Proceedings of the 2012 Industrial and Systems Engineering Research Conference
G. Lim and J.W. Herrmann, eds.

Long-Term Demand Prediction using Long-Run Equilibrium Relationship of Intrinsic Time-Scale Decomposition Components

Akkarapol Sa-ngasoongsong

School of Industrial Engineering and Management
Oklahoma State University
Stillwater, OK 74078, USA

Abstract

Long-term demand prediction is crucial for industries with long product development cycles. Especially in case of automobile products, where typical concept-to-release times are long (12-60 months), the process of demand evolution follows a nonlinear, nonstationary dynamics hindering accurate prediction of the futures of automobile demand. This paper presents a new prediction approach based on Intrinsic Time-Scale Decomposition (ITD) to forecast automobile demand over extended time-horizons. The ITD algorithm decomposes a time series signal into a sum of components called proper rotation components and monotonic trend. A key advantage of this approach is that this method can be used to identify long-run equilibrium relationship of automobile demand and ITD components of economic indicator. Once the relationship is identified, it takes advantage of the cointegration property of a Vector Error Correction Model (VECM) to forecast automobile demand. Hence, long-term impact of economic indicator components on automobile demand can be quantified. The empirical results suggest that this ITD-based prediction approach can significantly improve prediction accuracy in terms of RMSE (23%) and MAPE (26%) for long-term prediction of automobile demand (12-month ahead prediction), compared to classical and advanced time series techniques.

Keywords
Long-Term Demand Prediction, Intrinsic Time-Scale Decomposition, Vector Error Correction Model

1. Introduction
Demand prediction is essential part of any business activity. For industries with long product development cycles, long-term demand prediction serves as an input to many business decisions which usually affect profitability of the organization. As in the case of automobile products, where typical concept-to-release times are long (12-60 months), reliable long-term demand prediction makes an important contribution to successful service, revenue, production and inventory planning.

Automobile demand prediction has received significant attention in the literature. Many theoretical models for automobile demand prediction have been proposed [1-5]. Most of them are econometric approaches imposing a certain structure of economic theory on the data. Only recently, a few efforts have been made to address the automobile demand prediction problem using time series and data-driven approach [6-8]. However, none of them have addressed an automobile demand prediction for long-term horizon. In economic area, some recent developments on time series techniques have been specifically designed to quantify relationships among endogenous and exogenous variables. These techniques include Vector autoregressive (VAR) and Vector error correction Model (VECM) [9, 10]. Especially in the case of nonstationary variables, VECM has been broadly recognized as a powerful theory-driven model that can be used to describe the long-run dynamic behavior of multivariate time series. However, the choice of endogenous and exogenous variables in VECM is problematic for automobile demand prediction.

Demand is known to be influenced by many exogenous factors, such as advertising, sales promotions, retail price and technological sophistication [11]. In automobile market, advertising and sales promotions tend to have substantial effects, however, these effects on demand are rarely persistent [12]. To select variables for automobile demand prediction, some economic indicators, such as Consumer Price Index (CPI), unemployment rate, etc., have been suggested to have persistent effects on automobile demand, but the results of empirical model show that long-run equilibrium relationship among these economic indicators and automobile demand does not exist. Considering a nonlinear behavior of these economic indicators, one would expect that a statistical hypothesis test of long-run equilibrium using linear assumption may perform poorly due to model misspecification.

This paper presents a new prediction approach based on Intrinsic Time-Scale Decomposition (ITD) technique. The ITD method is a recently developed nonparametric decomposition technique for signals that are nonlinear and/or nonstationary in nature. ITD decomposes a signal into a sum of components called proper rotation components and monotonic trend. As advantages of nonparametric method, ITD requires very limited assumptions about the data. This advantage makes ITD suitable for nonlinear data from unknown underlying processes. ITD also provides unbiased decomposing components, compared to parametric algorithm due to parameter estimation. The key aspect of this new prediction approach is to develop a methodology to identify a causal and long-run equilibrium relationship from variables of interest and ITD components of related indicators, and use this identified relationship for prediction. This new prediction approach can be applied to any system of nonlinear and/or nonstationary variables. In this paper, the property of ITD leads to an application of determining long-run equilibrium relationship of automobile demand and ITD components of nonlinear economic indicator. Our investigation indicates that the long-run equilibrium of automobile demand and ITD components of unemployment rate does exist, and the ITD-based prediction can significantly improve prediction accuracy in terms of RMSE (23%) and MAPE (26%) for long-term prediction of automobile demand (12-month ahead prediction), compared to classical and advanced time series techniques. The organization of this paper is as follows: Section 2 present ITD algorithm in details. VECM and related statistical hypothesis tests are presented in Section 3. Section 4 is a methodology section of ITD-based prediction approach. The implementation details and empirical results are given in Section 5. Conclusion and suggested future work are presented in the last section of this paper.

