The Australian Journal of
Agricultural and
Resource Economics

Journal of the Australian Agricultural and Resource Economics Society

The Australian Journal of Agricultural and Resource Economics, 55, pp. 180–198

A hybrid commodity price-forecasting model applied to the sugar–alcohol sector

Celma O. Ribeiro and Sydnei M. Oliveira[†]

Accurate price forecasting for agricultural commodities can have significant decision-making implications for suppliers, especially those of biofuels, where the agriculture and energy sectors intersect. Environmental pressures and high oil prices affect demand for biofuels and have reignited the discussion about effects on food prices. Suppliers in the sugar–alcohol sector need to decide the ideal proportion of ethanol and sugar to optimise their financial strategy. Prices can be affected by exogenous factors, such as exchange rates and interest rates, as well as non-observable variables like the convenience yield, which is related to supply shortages. The literature generally uses two approaches: artificial neural networks (ANNs), which are recognised as being in the forefront of exogenous-variable analysis, and stochastic models such as the Kalman filter, which is able to account for non-observable variables. This article proposes a hybrid model for forecasting the prices of agricultural commodities that is built upon both approaches and is applied to forecast the price of sugar. The Kalman filter considers the structure of the stochastic process that describes the evolution of prices. Neural networks allow variables that can impact asset prices in an indirect, nonlinear way, what cannot be incorporated easily into traditional econometric models.

Key words: commodities, forecasting models, Kalman filter, neural networks.

1. Introduction

Forecasts of price and production in commodity markets have been studied for almost one century. The government, farmers, the industry and even players from the financial market are the main users of agricultural forecasts. The first econometric forecast for an agricultural commodity was presented in 1917 (Allen 1994). Since then, a large number of articles, with different approaches, have been published in the area of agricultural forecasting. The objective of this article is to present a hybrid model to forecast spot prices of agricultural commodities and analyse the model's application to the prices of goods derived from sugar cane.

A common approach to forecast prices in commodities markets consists of describing the behaviour of commodity prices and their volatility through stochastic models as presented by Gibson and Schwartz (1990), Schwartz (1997, 1998), Schwartz and Smith (2000) and Manoliu and Tompaidis (2002). Because some variables associated with commodities are non-observable, these authors use the Kalman filter along with futures prices data to calibrate

[†] Celma O. Ribeiro is a Professor, Department of Production Engineering, Escola Politécnica, University of São Paulo, São Paulo, São Paulo, Brazil. Sydnei M. Oliveira (email: sydneim@gmail.com) is a PhD student, Department of Production Engineering, Escola Politécnica, University of São Paulo, São Paulo, São Paulo, Brazil.

doi: 10.1111/j.1467-8489.2011.00534.x

the models. In this sense, this article adopts a new approach, considering exogenous variables whose stochastic behaviour is not described through the Kalman filter. The hybrid model proposed in this article adopts stochastic models to describe commodity prices and includes exogenous variables in the forecast through ANNs. Hu and Tsoukalas (1999), Chen *et al.* (2003), Hamid and Iqbal (2004), Zhang and Qi (2005), Zou *et al.* (2007) among others use neural networks as a forecasting tool in financial and agricultural markets. However, the use of hybrid approaches in agricultural commodities forecasting is new as far as we know. The model is analysed with information of sugar prices in the Brazilian market since March 2002 to December 2008 and in the Indian market from September 2004 to December 2009.

The search for renewable sources of energy at the global level has increased the automobile industry's interest in the use of non-fossil fuels. Environmental pressures for biofuel production and the volatility of oil prices have reignited global discourse about the impact of biofuel production on food prices. This is an important issue that has been the focus of much academic literature that has addressed questions about the potentially detrimental effect of fuel production on food supplies (Zhang *et al.* 2010).

Brazil has a long history of using ethanol as an automotive fuel. Given that, in Brazil, this fuel is derived from sugar cane, producers can choose between different proportions of ethanol and sugar, depending on the financial return of each of the two commodities.

