A General Framework For Forecasting Futures Markets Prices: Non-Parametric Regression Analysis*

Morteza Haghiri
Dept. of Agricultural Economics
University of Saskatchewan, Canada

ABSTRACT

A particular nonparametric regression analysis approach is used to forecast the canola futures contract prices in Canada. We used smoothing spline in a generalized additive model to estimate the parameters of the forecasting function. Diagnostic checking tools indicated that the time series underlying forecast followed at most an AR (11) process. The results showed that the canola futures contract prices, on average, would decrease by 1.4 percent of the origin futures contract prices in November 2000, but individually for the last three forecasting points, they will increase by 0.2 percent in compare to the previous ones.

* Paper presented for the 4th International conference in Applied Economics and Econometrics at Economic Research Centre/ Middle East Technical University (METU), Ankara, Turkey, September 13- 16, 2000.

A General Framework For Forecasting Futures Markets Prices: Non-Parametric Regression Analysis

MORTEZA HAGHIRI*

JEL Classification System: G13, G14

Keywords: Futures Markets, Generalized Additive Models, Smoothing Spline.

Abstract- A particular nonparametric regression analysis approach is used to forecast the canola futures contract prices in Canada. We used smoothing spline in a generalized additive model to estimate the parameters of the forecasting function. Diagnostic checking tools indicated that the time series underlying forecast followed at most an AR (11) process. The results showed that the canola futures contract prices, on average, would decrease by 1.4 percent of the origin futures contract prices in November 2000, but individually for the last three forecasting points, they will increase by 0.2 percent in compare to the previous ones.

Introduction

An old definition about market says: "*a market is a place or situation that puts sellers and buyers in communication with one another where prices are discovered, and where ownership may be transferred.*" This global definition applies to wide variety of markets from local market as well as to futures markets. *Futures markets* exist to assist in the process of price discovery and provide a mechanism that allows producers, holders, and users of a commodity to transfer and reduce price risk. The transfer is accomplished by taking equal and opposite positions in the futures and cash markets. There is also no question that futures markets move quickly, although its motives in the markets and the accuracy of the prices it generates, is always being debated.

Price analysis is the key to successful use of any futures or options market. Whether a trader is speculating or hedging, some expectation of prices is required for development of a trading strategy. These price expectations can be derived in a number

---

* Morteza Haghiri is a Ph.D. candidate in the department of Agricultural Economics, University of Saskatchewan, 51 Campus Drive, Saskatoon SK. S7N 5A8, Canada. Tel: (306) 966-4024, Fax: (306) 966-8413, E-mail: Morteza.Haghiri@sask.usask.ca

of ways. As date, two general categories of price analysis have been used in futures and options markets. First, *fundamental analysis* that is a way of looking forward, gathering data concerning supply and demand conditions expected in the future. Second, *technical analysis,* which is a way of looking backward, making forecasts based on historical data.

This study tries to introduce a general framework for forecasting futures market prices using a *non-parametric* analysis. Non-parametric analysis, unlike the parametric one, has the advantage of not having *a known priori* functional form in advance. In this type of analysis we let *data speak for them* to have some particular form of distribution. There are five main sections in this paper. Section one introduces basic characteristics of futures markets. In section two, both methods of futures market analyses are discussed in detail. Section three introduces the principle of forecasting. In section four, non-parametric regression analysis is introduced by using an empirical analysis of the canola futures contract prices. Section five concludes some remarkable points.

1. Futures Markets

Markets are places where assets are traded. This definition does not necessarily mean that in order to have a market, first, we must have a physical place. An asset is something having or perceived to have value which may be of a financial or nonfinancial nature. Financial assets represent financial claims, which may be direct or indirect. These types of assets or instruments trade in financial markets include capital markets, money markets, and currency markets which are not the case in this paper. Nonfinancial assets are assets of a more physical nature, such as physical commodities, real estate, and collectibles (Marshall 1989). Physical commodities can be further divided into agricultural, industrial material, and precious metal groupings. Agricultural futures commodity markets are of interest of this paper in the empirical section.

Futures markets have captured the public's attention since mid of 1970s, because world economic boom have caused growth activities and expanded markets that led to attract a wide variety of new traders to those markets. That boom created more volatile in prices and interest rates which results in greater uncertainty in decision making. Many markets are now much more international in nature than just few decades ago. Large capital flows seem to know no boundaries, responding to alternative investment opportunities around the world. Agricultural contracts helped lead futures trading to large volumes beginning in the 1970s due to rising commodity and energy prices. Adverse weather around the world, changing governments policies in agriculture, growing international trade, formation of regional and international agreement in trade such as the G8 countries, NAFTA, EU, and WTO dramatically altered traditional markets.

Futures markets perform several functions. Primarily, they (1) facilitate the management of risk; (2) aid firms in discovering forward prices; (3) provide a means for firms to request and secure additional operating capital; and (4) provide a source of information for decision making. Futures markets are effective, because of their close approximation to the classical concepts of pure competition. We will extend our discussion in this matter when we explain the principles of market analysis in next section.

