_{t-1}$, $GASOUT_{t-2}$		
6	0.07182	2	$GASIN_{t-6}$,	$GASOUT_{t-1}$	
7	0.07094	1	$GASIN_{t-3}$,	$GASOUT_{t-2}$	
8	0.07169	4	$GASIN_{t-6}$,	$GASOUT_{t-1}$	$GASOUT_{t-2}$
9	0.06081	1	$GASIN_{t-3}$,	$GASIN_{t-6}$,	$GASOUT_{t-1}$
10	0.06732	1	$GASIN_{t-3}$,	$GASOUT_{t-1}$	$GASOUT_{t-2}$

Piecewise cubic fit on 16 basis functions, MGCV = 0.7467E-01.
Sum of squared residuals using piecewise-linear MARS model = 12.1034255067554.

Mean for G = 3.7877610.
Mean for L = 0.97401226.

Footnotes

The authors thank Diana Stokes for editorial assistance. Any remaining errors are our responsibility.

1. Estimation was performed, using a modified form of the WMTS-1 model developed by Tiao, Box, Grupe, Hudak, Bell and Chang (1979) contained in B34S (Stokes 1991, 1997). The authors of this paper thank the developers of WMTS-1 for making the code available.

2. One of the authors of this paper discussed these results informally with George Box. It was his impression that the finding that the gas input series was nonlinear was highly plausible in view of the data source.

3. The MARS approach was first proposed by Friedman (1991) and was used successfully by Lewis and Stevens (1991) in a time series context. Lewis and Stevens (1991) and Lewis and Ray (1997, p. 883) showed how the Tong (1983, 1990) threshold autoregression (TR) model could be extended to fit nonlinear threshold models that are continuous in the domain of the predictor variables, allow interactions among lagged predictor variables and have multiple lagged predictor variable thresholds, thus overcoming the limitations of Tong s approach and admitting more general nonlinear threshold models. They called their approach TSMARS, or MARS related to time series. In related work, Chen and Tsay (1993a) used arranged, local regression models to model the chickenpox data, the sunspot data and simulated series. Chen and Tsay (1993b) used a nonlinear additive autoregressive model with exogenous variables for nonlinear time series model

fitting. The MARS model can be thought of as a generalization of this approach in that higher order interaction terms are allowed.

4. Stokes (1997) provides a more complete discussion of the MARS model. This description is based on that treatment.

References

Ashley, Richard, Patterson, Douglas M., and Hinich, Melvin. J. (1986) A Diagnostic Test for Nonlinear Serial Dependence in Time Series Fitting Errors. *Journal of Time Series Analysis* 7 no. 3, 165-178.

Box, G. E. P., and Jenkins, G. M. (1970, 1976, 1994) *Time Series Analysis- Forecasting and Control*, San Francisco: Holden Day, 1970, 1976. Third edition (with G. Reinsel) published 1994 by Prentice Hall.

Brillinger, David R. (1965) An Introduction to Polyspectrum. *Annals of Mathematica. Statistics*. 36, 1351-1374.

Brockett, P. L., Hinich, Melvin, and Wilson, Gary R. (1987) Nonlinear and Non-Gaussian Ocean Noise. *J. Acoust. Soc. Am*. 82 no. 4, 1386-1394.

Brockett, P. L., Hinich, Melvin, and Patterson, Douglas. (1988) Bispectral-Based Tests for the Detection of Gaussianity and Linearity in Time Series," *Journal of the American Statistical Association*, 83 no. 403. 657-664.

Chen, Rong and Ruey S. Tsay. (1993a) Functional-Coefficient Autoregressive Models. *Journal of the American Statistical Association*, 88, no. 421, 298-308.

Chen, Rong and Ruey S. Tsay. (1993b) Nonlinear Additive ARX Models. *Journal of the American Statistical Association*, 88, no. 423, 955-967.

Friedman, Jerome. (1991) Multivariate Adaptive Regression Splines. *The Annals of Statistics* 19, no. 1, 1-67

Hinich, Melvin J. (1982) Testing For Gaussianity and Linearity of a Stationary Time Series. *Journal of Time Series Analysis* 3 no. 3, 169-176.

Hinich, Melvin J. (1996) Testing for Dependence in the Input to a Linear Time Series Model. *Journal of Nonparametric Statistics* 6, 205-221.

Hinich, Melvin J. and Patterson, Douglas. (1985) Evidence of Nonlinearity in Daily Stock Returns. *Journal of Business and Economic Statistics*, 3, no. 1, 69-77.

Jenkins, G. and Watts D. (1968) *Spectral Analysis and Its Applications*, San Francisco, Holden-

Day.

Lemos, Maria and Houston H. Stokes. (1998) A Single-Blind Controlled Competition Among Tests for Nonlinearity and Chaos; Further Results. Unpublished memo UIC Department of Economics.

Lewis, P. A. and J. G. Stevens. (1991) Nonlinear Modeling of Time Series Using Multivariate Adaptive Regression Splines (MARS). *Journal of the American Statistical Association*, 86, no. 416, 864-877.

Lewis, P. A. and B. K. Ray. (1997) Modeling Long-Range Dependence, Nonlinearity, and Periodic Phenomena in Sea Surface Temperature using TSMARS. *Journal of the American Statistical Association* 92, no. 439, 881-893.

Priestley, M. (1981) *Spectral Analysis and Time Series* (Vol. 2), New York: Academic Press.

Stokes, Houston H. and Melvin Hinich. "Testing the Gas Furnace Model for Nonlinear Serial Dependence." Department of Economics, University of Illinois at Chicago, 1989. mimeo.

Stokes, Houston H. (1991, 1997) *Specifying and Diagnostically Testing Econometric Models*. New York: Quorum Books.

Tiao, G. C. and Box, G. E. P. (1981) Modeling Multiple Time Series With Applications. *Journal of the American Statistical Association*, 76, 802-816.

Tiao, G. C., Box, G. E. P., Grupe, M. R., Hudak, C. B., Bell, W. R. and Chang, I. (1979) The Wisconsin Multiple Time Series (WMTS-1) Program: A Preliminary Guide. Department of Statistics, University of Wisconsin, Madison, memo.

Tong, H. (1983) *Threshold Models in Non-Linear Time Series Analysis.* New York: Springer-Verlag.

Tong, H. (1990) *Nonlinear Time Series.* New York: Oxford University Press.