Expectations about spot-market and futures prices, which are driven by expected demand for each subproduct, inform how producers split production. This division can be set even before the sugar cane is planted. Intermediate considerations between planting the cane and delivering the end product need to be taken into account to protect against unfavourable price movements. Such concerns, including changing stock levels and other factors, can cause the division between sugar and ethanol production to be revised. All these factors reinforce the importance of accurate price forecasts.

The main contributions of this article are the following. It addresses using the Kalman filter in hybrid models, generalising simpler approaches that rely on models drawn from univariate time series analysis. The lack of multivariate time series models in this kind of model is a gap this article intends to fill. The second contribution is the construction of commodity price-forecasting models that simultaneously employ both the stochastic process structure for describing prices and exogenous variables. The article also shows the applicability of the proposed model to forecast sugar prices, a complex market related to the energy sector and with increasing importance around the world.

The remainder of this paper is organised as follows. The theoretical framework used to build the model is presented in Section 2. Section 3 briefly discusses the characteristics of the sugar-ethanol markets in Brazil and India. The configuration choices of the model to predict prices in these two selected markets are also presented in this section. The results are reported in Section 4, and the paper is concluded in Section 5.

2. The hybrid forecasting model for agricultural commodity prices

Combination models, in which two or more models are aggregated to increase predictive capacity, have been gaining recognition in recent literature. Zhang (2003) presents a model that unifies ARIMA and neural network models and can be applied to univariate time series. Tseng *et al.* (2008) combine the ARIMA model for seasonal time series with feedforward neural networks to predict the stock market index in Istanbul. Wang (2009) refines the output of a Grey-GJR–GARCH using neural networks to price options.

Hybrid models allow the use of different structural characteristics by, for example, leveraging qualitative variables to refine the output of models that take into account nonlinearities to describe the evolution of prices. Even if these models in most cases do not provide perfect predictions, relative to other approaches, they are quite flexible.

Few works use hybrid models to forecast agricultural commodity prices. Zou *et al.* (2007) compare the use of the ARIMA model with the neural network model in addition to the combination of both models to forecast wheat prices in the Chinese market and conclude that the simple neural network model is most effective in such a case. The Kalman filter is more inclusive and thorough in terms of results than the autoregressive models that are more widely used in hybrid models. The proposed model fills this gap by analysing the Kalman filter with a neural network model. In an approach similar to those of Schwartz (1997), Schwartz and Smith (2000), and Manoliu and Tompaidis (2002), the stochastic process that describes price evolution over time is expressed in the language of state-space models. The model therefore uses the Kalman filter to make an initial price forecast. A second model then refines the output of this first stage through feedforward neural networks. The role of exogenous variables, chosen according to the specifics of the market being analysed, comes into play with the shift from the first model to neural networks.

2.1. Commodity prices and the Kalman filter

Agricultural commodity prices can be described through geometric Brownian motion (Luenberger 1998):

$$\mathrm{d}x(t) = a(x,t)\mathrm{d}t + b(x,t)\mathrm{d}B_t, \tag{1}$$

where $a(x,t)$ is a deterministic tendency represented by the average gain, $b(x,t)$ is a component associated with random error, and B_t is a standard Wiener process.

Let $x(t)$ be the price of a commodity in the spot market at time t ($x(t) = S_t$). Given $a(x,t)$ the expected return on the spot prices of the commodity ($a(x,t) = \mu$) and $b(x,t)$ the standard deviation of the returns ($b(x,t) = \sigma$), it follows that

$$\frac{\mathrm{d}S_t}{S_t} = \mu \mathrm{d}t + \sigma \mathrm{d}B_t \tag{2}$$

Stock-commodity markets exhibit an idiosyncrasy, the convenience yield. According to Frechette and Fackler (1999), the convenience yield can be compared to the dividend yield in equity markets. According to Geman (2005), the convenience yield is the difference between the positive benefit of holding a physical good minus handling costs. Therefore, the convenience yield can be positive or negative depending on the period, the type of commodity and stock levels. In this way, the convenience yield can be expressed as a rate that depends on the handling cost, which in turn depends on such costs as financing, insurance, transportation, storage and primary depreciation (loss of the commodity's value over time). Brennan and Schwartz (1985) note that the convenience yield is the average return, in terms of an interest rate, on the asset that the holder of the physical commodity receives for the possibility of future supply shortages.