The existence of futures markets has created a flow of information far exceeding what would otherwise be the case. This allows for much more intelligent decision making. Futures markets are a place where information is interpreted and exchanged, and the information they provide for decision making may well be their most significant economic role for both participants and nonparticipants. Futures markets can make cash markets more efficient through its information role, and as a matter of fact, their purpose is to make cash markets work better. Due to this relationship with cash markets, futures markets are primary price discovery markets for many commodities, instruments, and indices. Futures prices are base reference values for producers, processors, and consumers of many goods and transference of ownership of goods and services, and increase the information content of spot market prices. Futures markets attract speculators who search for information and whose trading should improve rational price information. Speculation is not costless because information is costly, but ultimately these activities can reduce total marketing costs for everyone.

Firms and business face risks as they buy inputs, perform a service on them, and then sell products. During this process price relationships may change so much that the firm cannot cover all its costs. In order to be successful, such firms need skills in managing risks. Futures markets is one of the firms' choices for managing risks. Effective risk management has financial consequences. Firms that can reduce risks should have a higher probability of making a profit. Such firms should be able to command additional operating capital. Lenders would be more willing to lend money to those enterprises that can show lower probabilities of financial loss and bankruptcy while they can manage these interest rate risks directly on the futures market.

Futures markets are a zero-sum game. It means that for every currency unit we obtain from the market, some one else must lose a unit of currency. Futures markets have the image of being fast-moving markets. It is not that their prices necessarily fluctuate more than other prices, but that futures trading is done on the margin. Small amounts of money command large contracts. The benefits of a futures market come from improving the scope and flexibility of decision making for a firm. Futures trading is a highly standardized way to enter into fixed-price forward contracts; it is economical in execution and thereby secure in performance. Ownership responsibilities are shifted to someone else, and firms can organize production operations more efficiently.

Futures markets have more functions and characteristics than we discussed in this section, which for the lack of the space we couldn't explain all of them. For more details in futures markets functions and characteristics see Leuthold, Junkus and Cordier (1989).

2. The Principles of Price Analysis

Price analysis is the key to successful use of any futures or options market. Whether a trader is speculating or hedging, some expectation of prices is required for development of a trading strategy. These price expectations can be derived in a number of ways. As date, two general categories of price analysis have been used in futures and options markets. First, fundamental analysis that is a way of looking forward, gathering data concerning supply and demand conditions expected in the future. Secondly, technical analysis which is a way of looking backward, making forecasts based on historical data. Fundamental analysis is usually considered to be more difficult than technical analysis because there is no so much information to evaluate. It also requires knowledge of a specific market, whereas the same technical rules can be applied to a number of different markets. In general, a fundamentalist tries to determine the intrinsic value of a product, which is a long-run approach to price analysis. In next subsection we briefly explain each of those methods followed by comparing them to give readers a rough idea about the principles of price analysis. Interested readers may find more detail explanations in Blank, Carter and Schmiesing (1991).

2.1 Fundamental Analysis Approach (FAA)

There are many traders who concentrate their efforts in futures and options markets based on the fundamental factors of supply and demand because these markets are thought to be examples of near perfect competition. That idea mainly comes from the environment that traders are performing under some circumstances that fulfill the requirements of competition:

- Large number of participants so no one has much influence on market prices.
- Product homogeneity is closely approached by the contract specifications.
- Although free entry and exit cannot be happened however they are closely approximated through small contract sizes and margin requirements.
- Full information is available which is one of the advantages of futures and options markets.
- No intervention by the third decision-maker exists and traders have their independence and impersonality in their operations which are dictated by the organization and regulation of the markets.

Fundamental traders believe that in such a market, the key point to price analysis is to interpret of information concerning factors affecting supply or demand as to their meaning for the long-term trend in prices. They are not as concerned about when prices will move up and down significantly as they are with the probability of whether prices will move in a given direction and the possible extent of such a move. For example, if the direction of price movements can be predicted correctly, speculators can make profits and hedgers can better identify what the best strategy for them is. One thing that is important in the viewpoint of the fundamental traders is the randomness of the prices. Therefore, they try to monitor and correctly evaluate new long-term information concerning general economic factors and specific market supply and demand variables.

Conventionally, a fundamental approach to price analysis begins when an analyst determines the general business climate and checks the overall trend of commodity or financial prices. This survey is similar to the situation when we want to analyze the macroeconomic factors aimed at identifying the trend in the nation's economy. The second stage of the fundamental approach involves when we want to determine the answers to a number of questions for a particular product. By answering to those questions, which are accounted as the microeconomic aspects of our analysis, we obtain detail information of both qualitative and quantitative analysis of data concerning specific variables.(For more details on such type of questions see Gold 1975).

2.2 Technical Analysis Approach (TAA)

Technical analysis is the study of historical prices, usually in chart form, in order to predict future price behavior. Traditional technical analysts, when forecasting prices, reject the use of statistical supply and demand data or interpretations of new events. They study the market by itself rather than external supply and demand factors. The development of microcomputers has eased the study of quantitative historical price analysis through technical approach. Technical analysts, some times called chartists, believe that supply and demand data are subject to interpretation and are not usually current. The vast volume of interaction in futures and options markets almost makes impossible to analyze the market through the fundamental approach. Moreover, in the long-term the forecasts obtained from such approach are too vulnerable to new information, hence the focus on recent price.

Technical analysts believe that these problems with fundamental analysis can be avoided if we *"let the market tells us"* when to trade long or short. They state that by merely watching the direction and amplitude of price movements, we are able to tell how others are thinking and trading. They usually criticize fundamental analysts for ignoring the major, and unpredictable, price influence: the mood of the trading public. Technical analysts keep a record of price movements in hopes of finding patterns that indicate that a recognizable price movement is underway or that the trend is turning.