2. Intrinsic Time-Scale Decomposition

Given a nonlinear time series signal X_t, where $t \in (1,2,\ldots,T)$, a general representation of a dynamical system of X_t can be represented using Volterra expansion [13] as

$$X_t = \delta_t + \sum_{i=0}^{\infty}\omega_i\eta_{t-i} + \sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\omega_{ij}\eta_{t-i}\eta_{t-j} + \sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{k=0}^{\infty}\omega_{ijk}\eta_{t-i}\eta_{t-j}\eta_{t-k} + \sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{k=0}^{\infty}\sum_{l=0}^{\infty}\omega_{ijkl}\eta_{t-i}\eta_{t-j}\eta_{t-k}\eta_{t-l} + \ldots \quad (1)$$

where δ_t is deterministic component, η_t ~ IID and $\{\omega_i, \omega_{ij}, \omega_{ijk}, \omega_{ijkl}, \ldots\}$ is a set of coefficients. Eq. (1) is nonlinear if it has nonzero coefficients $\{\omega_{ij}, \omega_{ijk}, \omega_{ijkl}, \ldots\}$ on the higher-order terms. The Volterra expansion in Eq. (1) can be transformed into an autoregressive representation [14] as

$$X_t = f(X_{t-1}, X_{t-2}, \ldots) + \eta_t \quad (2)$$

where $f(X_{t-1}, X_{t-2}, \ldots)$ is some nonlinear function of the past values of X. The difficulty in estimating Eq. (2) is that the functional form $f(\ )$ is unknown when working with real world data. There are number of procedures in the literature that have been developed to determine the form of nonlinearity. However, no set of tests can actually discern the correct form of nonlinearity. Rather, the tests can only suggest the form of the nonlinearity.

Intrinsic Time-Scale Decomposition (ITD) is a recently developed nonparametric algorithm aiming to enhance an analysis of nonlinear and nonstationary signals [15]. The advantage of ITD is that it does not require parametric functional form of the original series. The ITD method represents a nonlinear, nonstationary signal as a summation of components called proper rotation components and monotonic trend. ITD decomposes the signal into a sequence of proper rotations of successively decreasing instantaneous frequency at each subsequent level of the decomposition as

$$X_t = HX_t + LX_t = HX_t + (H+L)LX_t = (H(1+L)+L^2)X_t = (H\sum_{k=0}^{a-1}L^k + L^a)X_t \quad (3)$$

where H is defined as a proper-rotation-extracting operator and L is defined as baseline-extracting operator. $HL^k X_t$ is the $(k+1)^{st}$ level proper rotation and $L^a X_t$ is the monotonic trend. The following procedure of ITD is used to extract proper rotation components from a time series X_t:

(1) Identify the location of all local extrema of X_t denoted as τ_k where $k \in \{1,2,\ldots\}$

(2) Assume that X_t is available for $t \in [0, \tau_{k+2}]$, define a baseline signal L_t using a piece-wise linear baseline-extracting operator, L, between successive extrema as follows

$$LX_t = L_t = L_k + \left(\frac{L_{k+1} - L_k}{X_{k+1} - X_k}\right)(X_t - X_k), \qquad t \in (\tau_k, \tau_{k+1}] \quad (4)$$

where X_k and L_k denote $X(\tau_k)$ and $L(\tau_k)$ respectively, and

$$L_{k+1} = \alpha[X_k + (\frac{\tau_{k+1} - \tau_k}{\tau_{k+2} - \tau_k})(X_{k+2} - X_k)] + (1-\alpha)X_{k+1}, \qquad 0 < \alpha < 1 \tag{5}$$

(3) The first proper rotation can be extracted from X_t using a proper-rotation-extracting operator, H, as

$$HX_t \equiv (1-L)X_t = H_t = X_t - L_t \tag{6}$$

(4) The process (1) to (3) will perform iteratively, using the baseline signals as input, until the monotonic baseline signal is obtained. The following properties provide some key concepts of ITD:

Remark 1: The decomposition is nonlinear in the sense that the decomposition of two signals need not produce components equal to the sum of the components obtained from decomposing each signal individually, i.e. $S_t = X_t + W_t \;\Rightarrow HL^k S_t \neq HL^k X_t + HL^k W_t$

Remark 2: The edge effect of ITD process is confined to the interval $[0, \tau_2]$ at the beginning and $[\tau_k, \tau_{k+2}]$ at the ending of each proper rotation and monotonic trend.

3. Vector Error Correction Model

Vector Error Correction Model (VECM) [16] is a dynamical nonstationary multi-factor model with long-run relationship characteristics among variables called cointegrating relationships. Let $y_t = (y_{1t}, \ldots, y_{kt})'$, t = 1,2,… denote a k-dimensional time series vector of random variables of interest. The vector error correction model with the cointegration rank r (≤ k), denoted as VECM(p), can be written as

$$\Delta y_t = \delta + \prod y_{t-1} + \sum_{i=1}^{p-1} \Phi_i^* \Delta y_{t-1} + \varepsilon_t \tag{7}$$

where Δ is differencing operator, such that $\Delta y_t = y_t - y_{t-1}$; $\Pi = \alpha\beta'$, where α and β are $k \times r$ matrices; Φ_i^* is a $k \times k$ matrix. The cointegrating vector, β, is sometimes called the long-run parameters, and α is adjustment coefficient. In the case of cointegration with exogenous variables, the VECM with exogenous variables, VECMX(p, s), can be written as follows:

$$\Delta y_t = \delta + \prod y_{t-1} + \sum_{i=1}^{p-1} \Phi_i^* \Delta y_{t-1} + \sum_{i=0}^{s} \Theta_i^* z_{t-i} + \varepsilon_t \tag{8}$$

Theoretically, VECM is applicable if and only if y_t is cointegrating, as defined in the following definition:

Definition 1: The components of the vector $y_t = (y_{1t}, \ldots, y_{kt})'$ are said to be cointegrated of order d, b, denoted by y_t ~ CI(d, b) if (I) All components of y_t are integrated of order d, $I(d)$, and (II) There exists a vector $\beta = (\beta_1, \beta_2, \ldots, \beta_n)$ such that the linear combination $\beta y_t = \beta_1 y_{1t} + \beta_2 y_{2t} + \ldots + \beta_n y_{nt}$ is integrated of order (d - b), I(d - b), where b > 0.