Including the convenience yield, c, in Equation (2) gives a particular case of geometric Brownian motion, as shown in Equation (3):

$$\mathrm{d}S_t = (\mu - c)S_t \mathrm{d}t + \sigma S_t \mathrm{d}B_t \tag{3}$$

Considering $X_t = \ln(S_t)$, applying Ito's Lemma and using discreet time gives the equation that describes the spot price between two points in time, t and $t-1$ (Elliott and Hyndman 2007):

$$X_t = \left(\mu - c - \frac{1}{2}\sigma^2\right)\Delta t + X_{t-1} + \sigma\sqrt{\Delta t}\xi_t, \text{ for } t = 1, \ldots, N \tag{4}$$

where N is the number of periods, Δt is the length of the interval used in the discretisation, and ξ_t is a sequence of random i.i.d. variables with standard normal distribution.

It is common to use information from commodity futures markets to forecast spot prices. The equation that relates the prices of futures contracts with specified expiration dates to the spot price is the following:

$$F(S,t,T) = S_t e^{(r-c)(T-t)}, \tag{5}$$

where r is the risk-free rate, t is the current time period, and T is the time of the futures contract's expiration. Given $X_t = \ln(S_t)$ and $Y_t = \ln(F_t)$ and adding noise that represents the random error in relation to the tendency of the future price (Elliott and Hyndman 2007), it follows that

$$Y_t = X_t + (r - c)\,(T - t) + \upsilon\psi_t, \tag{6}$$

where υ is a measure of dispersion and ψ_k is a random variable.

In the case of futures contracts with $M + 1$ different maturities $\{T_i\}_{i=0}^{M}$, it follows that

$$Y_t^i = X_t + (r - c)\,(T_i - t) + \upsilon\psi_t, \text{ for } t = 1, \ldots, N \text{ and } i = 0, \ldots, M \tag{7}$$

Filtering methods such as the Kalman filter can be applied to formulations using discreet time and one or more non-observable variables (Harvey 1989). The Kalman filter is an iterative algorithm that estimates the value of state variables X (observable or non-observable) based on observable variables Y and noise ε.

The algorithm of the Kalman filter estimates the state $x \in \Re^n$ of a process in discreet time governed by the following stochastic model (Welch and Bishop 2001):

$$x_k = Ax_{k-1} + B + Qw_{k-1}, \tag{8}$$

with measurement

$$y_k = Cx_k + D + Rv_k, \tag{9}$$

where x_k is an $n \times 1$ matrix, y_k is an $m \times 1$ matrix, A is an $n \times n$ matrix, B is an $n \times 1$ matrix, C is an $m \times n$ matrix, and D is an $m \times 1$ matrix.

The random variables w_k and v_k represent process noise and measurement noise, respectively. It is assumed that these noises are i.i.d. and are normally distributed with an average of zero and matrices with covariance I_n and I_m respectively – that is, $w \approx N(0, \mathrm{I}_n)$ and $v \approx N(0, \mathrm{I}_m)$. I_n denotes the $n \times n$ identity matrix and I_m denotes the $m \times m$ identity matrix.

Applying Equations (4) and (7) to Equations (8) and (9) produces the following:

$$A = 1, \tag{10}$$

$$B = \left(\mu - c - \frac{1}{2}\sigma^2\right)\Delta t, \tag{11}$$

$$Q = \sigma\sqrt{\Delta t}, \tag{12}$$

$$C = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}, \tag{13}$$

$$D = \begin{bmatrix} (r-c)(T_1-t) \\ \vdots \\ (r-c)(T_M-t) \end{bmatrix}, \tag{14}$$

$$R = \upsilon \tag{15}$$

To use the Kalman filter algorithm, it is necessary to estimate the parameters μ, σ, c and R, which leads to an optimisation problem. The minimisation of the mean squared error (MSE) is used to estimate these parameters.