Technical analysis provides traders with empirical, rather theoretical, guidance in generating price forecast, and this advantage have made appeal to traders and market analysts to use this approach in their analyses. Although technical analysts discuss buying and selling pressures, however, there are significant economic assumptions required to justify use of technical analysis. This approach assumes that history repeats itself, which means the future can be predicted from the past. In other words, applying this method inherently contains the assumption of no *random walk* exists in day-to-day price changes. We will explain the theory of random walk in the following sections. This, in turn, implies that the markets are inefficient in viewpoint of economic definition. Fama (1976) defined an efficient market is one that accurately and rapidly reflects available supply and demand information.

There are numerous technical analysis methods used by futures and options traders. These methods may fall into four broad categories:

- Patterns on price charts
- Trend-following methods
- Character-of-market analysis
- Structural theories

2.3 Fundamental versus Technical Analysis Approach

Futures brokers and managers of futures funds are those agents who support technical analysis. In comparing fundamental and technical analysis, technical analysts' argument center on four reasons for using their method over the alternative:

1. It is almost impossible for any private trader to compete with large commercial companies in the rapid collection of accurate and costly information. In contrast, with technical analysis, a private trader can often determine whether the major commercial companies' interests are trading short or long.
2. Fundamental analysis may succeed in predicting long-term price moves, but it is not too helpful in the short term. Since there are high level of financial leverage in futures trading, private traders cannot afford to be wrong in the short term.
3. Preoccupation with daily news events and rumors makes it very difficult to maintain the systematic and disciplined approach to trading decisions required for successful trading.
4. Technical analysis approach has been working for long time. For decades, people have made money employing primarily technical methods. It is simple to use and the same techniques can be applied to many different markets.

The first three points all focus on the need of a trader to identify quickly what other market participants are thinking and doing. The fourth reason is the crucial one. People may ask whether or not technical analysis is a profitable exercise. Several studies have addressed this issue. They have tried to make an objective assessment of the performance of technical trading methods. For example, Irwin and Brorsen (1987) found that inflation rates affect returns to futures fund, implying that there are structural factors creating trends in prices which provide opportunities for systems trading profits. They also concluded that the amount of technical trading in a market does not affect returns to fund. Concerning specific technical trading methods, Neftci and Policano (1984) found that moving average systems seem to have some predictive power, while trend-following systems have mixed results. To explore more differences between fundamental and technical analysis approach see Blank, Carter and Schmiesing (1991).

3. The Principles of Forecasting

There is a general consensus that forecasting is important because forecasts are constantly made in business, economics, government, finance, and many other fields. Obviously, forecasts are made to guide decisions in a variety of fields. Operations planning and control, marketing, economics, capacity planning, business and government budgeting, demography, crisis management, financial risk management, and financial speculation are some applications indicates that how important forecasts are. For example, speculators in asset markets have an interest in forecasting asset returns such as stock returns, interest rates, exchange rates, commodity prices, etc and such forecasts are made routinely. The forecasting of asset return volatility is related to the forecasting of asset returns. For decades, practical methods for volatility forecasting have been developed and widely applied. Volatility forecasts are crucial for evaluating and insuring risks associated with asset portfolios. Volatility forecasts are also crucial for firms and inventors who need to price assets such as options and other derivatives.

With so many different forecasting applications, we may think that numerous variety of forecasting techniques exist, however a relatively small number of tools form the common core of almost all forecasting methods. Although predicting the population of country *A* is different with anticipating the stock market of *B*, the principles underlying the forecasts are identical. There are six basic points relevant to all forecasting tasks that we should pay attention in order to have successful forecasting. Diebold (1998) had done an excellent and comprehensive review of those points, which are discussed in this section very briefly:

(a) *Decision Environment*: The key to generating good and useful forecast is recognizing that forecasts are made to guide decisions, so recognition and awareness of the decision-making environment is the important point to effective design, use, and evaluation of forecasting models.

(b) *Nature of Forecast Object*: Based on the problem we have, there are many objects to forecast. Typically, the forecast object in business and economics contains one of these three choices: *event outcome*, *event timing*, and *time series*. Event outcome forecasts are relevant to situations in which an event is certain to take place at a given time but the outcome of that event is uncertain. Event timing forecasts are relevant when an event outcome is certain to take place and the outcome is known, but the timing is uncertain. Time series forecasts involve projecting the future value of a time series of interest. Among those three, the last one has been paid attention by most economists, and also we use in this paper. There are at least two reasons why time series forecasts are by far the most frequently encountered in practice. First, most business, economic, and financial data are in time series format. Second, the technology for making and evaluating time series forecasts is sufficiently well developed.

(c) *Forecast Statement*: The way a forecast is stated depends entirely upon the forecast type. When we make a forecast, we must know if the forecast will be (a) a single number (*point forecast)*; (b) a range of numbers into which the future value can be expected to fall a certain percentage of the time (*interval forecast*); or (c) an entire probability distribution for the future value (d*ensity forecast*). A good point forecast provides a simple guide to the future of a time series, but it is highly sensitive to the random and unpredictable shocks. As a result of such shock, we may expect some nonzero forecast error even that point forecast is a very good projection of the future value. Thus, we may need to know the degree of confidence available in a particular point forecast. The uncertainty surrounding point forecasts suggests the usefulness of an interval forecast. An interval forecast is a range of values in which we expect the realized value of the series to fall with some prespecified probability. Two points are worthy here to be mentioned: first, the length (size) of the intervals conveys information regarding forecast uncertainty. Second, interval forecasts convey more information than point forecasts. A density forecast gives the entire density or probability distribution of the future value of the series of interest. As with interval forecasts, density forecasts convey more information than point forecasts. They also contain more information than interval forecasts, because given a density; interval forecasts at any desired confidence level are readily constructed.