In order to identify the cointegrating vector from a system of variables, y_t, using Definition 1, some statistical hypothesis tests are required. The following subsections provide the details of statistical hypothesis tests involving in cointegrating vector identification.

3.1 Unit Root Test

By definition, cointegration necessitates that all variables be integrated of the same order. The Augmented Dickey-Fuller (ADF) test [17, 18] is selected to pretest each variable to determine its order of integration. There are three main versions of the test equation as shown in Eq. (10)-(12).

$$\Delta y_t = \gamma y_{t-1} + \delta_1 \Delta y_{t-1} + \ldots + \delta_{p-1} \Delta y_{t-p+1} + \varepsilon_t \tag{9}$$

$$\Delta y_t = \alpha + \gamma y_{t-1} + \delta_1 \Delta y_{t-1} + \ldots + \delta_{p-1} \Delta y_{t-p+1} + \varepsilon_t \tag{10}$$

$$\Delta y_t = \alpha + \gamma y_{t-1} + \beta_t + \delta_1 \Delta y_{t-1} + \ldots + \delta_{p-1} \Delta y_{t-p+1} + \varepsilon_t \tag{11}$$

The null hypothesis is $\gamma = 0$ against the alternative hypothesis of $\gamma < 0$. Eq. (9) is the test equation of a unit root with zero mean. Eq. (10) is the test equation of a unit root with drift and Eq. (11) is the test equation of a unit root test with drift and deterministic time trend. The appropriate lag length (p) can be selected using the general-to-specific methodology [19] or an information criterion such as the Akaike Information Criterion (AIC) or Bayesian Information Criterion (BIC). Dickey and Fuller [20] have shown that the test statistic of Dickey-Fuller test does not follow a t-distribution, so that the usual percentile of t-distribution cannot be used as comparison numbers for the calculated t. In their Monte Carlo study, they found that the critical values for $\gamma = 0$ depend on the form of the regression and on sample size. The statistics called τ, τ_μ, and τ_τ are the appropriate statistics to use for Eq. (9), (10), and (11), respectively.

3.2 Weak Exogeneity Test

In a cointegrated system, a variable is weakly exogenous if no useful information is lost when this variable is conditioned without specifying its generating process. In the other words, if a variable does not respond to the discrepancy from the long-run equilibrium relationship, it is weakly exogenous. The weak exogeneity test is used to identify the weak exogeneity effect of each variable to the others. To test which variables should be treated as endogenous, and which ones as exogenous in the equation, the k-vector of $I(d)$ random variables y_t is initially partitioned into the k_1-vector y_{1t} and the k_2-vector y_{2t}, where $y_t = (y_{1t}^{'}, y_{2t}^{'})$ and $k = k_1 + k_2$. From the VECM model in Eq. (7), the parameters can similarly be decomposed as $\delta = (\delta_1^{'}, \delta_2^{'})$, $\alpha = (\alpha_1^{'}, \alpha_2^{'})$, $\Phi_i^* = (\Phi_{1i}^*, \Phi_{2i}^*)$, and the variance-covariance matrix as

$$\Sigma = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix} \tag{12}$$

The conditional model for y_{1t} given y_{2t} is

$$\Delta y_{1t} = \omega \Delta y_{2t} + (\alpha_1 - \omega \alpha_2)\beta^{'} y_{t-1} + \sum_{i=1}^{p-1} (\Phi_{1i}^* - \omega \Phi_{2i}^*)\Delta y_{t-i} + (\delta_1 - \omega \delta_2) + (\varepsilon_{1t} - \omega \varepsilon_{2t}) \tag{13}$$

and the marginal model for y_{2t} is

$$\Delta y_{2t} = \alpha_2 \beta^{'} y_{t-1} + \sum_{i=1}^{p-1} \Phi_{2i}^* \Delta y_{t-i} + \delta_2 + \varepsilon_{2t} \tag{14}$$

where $\omega = \Sigma_{12}\Sigma_{22}^{'}$

The null hypothesis of the test of weak exogeneity is $\alpha_2 = 0$, and the weak exogeneity test statistic follows χ^2 distribution.

3.3 Cointegration Test

For a test of cointegration, the Johansen's reduced rank methodology [21, 22] is employed. Let $y_t = (y_{1t}, \ldots, y_{kt})'$, $t = 1,2,\ldots,k$ denote a k-dimensional time series vector. As in ADF test, Δy_t can be written as

$$\Delta y_t = \pi y_{t-1} + \sum_{i=1}^{p-1} \pi_i \Delta y_{t-i} + \varepsilon_t \tag{15}$$

The key of the rank methodology is rank of matrix π. The rank of π is equal to the number of independent cointegrating vectors which can be obtained by checking the significance of the characteristic root of π. The test for the number of characteristic roots can be conducted using two test statistics, the trace and maximum eigenvalue statistics:

$$\lambda_{trace}(r) = -T \sum_{i=r+1}^{n} \ln(1 - \lambda_i) \tag{16}$$

$$\lambda_{\max}(r, r+1) = -T \ln(1 - \lambda_{r+1}) \tag{17}$$

where λ_i is the estimated values of the characteristic roots (also called eigenvalues) obtained from the estimated π matrix, and T is the number of usable observations. The null hypothesis is $\lambda_i = 0$ for $i = r+1, \ldots, k$. The asymptotic distribution of the trace and maximum eigenvalue statistic is given by $tr(A)$ and $max(A)$ where A is the matrix defined as

$$A = \int_0^1 (dW)\tilde{W}^{'} \left(\int_0^1 \tilde{W}\tilde{W}^{'} dr \right)^{-1} \int_0^1 \tilde{W}(dW)^{'} \tag{18}$$

where $tr(A)$ is the trace of matrix A and $max(A)$ is the maximum eigenvalue of a matrix A. W is the $k - r$ dimensional Brownian motion, and $\tilde{W}$ is the detrended Brownian motion. Tables of the critical values of these statistics can be found in [23].