2.2. Artificial neural networks

In the last two decades, neural networks have been used in numerous applications across various disciplines (Jain and Kumar 2007). In particular, the literature discusses various interesting applications of neural networks to predict complex, nonlinear time series.

These models do not take as given a specific probability distribution for the variables in question. Artificial neural networks are nonlinear models that consist of a structure of neuronal units. The neurons are interconnected with weight w_{jk} determining the intensity of the connections. These weights, which are the parameters of the model, are derived from optimisation techniques that minimise some measures of error. The weights are refined as new data become available, thereby producing dynamic intelligent models.

A key model is the feedforward neural network, in which the neurons do not have connections with those in the previous layer or the same layer. There also is no feedback. A special case is the multilayer perceptron network, which has an input layer, hidden intermediate layers and an output layer. Inputs s_k for neuron k are weighted by the weights w_{jk}, and an activation function F_k generates result y_k, which is the neuron's output. This process is repeated across all layers, and the last output is compared with the target values, with the weights being adjusted such that a given measure of error is minimised.

The output for each neuron k at time t therefore can be represented as follows:

$$y_k(t+1) = F_k(s_k(t)) = F_k\left(\sum_j w_{jk}(t)y_j(t) + \theta_k(t)\right), \tag{16}$$

where $\theta_k(t)$ is an external input, or 'offset.'

The squared error is calculated based on the output of the network y_o^P and on the target value d_o^P:

$$E^P = \frac{1}{2}\sum_{o=1}^{N_o}(d_o^P - y_o^P)^2, \tag{17}$$

where o is the neuron being analysed.

Building a neural network consists first of selecting an architecture – that is, determining the number of layers, the number of neurons in each layer and the activation functions.

Funahashi (1989) and Hornik *et al*. (1990) demonstrate that just one hidden layer with a sigmoid activation function is adequate for the network to be a universal function approximator.

Wang *et al*. (1994) deal with determining the number of neurons in each layer in feedforward neural networks with two hidden layers. In this case, the number of neurons depends on the network's complexity and its input and output dimensions. However, according to Tu (1996), there is no theoretical consensus for predetermining the number of neurons in a hidden layer, which leads to a trial-and-error approach for making this determination.

The weights w_{jk} are obtained though a training method that uses part of the data and iteratively reduces the difference between the network's output value and the target until one of the stopping criteria is reached.

The algorithm first proposed by Levenberg (1944) and perfected by Marquardt (1963), which became known as the LM algorithm, was selected as the training method for the neural network. Hagan and Menhaj (1994), El-Bakry (2003), Kisi (2004), and Cigizoglu and Kişi (2005) all considered the Levenberg–Marquardt algorithm to be the fastest for neural network training and more powerful than other conventional techniques for gradient reduction.

Another important definition is related to the stopping criteria, which guarantee stops in cases of overfitting. Overfitting occurs when the error increases instead of decreasing during the training phase and causes results to deteriorate (Geman *et al*. 1992). Prechelt (1998) compares three stopping criteria by cross-validation and recommends choosing the criterion by generalisation loss to maximise the likelihood of finding a feasible solution, despite this approach increasing processing time. This criterion is the one used in the proposed model.

$$\mathrm{GL}(t) = 100 \cdot \left(\frac{E_{\mathrm{VA}}(t)}{E_{\mathrm{OPT}}(t)} - 1 \right) \tag{18}$$

Equation (18) shows the formula to calculate the generalisation loss at time t, $\mathrm{GL}(t)$, where $E_{\mathrm{VA}}(t)$ is the error that corresponds to the set of validation data used by the stopping criterion and $E_{\mathrm{OPT}}(t)$ is the smallest error of the validation data set obtained at time t.

$$E_{\mathrm{OPT}}(t) = \min_{t' \le t} E_{\mathrm{VA}}(t'). \tag{19}$$

The generalisation loss is a clear indicator to stop the training once overfitting appears. The training is interrupted when the generalisation loss reaches a pre-established minimum parameter.