(d) *Forecast Horizon*: The forecast horizon is defined as the number of periods between the current time and the date of the forecast we make. At least two things make the horizon so important: first, it is clear that the forecast changes if the horizon alters. Second, the best forecasting model will often change with the forecasting horizon as well. In general, we have an *h-step-ahead forecast*, for which the horizon *h* is at the discretion of the user. The choice of horizon depends on the decision that the forecast will guide. The nature of the decision environment typically dictates whether short-term, medium-term or long-term forecasts are needed.

(e) *Information Set*: Certainly, the quality of our forecasts is limited by the quality and quantity of information available when forecasts are made. The idea of information set is fundamental to constructing good forecasts. Sometimes in time series forecasting, we use the historical of one series variable, called *univariate information set.* Alternatively, sometimes we use the *multivariate information set* which contains al least two historical series of variables related somehow to each other. No matter what type of information set we have, it is always important to seriously consider what information is available, what additional information could be collected or made available, the form of the information such as qualitative or quantitative, and so on.

(f) *Choice of Forecasting Method*: Conventionally, the specifics of the situation will indicate the desirability of a specific method or modeling strategy. Although the complexity of phenomena in the real-world may cause us to think of using complicated tools and models for forecasting, fortunately this is not the case because decades of professional experience suggest using simple and parsimonious models tend to be best models for out-of-sample forecasting business, economics, and finance. Few reasons may convince us why smaller and simpler models are often more attractive than larger and more complicated ones. First, we can estimate the parameters of simpler models more precisely. Second, because simpler models are more easily interpreted, understood and scrutinized, anomalous behavior is easily spotted. Third, it is easier to communicate an intuitive feel for the behavior of simple models, which makes them more useful in the decision-making process. Finally, enforcing simplicity lessens the scope for data mining, but we should notice that using a simple model is quite different with using a naive one. In fact, as Diebold (1998) stated “it should be well formalized in the *KISS principle*: *Keep It Sophisticatedly Simple.*”

4. Nonparametric Regression Analysis

A fundamental problem in econometrics is to develop models based on a sample of observations so that further analyses can be carried out with statistical techniques applied in econometrics using the model so developed. During the past several years, econometrists used parametric models widely. A disadvantage of such models is that it may not be robust in the sense that slight contamination of the data by observations not following the particular parametric family might lead to erroneous conclusions. Furthermore, the data might be of such a type that there is no suitable parametric family that gives a good fit. Under these circumstances, we might take recourse to nonparametric modeling.

In general, there are three basic problems in statistics and other close majors such as econometrics: (1) devising scientific and valid statistical techniques for data collection; (2) developing statistical methods for statistical inference based on collection data; and (3) improving the statistical models for future used based on these experiences. The classical statistical methods are parametric. The assumption underlying every econometric model is that the sample of observations comes from a population with a known parametric family. The problem is either to estimate the unknown parameters or to devise tests and derive confidence region for the unknown parameters based on the sample. This assumption is rather strong because the assumed parametric model need not be the *actual* one if there is one, and the statistical methods developed for a particular parametric model could lead to erroneous conclusions when applied to a slightly perturbed model. These problems brought about a recent trend to develop nonparametric statistical methods for data analysis.

The fundamental problem is the estimation of the distribution function or the density function, if it exists, of the population from which the sample has been obtained. For a long time, statisticians and other researchers used histograms, which are a type of density estimators, as tools in drawing inferences form data. Recently a large class of nonparametric methods has been developed for the estimation of distribution functions, density functions, etc., for data of several types. One of these methods is the *Generalized Additive Model (GAM)* introduced in the next section.

4.1 Generalized Additive Models (GAM)

Generalized additive models proposed by Hastie and Tibshirani (1984) are the extension of a*dditive models*, which in turn, are the extended form of a family functions of standard linear models, called generalized linear models. In these models, the mean of the response is modeled as an additive sum of the predictors. Generalized linear models are themselves a generalization of linear regression models. Specifically, the predictor's effects are assumed to be in linear in the parameters, but the distribution of the responses, as well as the *link* between the predictors and this distribution, can be quite general. Many useful models fall into the class, for example, in a linear logistic model for binary data, the response Y is assumed to have a Bernoulli distribution with

$$\mu = \mathrm{pr}\left(Y|X_1,..., X_p\right)$$

And μ is linked to the predictors through $\log\{\mu/1-\mu\} = \alpha + \sum_j X_j\beta_j$. Other familiar models in this comprehensive class are log-linear models for categorical data and gamma regression models for responses with constant coefficient of variation. The family of generalized linear models provides an appealing framework for studying the common structure of such models. In short, generalized additive models extend generalized linear models in the same manner as that additive models extend linear ones. It means generalized additive models replace the linear form $\alpha + \sum_j X_j\beta_j$ with the additive form $\eta(X) = \alpha + \sum_j f_j\left(X_j\right)$ The name *additive* refers to the multivariate assumption underlying the model, namely that the p-predictor function η has a low-dimensional additive structure. Such models are attractive if they fit the data, since they are far easier to interpret than a p-dimensional multivariate surface.