4. ITD-based Prediction Methodology

As presented in the introduction section, the ITD-based prediction method takes an advantage of cointegration property of VECM by identifying long-run equilibrium relationship of automobile demand and ITD component of economic indicator. The ITD-based prediction methodology is consisting of the following three steps. The first step of the prediction methodology is an ITD step. In this step, economic indicator will be decomposed into a finite number of proper rotation components and monotonic trend using ITD method. The second step of methodology is variable selection step. This step is used to select endogenous and exogenous variables for VECM. In the second step, nonstationary condition and integration order for each component will be identified using unit root test. Since automobile demand is nonstationary, stationary ITD components will be treated as exogenous variables. Other exogenous variables will be identified using weakly exogenous test. Long-run equilibrium relationship of endogenous variables, including automobile demand and selected ITD components, will be identified using cointegration test. The third step of prediction methodology is a prediction step. The VECM model of selected endogenous and exogenous variables in the second step will be parameterized. Out-of-sample forecasting will be made for automobile demand. The comparison of out-of-sample forecasting with other classical and advanced time series techniques will be done in this step. The overall framework of ITD-based prediction methodology is shown in Figure 1.

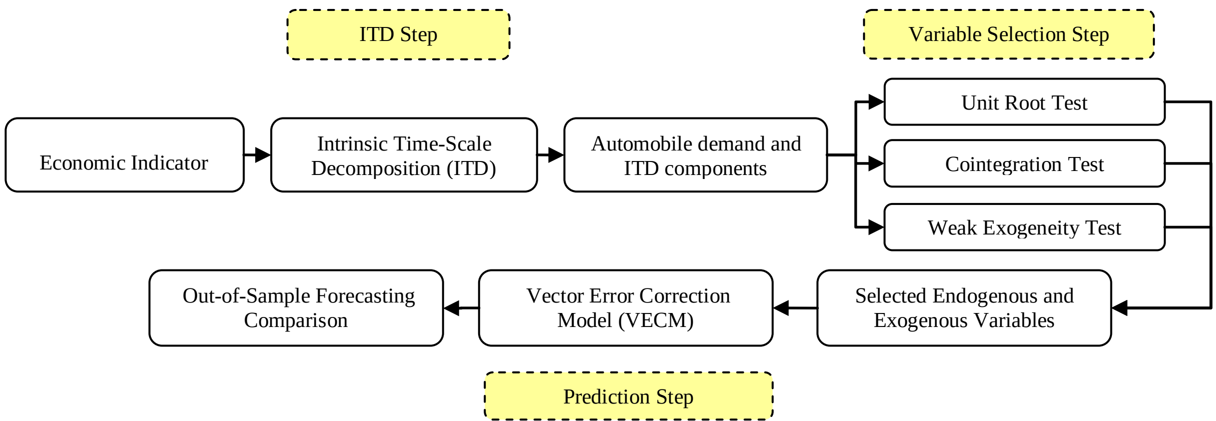
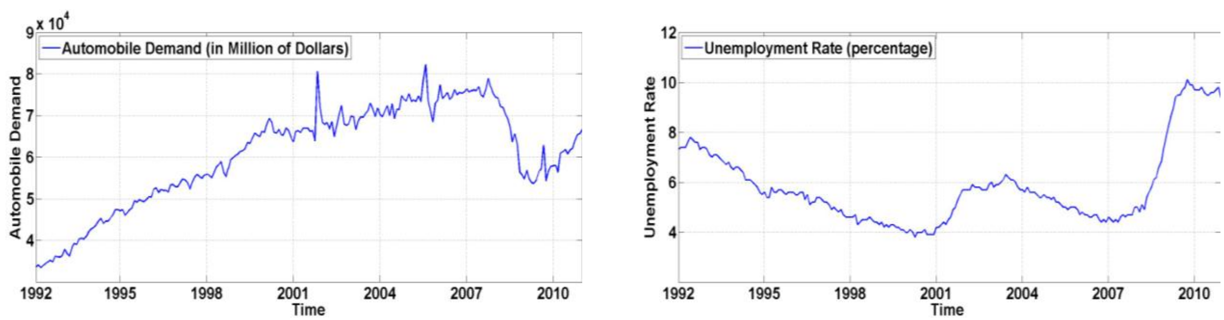


Figure 1: ITD-based Prediction Methodology Framework

5. Implementation Details and Empirical Results

In this paper, the number of monthly retail sales (Motor vehicle and parts dealers) in U.S. during a period of January 1992- December 2010 is used to represent an aggregate automobile demand. An unemployment rate during the same period of time is selected as economic indicator to help predict demand. In order to predict demand for long term, ITD components of unemployment rate are hypothesized to have long-run equilibrium relationship with automobile demand. This paper will investigate and test hypotheses of long-run relationship of automobile demand and ITD components of unemployment rate. Table 1 provides the details of demand and unemployment rate. For simplicity, logarithm transformation will be used with both variables in subsequent analysis. The time series in original scale of both variables are shown in Figure 2.