3. An application in the sugar–alcohol market

To analyse the proposed model, this paper focusses on the sugar–alcohol market because of its importance to the world economy. Ethanol is an important substitute for gasoline, and its production has been growing since the 1970's. Ethanol production is rising rapidly in different parts of the world, making ethanol more competitive (Ajanovic 2010). Sugar is an essential product for food security, and sugar cane is the most efficient feedstock for biofuels (Hira 2010).

The choice of the Brazilian market is because of its important leadership in sugar cane production. Brazil and India are responsible for more than 50 per cent of world sugar production, as shown in Table 1. United States and Brazil are the leading world producers of ethanol, their production was 89 per cent of worldwide ethanol production in 2008 (Mussatto *et al.* 2010).

3.1. The sugar–alcohol market

In the last decade, the ethanol market has gained more importance around the globe because of environmental pressures for reducing carbon dioxide emissions and as an alternative to expensive oil. Today, not just businesses are aware of opportunities in this sector, but government efforts are pushing the industry to grow. The RIRDC (2010) report shows the Australian government efforts to increase the bioenergy sector. As pointed out by Sumner (2000), there still exist barriers, such as domestic subsidies, that prevent faster expansion.

The Brazilian sugar–alcohol market has become complex in recent years. Ethanol and sugar are the principal products generated from sugar cane. According to Moraes (2007), sugar production from sugar cane began in the colonial period, roughly in 1520, and was Brazil's primary export product.

Table 1 Top ten sugar cane producers – 2008

Country	Production (million tonnes)
Brazil	649
India	348
China	125
Thailand	74
Pakistan	64
Mexico	51
Colombia	39
Australia	34
Argentina	30
United States	28
World	1743

Source: Food and Agricultural Organization of United Nations: Economic and Social Department: The Statistical Division. Retrieved 2010-06-17.

The production of fuel alcohol from sugar cane, on the other hand, began in the 1970s with the Proálcool programme, which came into being to promote an alternative to oil, which was very expensive at the time. The processes for producing sugar and alcohol are similar. They differ only in the steps that follow the extraction of cane juice (ATR), when the producer determines the proportion of sugar and alcohol to be produced (Goldemberg and Moreira 1999).

Over the last decade in Brazil, the availability of flex-fuel cars capable of running on any mixture of gasoline and alcohol has driven up the domestic demand for alcohol significantly. Despite commercial barriers, demand for ethanol also has increased worldwide. The growth in demand for ethanol in Brazil altered the balance of production between alcohol and sugar, thanks to the ease with which producers could switch between sugar and alcohol. The producer studies price levels and decides on the level of production that offers the best return, keeping in mind the minima required to meet factory-operating costs. This change in the profile of the market, in conjunction with the existence of innumerable financial instruments to manage risk, helped the sugar–alcohol sector mature and led to an increased need for price-forecasting models.

In Brazil, the government offers price floors for alcohol as a guarantee to producers. Besides supply, the price of alcohol also depends on domestic demand, which in turn depends on the prices of gasoline and the quantity of flex vehicles that run on alcohol and/or gasoline. Because of market fragmentation, there is no easily ascertainable market price. The price of sugar, on the other hand, is strongly affected by the international market and depends on the global harvest and worldwide demand. The exchange rate can affect the relative prices of alcohol and sugar and can cause producers to opt to produce more alcohol and less sugar. This same phenomenon can occur because of expectations about demand for alcohol.