In this paper we intend to use a flexible method for modeling an individual term in an additive model. The functions are fitted using *scatterplot smoothers*, nonparametric techniques for fitting a regression function in a flexible data-defined manner. Several smooth terms are fitted simultaneously in an additive model by using the scatterplot smoothers iteratively. However, this flexible method cannot be applied with any cost. The cost involved in such method is that for nonparametric techniques to be successful and remain parsimonious, the underlying functions need to be reasonably smooth.

The additive predictor η can be used in all different situations where the linear predictor is used for the generalized linear models. A *univariate smoother*, a *semiparametric model*, or a *nonparametric logistic function* are some examples that we can classify them under the family of generalized additive models. There are many approaches that we can estimate any nonparametric regression coefficients. One approach is based on *penalized likelihood* and *smoothing splines*. Others are based on the idea of *local estimation*. In this paper, we used the spline smoothing approach to estimate the coefficients of the forecasting function.

4.2 Empirical Analysis

In order to test nonparametric regression analysis in forecasting agricultural commodity futures price, we utilized the database obtained from technical tools, which is available in the World Wide Web at http://www.techtool.com/. The source of data is the *Winnipeg Commodity Exchange (WCE)* and that is in time series (daily) starts from 1981 to 2000. The WCE is a public exchange in which demand and supply, reflected through buyers and sellers respectively, contribute to price discovery through open auction trading. It is the centralized nature of the auction and the characteristics of the futures contracts which make the institution unique compared with other public markets for agricultural commodities in Canada.

Among all the commodity information available in the WCE database, we chose canola as of our interest. For every daily canola futures contract (i.e. a type of forward contract), the data contains date, open, high, low, and settle (close) prices, volume, and open interest. The price of contract is determined when buyers and sellers, or their agents, agree on a price on the floor of the Exchange in a specific transaction. The buyer and seller of the contract publicly agree on a price and date for subsequent delivery of a fixed amount of a certain grade of a designated commodity, within a given geographical area. That futures contract commits the buyer and seller to make equal and opposite legally binding commitments. For example, the seller of one September canola contract is agreeing to deliver 20 tonnes of No.1 canola to a particular port no later than the last business day of September. On the other side of the transaction, the buyer of the canola futures contract agrees to accept delivery in September and, upon delivery, to complete payment for the contracted account. We may notice that the buyers and sellers are not actually dealing in canola when they agree on a canola futures price. In contrast, they are agreeing on a price which completes the terms of the contract for future delivery if that occurs.

The contract delivery month is different for each commodity. For canola, the contract delivery months are January, March, May, July, August, September and

November. Any forecast futures contract prices for canola must be done for only these seven contract delivery months.

4.2.1 Model Framework

We followed the Box-Jenkins (1970) methodology, which abandoned the econometric modeling approach of using explanatory variables (predictors) suggested by economic theory to forecast, choosing instead to rely on the past behavior of the variable being forecast (Kennedy 1998). Basically, Box-Jenkins method begins by transforming the response variable from non-stationary to become stationary, namely that its stochastic properties are invariant with respect to time (i.e., that the mean of response variable, its variance, and its covariance with other response values associated with lags, do not depend on time). That transformation is necessary because in the presence of non-stationary, all the conventional hypothesis-testing procedure based on t, F, chi-square tests, and the like may be suspect (Gujarati 1995).

Figure (1) shows the sample correlogram of futures contract price of canola in the level case obtained from EViews version 3.1, that is shown up to 36 lags. The number of lags came from the fact that in practice, lags up to one-third of the sample size are generally used. This figure indicates that our sample is not a stationary time series one, because it starts at a very high value (0.944 at lag 1) and tapers off very gradually. There are some evidences to confirm this pledge: First, we see two different patterns in the estimated autocorrelation coefficients through the sample size. The first pattern, from the first lag up to the eighth one, all the estimated autocorrelation coefficients are individually statistically different from zero at the 95% level of significant. A similar pattern is exactly seen from the twenty-eighth lag up to the end one. However, the estimated autocorrelation coefficients between these two categories are not individually statistically different from zero. Second, the Box-Pierce Q statistics, (Q=495.75), tells us we should reject the null hypothesis that all estimated autocorrelation are zero; at least some of them must be nonzero. Finally, the Augmented Dickey-Fuller (ADF) unit root test, which is shown in table (1) indicates that we cannot reject the null hypothesis that the sample data is stationary. In other words, the response in time series being studied is a random walk.