(a) (b)

Figure 2: (a) Automobile Demand and (b) Unemployment rate

From literature [24], many research papers have indicated an existence of nonlinear behavior in the U.S. unemployment rate. To present the effect of nonlinear behavior in the unemployment rate on the long-run equilibrium relationship with automobile demand, the result of cointegration test between unemployment rate and automobile demand is shown in Table 2.

Table 1: Summary of variables

Variables	Source	Publishing Interval	Description
Automobile Demand	BLS*	Monthly	A monthly retail sales of motor vehicles and part accessories
Unemployment rate	BLS*	Monthly	A national unemployment rate (16 years or over)

* Bureau of Labor Statistics (Seasonally adjusted data)

Table 2: Cointegration Rank Tests (Automobile Demand and Unemployment Rate)

		Trace Test			Maximum Eigenvalue Test		
H_0: Rank = r	H_1: Rank > r	Trace Statistic	5%	10%	Max Statistic	5%	10%
0	0	8.6416	15.34	13.31	7.8798	14.07	12.07

Table 2 reports the test statistics and the corresponding asymptotic critical values of both Trace and Maximum Eigenvalue tests at the 5% and 10% significance levels. The empirical results of cointegration tests show that the null hypothesis of no cointegrating vector cannot be rejected for both tests at 5% and 10% significance levels. The test results indicate that there is statistically no long-run equilibrium relationship between unemployment rate and automobile demand. Considering cointegrating vector requirement, VECM cannot be directly applied to these variables. As discussed in the introduction section, the cointegration tests using linear assumption may perform poorly in the case of nonlinear variables. The next step in this section is to apply the ITD-based prediction methodology presented in Section 4 to investigate and test the long-run equilibrium relationship between automobile demand and ITD components of unemployment rate. The implementation details are as follows:

5.1 Intrinsic Time-Scale Decomposition Step

The first step of prediction methodology is to decompose unemployment rate using ITD method described in Section 2. The ITD algorithm decomposes unemployment rate into four proper rotation components and monotonic trend as shown in Figure 3. The procedure of ITD algorithm allows a perfect reconstruction of original unemployment rate as presented in Eq. (3).

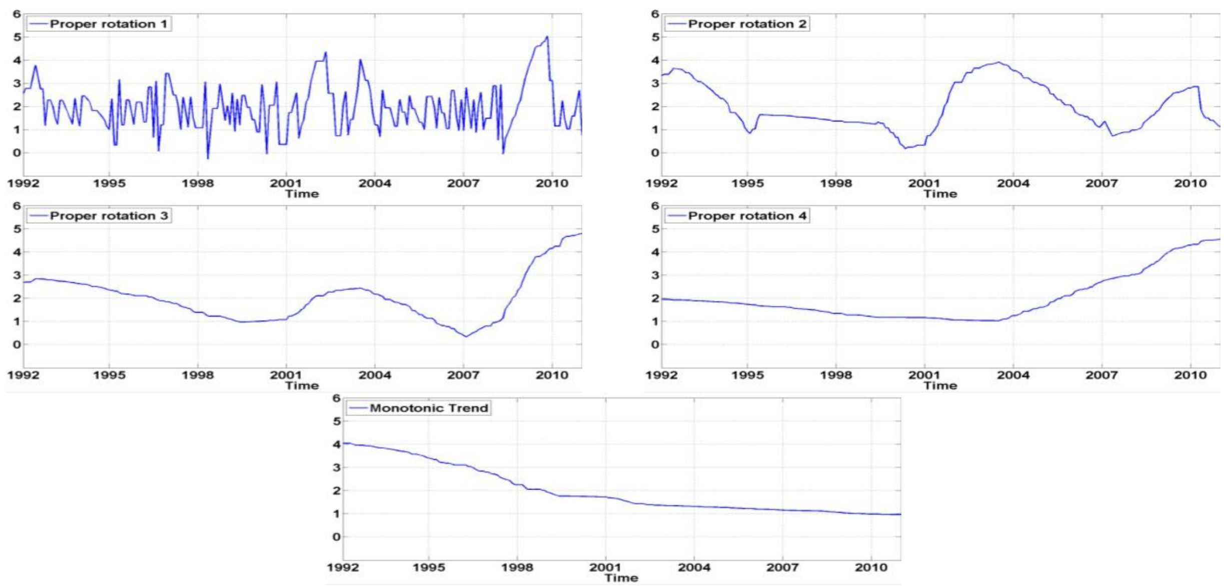


Figure 3: ITD components of Unemployment Rate

5.2 Variable Selection Step

In this second step, endogenous and exogenous variables for VECM are selected from automobile demand and ITD components of unemployment rate obtained from the first step. Stationarity characteristic of automobile demand and each ITD components of unemployment rate is identified using unit root tests. Exogenous variables are selected, based on results from weak exogeneity tests. In order to conduct the Augmented Dickey-Fuller unit root test, it is important to use the correct number of lags in the equation. The appropriate lag length is selected, based on the general-to-specific methodology. The procedure begins with estimating an autoregressive model with relatively long lag length. The idea is to pare down the model by the usual *t*-test and/or *F*-tests. If the *t*-statistic on the last lag (p^*) is insignificant at some specified significance level, reestimate the model with a lag length of p^*-1, then repeat the process until the last lag is significantly different from zero. The ADF test results with optimal lag length selected using a general-to-specific methodology are shown in Table 3.