It is important to note that commodity prices in the futures, contracts and spot markets are strongly related (Geman 2005). Commodity futures prices set though trading on exchanges are important because they reflect the expectations of global supply and demand, while spot prices often depend on the peculiarities of the local markets in which negotiations take place (Dooley and Lenihan 2005). In Brazil, the sugar futures market has existed since 1995 and the alcohol futures market since 2000. The sugar futures market, which sets prices for crystalline sugar, is liquid, at least for first-expiration contracts, despite not having matched the growth in the volume of the physical market. Liquidity in the alcohol futures market, however, is low regardless of expiration and has been drying up over time, which indicates a low level of representativeness for future spot prices.

Various authors such as Gibson and Schwartz (1990), Schwartz (1997, 1998), Schwartz and Smith (2000), and Manoliu and Tompaidis (2002) use futures prices to forecast commodity spot prices. The literature also cites other variables that are related to commodity prices. Akram (2009) confirms the strong relationships between commodity prices, interest rates and the value of the dollar. Thompson *et al*. (2009) address the impact of the prices of oil and corn and regulations specifying the mixture of ethanol with gasoline on the price of ethanol. Farinelli *et al*. (2009) discuss the impact of demand in countries that import Brazilian ethanol on the price of ethanol to predict the ways prices will react to policy changes in those countries. Besides depending on exogenous variables, commodity prices like sugar and ethanol present idiosyncrasies, such as the presence of the convenience yield. To accommodate the structure of the stochastic process and to incorporate the effects of exogenous variables, the hybrid model will be used to forecast prices.

3.2. Forecasting sugar prices

A study using actual data from Brazil and India to forecast sugar spot prices 1 month in advance was used to evaluate the proposed hierarchical model. The sugar spot-market data come from contracts for crystalline sugar, as is the case with futures contracts negotiated on the Brazilian Mercantile and Futures Exchange or on the New York Mercantile Exchange.

The lack of a liquid market in alcohol futures, or even an average spot price, rendered impossible the same study for alcohol prices, because the neural network model uses spot-market data for calibration.

3.2.1. Stochastic processes and Kalman filter model

In the Brazilian study, 82 observations of monthly spot prices for sugar, from March 2002 to December 2008, supplied by the Center for Advanced Studies in Applied Economics (*Centro de Estudos Avançados em Economia Aplicada*, or CEPEA/Esalq), are the starting point for the first model, along with first-expiration sugar futures prices taken from the Mercantile and Futures Exchange (*Bolsa de Mercadorias e Futuros*, or BM&FBOVESPA).

In the Indian study, spot and future prices were used to adjust the Kalman filter. Sixty-four observations of monthly spot prices for sugar, from September 2004 to December 2009, supplied by National Commodity & Derivatives Ltd (CDEX) were used. Futures prices from New York Mercantile Exchange were considered in this case. Estimates of the average return μ, the volatility of the spot price σ, and the convenience yield c are made by determining the parameters of the Kalman filter through minimisation of the MSE.

A range of measures including the mean absolute error (MAE), the MSE, the mean absolute percentage error (MAPE) and the variance was analysed. The values of each measure are presented in Equations (20–23).

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n} |e_i|, \tag{20}$$

$$\text{MSE} = \frac{1}{n}\sum_{i=1}^{n} e_i^2, \tag{21}$$

$$\text{MAPE} = \frac{1}{n}\sum_{i=1}^{n} \left|\frac{e_i}{A_i}\right|. \tag{22}$$

$$\text{Variance} = \frac{1}{n}\sum_{i=1}^{n} (e_i - \bar{e})^2, \tag{23}$$

In the expressions above, $|e_i|$ is the absolute error, or difference between the forecasted and actual values, at time i. The actual value is $|A_i|$, and the average error is $\bar{e}$.

A base model that assumes that the price in the following month is equal to the price in the current month was used for purposes of comparison. Table 2 shows various measures of error produced by the base model and the Kalman filter model for both markets.

For Brazil, the Kalman model reduced MAE by 5.4 per cent, the MSE by 64.7 per cent, the MAPE by 40.3 per cent and the variance by 64.9 per cent relative to the base model.