Figure(1): Correlogram of Futures Contract Prices(level)

Canola
Sample: 1 112
Included observations: 112

Lag	AC	PAC	Q-Stat	Prob
1	0.944	0.944	102.42	0.000
2	0.863	-0.248	188.90	0.000
3	0.764	-0.181	257.24	0.000
4	0.671	0.068	310.48	0.000
5	0.573	-0.122	349.66	0.000
6	0.474	-0.079	376.77	0.000
7	0.363	-0.171	392.77	0.000
8	0.248	-0.086	400.29	0.000
9	0.144	0.063	402.85	0.000
10	0.066	0.121	403.39	0.000
11	0.010	0.041	403.41	0.000
12	-0.019	0.123	403.45	0.000
13	-0.042	-0.049	403.68	0.000
14	-0.065	-0.105	404.23	0.000
15	-0.086	-0.025	405.19	0.000
16	-0.091	0.043	406.30	0.000
17	-0.083	0.014	407.23	0.000
18	-0.062	0.042	407.75	0.000
19	-0.043	-0.054	408.00	0.000
20	-0.035	-0.068	408.18	0.000
21	-0.047	-0.099	408.48	0.000
22	-0.071	-0.119	409.20	0.000
23	-0.091	0.057	410.40	0.000
24	-0.111	-0.040	412.17	0.000
25	-0.128	-0.018	414.59	0.000
26	-0.155	-0.028	418.14	0.000
27	-0.187	-0.001	423.41	0.000
28	-0.235	-0.140	431.80	0.000
29	-0.274	0.036	443.37	0.000
30	-0.300	0.025	457.35	0.000
31	-0.311	-0.032	472.59	0.000
32	-0.317	-0.025	488.63	0.000
33	-0.320	-0.021	505.22	0.000
34	-0.326	-0.017	522.61	0.000
35	-0.331	-0.054	540.72	0.000
36	-0.332	-0.110	559.30	0.000

Table (1): Augmented Dickey-Fuller Unit Root Test (Model 1)

ADF Test Statistic	-2.048455	1% Critical Value*	-3.4906
		5% Critical Value	-2.8877
		10% Critical Value	-2.5805

*MacKinnon critical values for rejection of hypothesis of a unit root.

Augmented Dickey-Fuller Test Equation
Dependent Variable: D(Canola Future Contract, Level Price)
Method: Least Squares
Sample(adjusted): 3 112
Included observations: 110 after adjusting endpoints

Variable	Coefficient	Std. Error	t-Statistic	Prob.
P(-1)	-0.061187	0.029870	-2.048455	0.0430
D(P(-1))	0.279520	0.093381	2.993338	0.0034
C	21.08215	10.52917	2.002262	0.0478

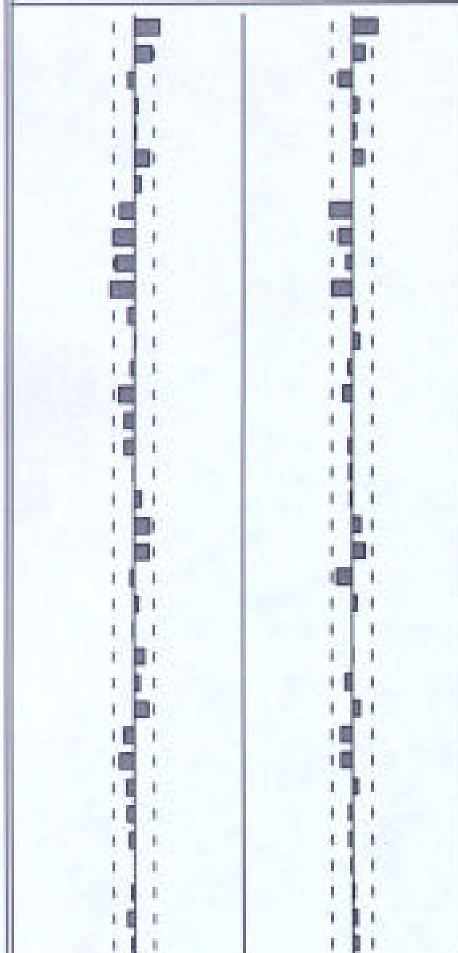
R-squared	0.095951	Mean dependent var	-0.397198
Adjusted R-squared	0.079053	S.D. dependent var	14.98146
S.E. of regression	14.37710	Akaike info criterion	8.196045
Sum squared resid	22117.02	Schwarz criterion	8.269694
Log likelihood	-447.7825	F-statistic	5.678214
Durbin-Watson stat	2.082669	Prob(F-statistic)	0.004532

With these evidences, we have no choice but to transform the response variable in order to remove the unit root problem and change it to stationary. To do this, we took the first differences of the response variable, and again obtain the correlogram of the new data series. Figure (2) shows the sample correlogram of canola futures contract prices where the prices were taken first differences. In order to facilitate in comparison, the sample correlogram is computed in the same number of lag points as Figure (1). In Figure (2), we notice that all of the estimated correlated coefficients, except the first, eighth, and eleventh lags are not individually statistically different from zero at the 95 % level of significant. This means the sample that is being studied does not have unit root problem any more, however further investigation is required. To make sure the error terms in the new data set are purely random, we again utilized the augmented Dickey-Fuller test which the results are shown in Table (2).