Table 3: ADF Unit Root Test of Original and First Differenced Variables

Variables	Optimal Lag Length	Unit Root with Drift		Unit Root with Drift and Deterministic Time Trend	
Original Variables	p^*	τ_μ	Pr < τ_μ	τ_τ	Pr < τ_τ
Automobile Demand (AD)	3	-3.21	0.0213	-1.70	0.75
Proper rotation 1 (H1)	4	-4.83	0.0001	-4.90	0.0004
Proper rotation 2 (H2)	4	-2.71	0.0751	-2.69	0.2405
Proper rotation 3 (H3)	5	-2.51	0.1146	-2.33	0.4130
Proper rotation 4 (H4)	5	0.14	0.9683	-0.66	0.9739
Monotonic Trend (R)	5	-2.62	0.0905	-0.68	0.9725
First Differenced Variables	p^*	τ_μ	Pr < τ_μ	τ_τ	Pr < τ_τ
ΔAD	6	-4.74	0.0002	-5.19	0.0002
ΔH2	3	-4.34	0.0006	-4.34	0.0033
ΔH3	3	-2.99	0.0377	-3.45	0.0475
ΔH4	3	-2.97	0.0395	-4.36	0.0030
ΔR	4	-3.20	0.0216	-4.14	0.0065

From the ADF unit root test results, only proper rotation 2 (H2) is statistically stationary at 1% significance level. The null hypothesis of unit root cannot be rejected for automobile demand and all other ITD components of unemployment rate at 1% significance level. Applying first difference to all variables, except H2, the ADF test statistics reject the null hypothesis of unit root for all differenced variables at 5% significance level. Based on the results of ADF tests, all variables, excluding H2, are treated as nonstationary I(1) series. To select exogenous variables for VECM, the results of weak exogeneity tests on I(1) series as shown in Table 4.

Table 4: Weak Exogeneity Tests (Selected Variables)

Variables	Degree of Freedom	$\chi^2(4)$	Pr > χ^2
AD	4	22.74	0.0001
H2	4	24.25	<.0001
H3	4	12.43	0.0144
H4	4	23.95	<0.0001
R	4	6.58	0.1595

The weak exogeneity test begins with the vector of five variables (AD, H2, H3, H4 and R) in Table 3. For a system of five variables, there can exist at most four cointegrating vector. The results in Table 4 show that the null hypothesis of weak exogeneity cannot be rejected for monotonic trend (R).

Table 5: Weak Exogeneity Tests (Re-estimate)

Variables	Degree of Freedom	$\chi^2(4)$	Pr > χ^2
AD	3	16.95	0.0007
H2	3	12.01	0.0074
H3	3	8.53	0.0363
H4	3	19.82	0.0002

Hall et al. [25] suggest re-estimating the model, using only rejecting weakly exogenous variables as endogenous variables, to avoid sensitivity on the model specification. Table 5 shows the results of re-estimating weak exogeneity tests of four endogenous variables, excluding monotonic trend. The test statistics continue to reject the null hypothesis of weak exogeneity at 5% significance level for all four variables. Based on results of unit root and weak exogeneity tests, automobile demand (AD), and three ITD components of unemployment rate (H2, H3 and H4) are selected as endogenous variables. For exogenous variables, due to possible colinearity issue between automobile demand (AD) and monotonic trend (R), only proper rotation 1 (H1) is selected as exogenous variable.

To investigate an existence of long-run equilibrium relationship among endogenous variables, Table 6 reports the test statistics and corresponding asymptotic critical values at 5% and 10% significance level of cointegration rank tests. The test results show that the null hypothesis of one cointegrating vector cannot be rejected at both 5% and 10% significance level. There is statistically long-run equilibrium relationship or cointegrating vector among four endogenous variables (AD, H2, H3 and H4) as shown in Eq. (19).

$$\beta \cdot AD = \beta_{H2} \cdot H2 + \beta_{H3} \cdot H3 + \beta_{H4} \cdot H4 \quad (19)$$

The underlying process of these variables are random in the short term but tend to move together in the long term horizon. Since all four variables are I(1) series, the cointegrating vector in Eq. (19) is stationary process ($I(d\text{-}b) = I(0)$ where $d = b = 1$). Long-Run parameter (β) estimates of the cointegrating vector are presented in Table 7, where the parameter estimate of automobile demand is normalized ($\beta=1$).

Table 6: Cointegration Rank Tests (Automobile Demand and selected ITD components of unemployment rate)

		Trace Test			Maximum Eigenvalue Test		
H_0: Rank = r	H_1: Rank > r	Trace Statistic	5%	10%	Max Statistic	5%	10%
0	0	52.5057	47.21	43.84	29.5151	27.07	24.73
1	1	22.9906	29.38	26.70	14.4999	20.97	18.60

Table 7: Long-Run Parameter Beta Estimates of Cointegrating Vector (AD, H2, H3 and H4)

Variable	AD	H2	H3	H4
Beta	β	B_{H2}	B_{H3}	B_{H4}
Parameter Estimate	1	0.41389	-0.46206	0.68588