For India, all error measures were also improved, showing similar results to those obtained for Brazil. Figure 1 shows the results of the Kalman filter's forecasts based on sugar futures prices using Brazilian Data.

3.2.2. Artificial neural network correction

It is expected that the neural network can further reduce measures of error and the variance of the error, even with the inclusion of new variables. To this end, in addition to forecasted values from the Kalman filter model and futures prices, new variables were tested. Table 3 presents exogenous variables taken into account and the data sources used for the Brazilian forecast. Table 4 presents the variables for the Indian forecast.

Table 2 Results of the model using stochastic processes and the Kalman filter

Error measure	Brazil		India	
	Base	Kalman filter	Base	Kalman filter
MAE	0.0822	0.0778	0.3495	0.0423
MSE	2.1362	0.7539	1.4306	0.6661
MAPE (%)	8.88	5.30	3.53	3.51
Variance	2.1560	0.7572	1.3293	0.6749

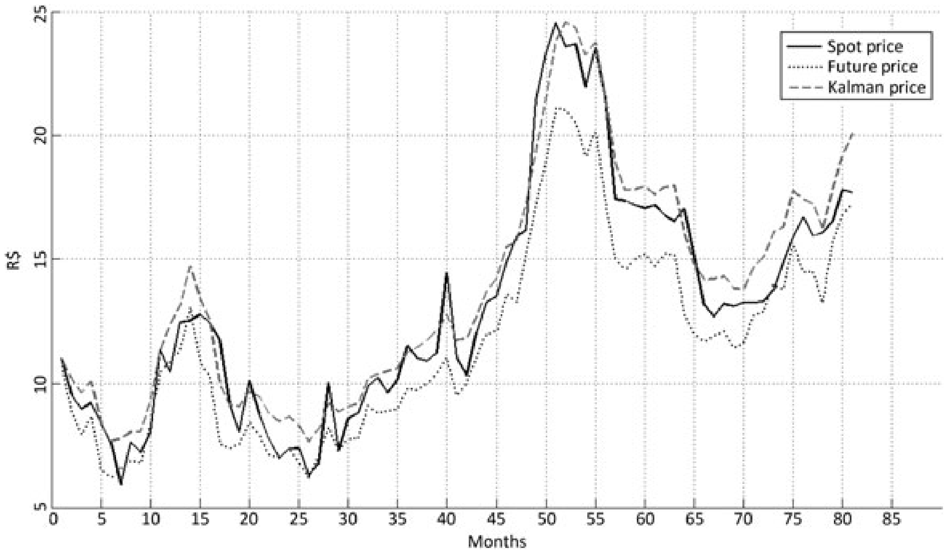

Figure 1 Results of the model with stochastic processes and the Kalman filter (Brazilian Data).

Table 3 Exogenous variables used in the neural network model (Brazilian Data)

Variable (Brazil)	Abbreviation	Source
Sugar cane production	SCP	Sugar Cane Industry Association (UNICA)
Sugar production	SPB	Sugar Cane Industry Association (UNICA)
Agricultural GDP	GDP	Institute for Applied Economic Research (IPEA)
Exchange rate (Dollar/Real)	ERR	Mercantile and Futures Exchange (BM&FBOVESPA)
Oil price per barrel	OPB	New York Board of Trade (NYBOT)
Sales of biofuel vehicles	SBV	National Association of Automobile Manufacturers (ANFAVEA)
Interest rate (SELIC)	SLC	Central Bank of Brazil

Table 4 Exogenous variables used in the neural network model (Indian Data)

Variable (India)	Abbreviation	Source
Exchange rate (Dollar/Rupee)	ERR	Reserve Bank of India
Oil price per barrel	OPB	New York Board of Trade (NYBOT)

The process of evaluating each variable's capacity to reduce mean error and its variance followed the following steps:

1. Choosing one of the variables.
2. Testing various neural network models, varying the number of intermediate-layer neurons.
3. Training and simulation.