Figure(2):Correlogram of Futures Contract Prices

Canola (First-Differenced)
Sample: 1 112
Included observations: 111

Autocorrelation | Partial Correlation



	AC	PAC	Q-Stat	Prob
1	0.246	0.246	6.8920	0.009
2	0.172	0.118	10.282	0.006
3	-0.054	-0.129	10.615	0.014
4	0.035	0.059	10.758	0.029
5	0.023	0.035	10.820	0.055
6	0.139	0.115	13.144	0.041
7	0.063	0.004	13.617	0.058
8	-0.137	-0.211	15.899	0.044
9	-0.201	-0.130	20.889	0.013
10	-0.174	-0.054	24.656	0.006
11	-0.229	-0.189	31.242	0.001
12	-0.050	0.035	31.562	0.002
13	0.010	0.068	31.576	0.003
14	-0.029	-0.036	31.685	0.004
15	-0.146	-0.086	34.472	0.003
16	-0.089	-0.007	35.511	0.003
17	-0.083	-0.027	36.431	0.004
18	-0.005	-0.020	36.435	0.006
19	0.069	-0.009	37.093	0.008
20	0.152	0.090	40.264	0.005
21	0.145	0.131	43.213	0.003
22	-0.027	-0.143	43.318	0.004
23	0.045	0.042	43.604	0.006
24	-0.004	0.007	43.607	0.008
25	0.100	0.009	45.060	0.008
26	0.052	-0.065	45.460	0.010
27	0.134	0.079	48.159	0.007
28	-0.093	-0.106	49.463	0.007
29	-0.136	-0.099	52.309	0.005
30	-0.068	0.053	53.021	0.006
31	-0.063	-0.025	53.651	0.007
32	-0.046	-0.026	53.991	0.009
33	0.005	-0.014	53.995	0.012
34	-0.019	0.023	54.053	0.016
35	-0.065	0.043	54.754	0.018
36	-0.017	0.051	54.801	0.023

Table (2): Augmented Dickey-Fuller Unit Root Test (Model2)

ADF Test Statistic	-5.618282	1% Critical Value*	-3.4911
		5% Critical Value	-2.8879
		10% Critical Value	-2.5807

*MacKinnon critical values for rejection of hypothesis of a unit root.

Augmented Dickey-Fuller Test Equation
Dependent Variable: D(Canola Future Contract, First-Difference Price)
Method: Least Squares
Sample(adjusted): 4 112
Included observations: 109 after adjusting endpoints

Variable	Coefficient	Std. Error	t-Statistic	Prob.
D(P(-1))	-0.663903	0.118168	-5.618282	0.0000
D(P(-1),2)	-0.118150	0.096397	-1.225669	0.2230
C	-0.357158	1.397841	-0.255507	0.7988

R-squared	0.385994	Mean dependent var	-0.129804
Adjusted R-squared	0.374409	S.D. dependent var	18.43994
S.E. of regression	14.58495	Akaike info criterion	8.224994
Sum squared resid	22548.39	Schwarz criterion	8.299068
Log likelihood	-445.2622	F-statistic	33.31836
Durbin-Watson stat	1.970474	Prob(F-statistic)	0.000000

As it is seen in Table (2), the estimated ADF statistical value in absolute term, (5.62), is greater than the critical values at 1, 5, and 10 percent level of significant, which means we are now able to state with confidence that the response data set is become stationary. Therefore, we can proceed our goal to estimate canola futures contract price for the delivery months. One thing is worthy to point out here and that is the estimated ADF statistical value is also greater than the MacKinnon critical value test. Since the time series is being studied becomes stationary after taking first difference, we can say that the original (random walk) series is integrated of order 1, I (1). It is obvious that the first difference of response variable is an I (0) process.

4.2.2 Identification Selection

Identification selection is one of the steps of the Box-Jenkins methodology. The chief tools in identification are the *autocorrelation function (ACF)*, the *partial autocorrelation function (PAFC)*, and the resulting correlograms. As we pointed out in the previous section, in Figure (2), all estimated partial autocorrelation coefficients were statistically non different from zero, except the first, eighth, and eleventh lags. If the partial correlation coefficient were significant only at lag 1, we could have identified this as an AR(1) model, but this is not the case. Instead, we can conclude that the process that generated the first-differenced of response variable is at the most an AR (11) process.

In sum, the partial autocorrelations with spikes at 1,8, and 11 seem statistically significant, which lead to put them into the estimation process, i.e. the second step of Box-Jenkins (1970) methodology.

4.2.3 Process Estimation Using Nonparametric Regression

In the introduction of section five, and in the subsection of 5.1, we discussed the nonparametric regression functions and particularly how to estimate the parameters of a model using one of the techniques introduced. In this section we utilized the generalized additive model (GAM) with spline smoothing approach in order to estimate the parameters of the model.

Let P_t denote the first differences of canola futures contract prices. Then our tentatively identified AR model is

$$P_t = \delta + \alpha_1 P_{t-1} + \alpha_8 P_{t-8} + \alpha_{11} P_{t-11} \qquad (1)$$

Using S-Plus version 4.5 by writing proper spline smoothing interface codes in Unix Operation System (UOS), we obtained the following estimates:

$$P_t = 0.1706 + 0.0483\ s(P_{t-1}) + 0.024\ s(P_{t-8}) + 0.0734\ s(P_{t-11}) \qquad (2)$$

Deviance Residuals:

Min	*1Q*	*Median*	*3Q*	*Max*
-90.87091	-7.910196	0.82399	8.399606	110.5906

Dispersion Parameter for Gaussian family taken to be 820.1527
Null Deviance: 78394.72 on 99 degrees of freedom
Residual Deviance: 71351.17 on 86.99742 degrees of freedom
Number of Local Scoring Iterations: 1

Table (3) shows the degrees of freedom for terms and F-values for nonparametric estimated coefficients using general additive model approach by smoothing spline.