5.3 Prediction Step

From results of steps 1 and 2, it shows that VECM is applicable for an automobile demand and ITD components of unemployment rate since cointegrating vector of these variables can be identified. The ITD-based prediction using VECMX with 4 endogenous (AD, H2, H3 and H4) and 1 exogenous (H1) variable is estimated using Eq. (9). To evaluate the forecasting performance of the ITD-based prediction using VECMX, the out-of-sample forecasting of this model is compared with those from three rival models. The first model is an autoregressive integrated moving average (ARIMA) model. The ARIMA model is the most general class of univariate time series models which can be applied to nonstationary variable. The second model is an autoregressive integrated moving average with exogenous variables (ARIMAX) model. The ARIMAX model of automobile demand is selected as a baseline model since it allows to use unemployment rate as exogenous variable in the model, however it cannot handle feedback from one variable to the other. The third model is vector autoregressive (VAR) model. The VAR model can handle feedback from one variable to the other, but does not impose cointegrating restrictions as in the VECMX. It treats automobile demand and unemployment rate symmetrically. Two model selection criterions are selected for out-of-sample comparison. They are root mean square error (RMSE) and mean absolute percentage error (MAPE). The out-of-sample forecasting initially identified each model specification over the sample period of January 1992 to May 2007, then each model is used to generate forecasts of four-, eight- and twelve-step ahead predictions. The sample is then rolled forward for one month, and another set of four- to twelve-step ahead predictions is generated. As a result, 24 four-, eight- and twelve-step ahead predictions are obtained. Table 8 provides RMSE and MAPE values of four-, eight- and twelve-step ahead predictions of automobile demand (AD) for each model. Figure 4(a) and 4(b) show the out-of-sample forecasting comparison of four models using RMSE and MAPE.

Table 8: Out-of-Sample Forecasting Comparison

Model	4-step ahead prediction		8-step ahead prediction		12-step ahead prediction	
	RMSE	MAPE	RMSE	MAPE	RMSE	MAPE
ARIMA	0.0846	0.0071	0.1562	0.0132	0.2182	0.0190
ARIMAX	0.0859	0.0072	0.1585	0.0134	0.2206	0.0192
VAR	0.0772	0.0064	0.1440	0.0121	0.1999	0.0173
ITD-based Prediction	0.0578	0.0045	0.1100	0.0091	0.1531	0.0128

Comparing four models in Table 8, ITD-based prediction model is the best model in terms of RMSE and MAPE. It can significantly improve a prediction accuracy of automobile demand for all four-, eight-, and twelve-step ahead predictions. For 4-step ahead prediction, ITD-based prediction model can reduce RMSE and MAPE by 25% and 30% respectively, compared to the second best model (VAR model). For 8-step ahead prediction, ITD-based prediction model can reduce RMSE and MAPE by 24% and 25% respectively, compared to VAR model. In the case of 12-step ahead prediction, ITD can reduce RMSE and MAPE by 23% and 26% respectively, compared to VAR model.

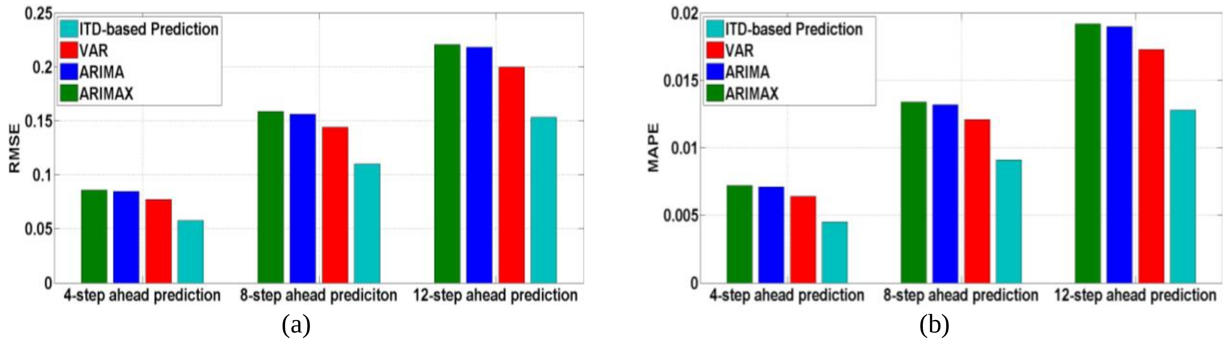


Figure 4: (a) Out-of-Sample Forecasting Comparison using RMSE and (b) Out-of-Sample Forecasting Comparison using MAPE

6. Conclusion and Suggested Future Work

In this paper, a new prediction approach based on Intrinsic Time-Scale Decomposition (ITD) is proposed for long-term prediction application of nonlinear and/or nonstationary variables. ITD is a recently developed nonparametric algorithm aiming to enhance an analysis of nonlinear and nonstationary signals. The ITD method represents a nonlinear and/or nonstationary signal as a summation of components called proper rotation and monotonic trend. The advantage of ITD-based prediction approach is that, from system of variables with no long-run equilibrium relationship or cointegrating vector, this approach can be used to identify cointegrating vector of one variable and ITD components of other variables. In this study, the ITD-based prediction approach is applied to a long-term prediction of automobile demand with the aid of economic indicator. Considering economic indicator as a combination of multiple components with different characteristics resulting from factors underlying each component, the property of ITD leads to a prediction application using a long-run equilibrium relationship or cointegrating vector of automobile demand and ITD components of economic indicator. The empirical results show that automobile demand has a long-run equilibrium relationship with some of ITD components of unemployment rate, and ITD-based prediction approach can significantly improve a prediction accuracy for long-term prediction of automobile demand.

To improve a prediction accuracy using ITD algorithm further, one may consider the edge effect of ITD as discussed in Section 2 of this paper as suggested future work. The edge effect of ITD on the most recent data is confined to the interval of the last two extremas. This is because the most recent data point of the signal is treated as an extremum in which it may not necessarily be true when more data become available. In a prediction application, the edge effect of ITD has a significant effect on a prediction accuracy. One of possible alternatives to solve this issue is to construct the decomposition beyond the last available data point. However, this extension of the data is a risky procedure because the edge effect will depend on an accuracy of the extended data. Currently, no simple approach or general theory exists to solve the edge effect issue of ITD.

However, with property of ITD that a proper rotation is monotonic between two consecutive local extremas, ITD has an advantage to help in the analysis. That is the whole data span needed not to be extended. The only things needed are the value and location of the next two extremas. Considering the advantage of ITD, this task of improving a prediction accuracy by solving the edge effect issue is still challenging.

Acknowledgements

This research is supported by National Science Foundation (CMMI 0729552 and CMMI 0830023).

References

1. Berkovec, J., *Forecasting automobile demand using disaggregate choice models.* Transportation Research Part B: Methodological, 1985. **19**(4): p. 315-329.
2. Franses, P.H., *Modeling New Product Sales; An Applicaiton of Cointegration Analysis.* International Journal of Research in Marketing, 1994. **11**: p. 491-502.
3. Greenspan, A. and D. Cohen, *Motor Vehicle Stocks, Scrappage, and Sales.* The Review of Economics and Statistics, 1999. **81**(3): p. 369-383.
4. Lave, C.A. and K. Train, *A disaggregate model of auto-type choice.* Transportation Research Part A: General, 1979. **13**(1): p. 1-9.
5. Mannering, F.L. and K. Train, *Recent directions in automobile demand modeling.* Transportation Research Part B: Methodological, 1985. **19**(4): p. 265-274.
6. Brühl, B., et al., *A Sales Forecast Model for the German Automobile Market Based on Time Series Analysis and Data Mining Methods*, in *Advances in Data Mining. Applications and Theoretical Aspects*, P. Perner, Editor. 2009, Springer Berlin / Heidelberg. p. 146-160.
7. Shahabuddin, S., *Forecasting Automobile Sales.* Management Research News, 2009. **32**(7): p. 670-672.
8. Wang, F.-K., K.-K. Chang, and C.-W. Tzeng, *Using adaptive network-based fuzzy inference system to forecast automobile sales.* Expert Systems with Applications, 2011. **38**(8): p. 10587-10593.
9. Dekimpe, M.G., et al., *Time-Series Models in Marketing*, in *Handbook of Marketing Decision Models*, B. Wierenga, Editor. 2008, Springer US. p. 373-398.
10. Dekimpe, M.G. and D.M. Hanssens, *The Persistence of Marketing Effects on Sales.* Marketing Science, 1995. **14**(1): p. 1-21.
11. Landwehr, J.R., A.A. Labroo, and A. Herrmann, *Gut Liking for the Ordinary: Incorporating Design Fluency Improves Automobile Sales Forecasts.* Marketing Science, 2011. **30**(3): p. 416-429.
12. Dekimpe, M.G., D.M. Hanssens, and J.M. Silva-Risso, *Long-run effects of price promotions in scanner markets.* Journal of Econometrics, 1998. **89**(1-2): p. 269-291.
13. Schetzen, M., *The Volterra and Wiener Theories of Nonlinear Systems*. 1980: {John Wiley & Sons}.
14. Gerald P. Dwyer, J., *Nonlinear Time Series and Financial Applications*. 2003, Federal Reserve Bank of Atlanta.
15. Frei, M.G. and I. Osorio, *Intrinsic time-scale decomposition: time-frequency-energy analysis and real-time filtering of non-stationary signals.* Proceedings of the Royal Society A-Mathematical Physical and Enigneering Sciences, 2007. **463**(2078): p. 321-342.
16. Engle, R.F. and B.S. Yoo, *Forecasting and testing in co-integrated systems.* Journal of Econometrics, 1987. **35**(1): p. 143-159.
17. Dickey, D.A. and W.A. Fuller, *Likelihood Ratio Statistics for Autoregressive Time Series with a Unit Root.* Econometrica, 1981. **49**(4): p. 1057-1072.
18. Said, S.E. and D.A. Dickey, *Testing for unit roots in autoregressive-moving average models of unknown order.* Biometrika, 1984. **71**(3): p. 599-607.
19. Enders, W., *Applied Economietric Time Series*. 2010: John Wiley & Sons.
20. Dickey, D.A. and W.A. Fuller, *Distribution of the Estimators for Autoregressive Time Series With a Unit Root.* Journal of the American Statistical Association, 1979. **74**(366): p. 427-431.
21. Johansen, S., *Statistical analysis of cointegration vectors.* Journal of Economic Dynamics and Control, 1988. **12**(2-3): p. 231-254.
22. Johansen, S., *A Stastistical Analysis of Cointegration for I(2) Variables.* Econometric Theory, 1995. **11**(01): p. 25-59.
23. Osterwald-Lenum, M., *A Note with Quantiles of the Asymptotic Distribution of the Maximum Likelihood Cointegration Rank Test Statistics1.* Oxford Bulletin of Economics and Statistics, 1992. **54**(3): p. 461-472.
24. Rothman, P., *Forecasting Asymmetric Unemployment Rates.* The Review of Economics and Statistics, 1998. **80**(1): p. 164-168.
25. Greenslade, J.V., S.G. Hall, and S.G.B. Henry, *On the identification of cointegrated systems in small samples: a modelling strategy with an application to UK wages and prices.* Journal of Economic Dynamics and Control, 2002. **26**(9-10): p. 1517-1537.