Table (3): DF for Terms and F-values for Nonparametric Effects

Terms	DF	Npar. DF	Npar. F	Pr(F)
Intercept	1			
P^*_{t-1}	1	3	0.8234	0.4844
P^*_{t-8}	1	3	1.2930	0.2820
P^*_{t-11}	1	3	0.9520	0.4193

Before we are going to do our last task to forecast canola futures contract prices for the specific delivery months, we must examine our ARIMA model, in general and AR model in our case, using diagnostic tests. One simple diagnostic is to get the residuals from equation (2) and obtain the autocorrelation and partial autocorrelation function of these residuals for appropriate number of lags, which in our case it is 36. For the lack of space, we do not report the results of the diagnostic check, but none of the autocorrelations and partial autocorrelations was individually significant. Nor was the sum of the 36 squared autocorrelations, as shown by the Box-Pierce Q statistics, statistically significant. With these results, we concluded that the correlograms of both autocorrelation and partial autocorrelation gave the impression that the residuals estimated from equation (2) were purely random. Therefore, it is not necessary to look for another ARIMA model.

4.2.4 Forecasting

We reach the point of forecasting canola futures contract prices after using diagnostic tools for checking the model. To do this, we emphasize again that the forecasts of canola futures contract must be done in certain time, that is the seven futures contract months. Following the settle price in November 2000, i.e., the last data available for canola futures contract, we have forecasted the futures contract prices for the next 12 contract months starting January 2001 till August 2002. We should notice that in equation (2), the response variable is *change* in the futures contract price over the previous period of time. Therefore, to obtain the forecast of the price level rather than its changes, we must *undo* the first-difference transformation that we had used to obtain the changes. In other words, we *integrated* the first-differenced series. Table (4) shows the results of forecasting. They indicate that the canola futures contract prices will have a swing movement around the futures contract prices in November 2000. Although those prices reduce to 1.4 percent of the ones in November 2000, they increase for the last three forecasting points by 0.2 percent in compare to the previous ones.

Table (4): Forecasts of Canola Futures Contract Prices ($CDN/Tonne)

2001		2002	
January	295.372	January	291.726
March	293.797	March	291.497
May	291.725	May	292.184
July	292.502	July	292.011
August	291.836	August	292.123
September	291.099		
November	291.726		

5. Conclusions

Futures markets have captured the public's attention since mid of 1970s, because world economic boom have caused growth activities and expanded markets that led to attract a wide variety of new traders to those markets. In Canada, canola futures contract prices are traded in exchange markets in certain period of time. In this paper we utilized the Winnipeg Commodity Exchange (WCE) database to forecast the canola future contract prices using a general framework of regression, nonparametric regression analysis. There are many nonparametric approaches to estimate the parameters of a function through general additive model (GAM). We may refer to local estimation, penalized likelihood, and smoothing spline as some examples of nonparametric approaches.

Price analysis is the key to successful use of any futures or options market. As date, two general categories of price analysis have been used in futures and options markets. First, fundamental analysis that is a way of looking forward, gathering data concerning supply and demand conditions expected in the future. Secondly, technical analysis which is a way of looking backward, making forecasts based on historical data

There are six basic points relevant to all forecasting tasks that one should pay attention in order to have successful forecasting. Those are decision environment, nature of forecast object, forecast statement, forecast horizon, information set, and choice of forecasting method.

Followed the Box-Jenkins methodology, we forecasted the canola futures contract prices. Since we had I (1) process in the response variable, taking the first-differenced made it change to I (0). To recognize the type of time series, we used autocorrelation function, partial autocorrelation function and the correlograms. The results indicated that the model followed at most an AR (11) process. By estimating the parameters using smoothing splines, we forecasted the canola futures contract prices for the 12 contract months. The results showed that the canola futures contract prices, on average, would decrease by 1.4 percent of the origin futures contract prices in November 2000, but individually for the last three forecasting points, they will increase by 0.2 percent in compare to the previous ones.

References

Blank, S. C., Carter, C. A., and B. H. Schmiesing (1991), "Futures and Options Markets: Trading in Financial and Commodity", USA: Prentice-Hall Inc.

Box, G. P. E., and G. M. Jenkins (1970), "Time Series Analysis: Forecasting and Control", San Francisco, USA: Holden Day (revised edition 1976).

Diebold, F. X. (1998), "Elements of Forecasting", USA: South-western Publishing Co.

Fama, E. F. (1976), "Foundations of Finance", New York: Basic Books.

Gold, J. (1975), "Modern Commodity Futures Trading", New York: Commodity Research Bureau.

Gujarati, D. N. (1995), "Basic Econometrics", USA: McGraw-Hill Inc.

Hastie, T. J., and R. J. Tibshirani (1990), "Generalized Additive Models", UK: Champman and Hall.

Irwin, S., and W. Brorsen (1987), "A Note on the Factors Affecting Technical Trading System Returns", Journal of Futures Markets, Vol.7, pp.591- 595.

Kennedy, P. (1998), "A Guide to Econometrics", Cambridge, Massachusetts, USA: The MIT Press.

Leuthold, R. M., Junkus, J. C., and J. E. Cordier (1989), "The Theory and Practice of Futures Markets", Massachusetts, USA: Lexington Books.

Marshall, J. F. (1989), "Futures and Option Contracting: Theory and Practice", Cincinnati, USA: South-western Publishing Co.

Neftci, S., and A. Policano (1984), "Can Chartists Outperform the Market? Market Efficiency Tests for 'Technical Analysis'", Journal of Futures Markets, Vol.4, pp.465- 478.