Geman and Ohana (2009) highlight the importance of the volume of stocks to describe the behaviour of prices, insofar as this variable is capable of explaining differences between prices in spot and futures markets.

For Pereira (2009), on certain occasions, the spot price of a good can remain above the price on the futures market. This occurs because the holder of the physical good benefits from possession by being able to search for the highest price and wait for the best moment to sell. In making the decision about when to sell, the holder must take into account certain factors, such as interest rates, the seasonality of prices, the cost of storage and any expected increase in prices.

3.2.2.1. Brazilian data

In Brazil, information about stocks of sugar and alcohol is not updated with enough regularity and reliability to permit this data to reflect the actual levels of domestic stocks and therefore cannot be used. Sugar cane and sugar production directly reflect supply in a given year and in turn affect stock levels and price.

The agricultural component of GDP, which measures all Brazil's agricultural output, also can indirectly indicate the supply of sugar cane in the market. In Brazil, data on agricultural GDP and the production of sugar cane and sugar are updated annually but need to be estimated monthly to be applicable.

Although it has grown significantly in the last 3 years, the exportation of ethanol remains at low levels. Still, Brazilian exports are expected to increase substantially because of new carbon dioxide emissions targets and the growth of biofuel programmes in European and Asian countries.

In Brazil, the interest rate is represented by the base rate, called the SELIC, set by the central bank as the rate on Brazilian government bonds. The interest rate affects decisions about stocks and therefore affects supply and price. The interest rate appears explicitly in the part of the model that relates prices in the spot and futures markets (see Eqn 6).

As there are no data sources for monthly sales of or demand for alcohol, the monthly sales of vehicles powered by biofuel can be taken as a proxy for potential demand. According to the Sugar Cane Industry Association (UNICA, 2009), 50 per cent of the on-road vehicle fleet will run on biofuel in 2012, and this figure will increase to 65 per cent by 2015. UNICA, 2009 also reports that 50 per cent of gasoline consumption has been replaced with ethanol; this explicitly defines the rate of substitution of ethanol for oil. The exchange rate has the capacity significantly to improve the forecasting of commodity prices as shown by Chen *et al.* (2008) and Engel and West (2005).

3.2.2.2. Indian data

The study of the Indian market was conducted after the Brazilian study was concluded. Because of the lack of information about the Indian market, only the two variables that gave good results in the Brazilian case were used in the

Indian case. The exogenous variables considered were the rupee exchange rate and oil prices.

3.2.2.3. Architecture and procedures

The architecture of the network was determined according to the relevant literature. The neural network used possesses three input neurons in the first layer: the spot price forecasted one period into the future, obtained from the Kalman filter, the futures-market price and one of the seven variables listed in Table 3. The decision to analyse only one exogenous variable as input was because of the presence of overfitting in analyses performed with two or more variables. Attempts were made with two or more variables, and, even relaxing the stopping criteria and performing various simulations, the level of overfitting impeded any improvement in the results. The network used has only one hidden layer, as recommended in the literature (Funahashi 1989; Hornik *et al*. 1990). Because there is no consensus on the number of neurons in the hidden layer (Tu 1996), it is necessary to test various configurations. This study tested configurations with one, two and three neurons in the hidden layer; the best results were found with two neurons in the hidden layer. There is a range of options for activation functions, including the linear, the sigmoid, the hyperbolic tangent and the Gaussian functions. The convergence of the neural network depends upon choosing the proper activation function. The sigmoid function is the one most commonly found in financial applications and has been chosen for the model at hand (Funahashi 1989; Hornik *et al*. 1990). Figure 2 illustrates the architecture of the neural network employed.

The procedure that follows was used to train the network. In the Brazilian study, of the 82 months considered, a set of prices from the first 40 months was used to train the network. In the Indian study, of the 64 months considered, a set of prices from the first 35 months was used to train the network. In accordance with Zhang (2003), the total sample size has a strong influence on the convergence of the error, and large data sets normally are necessary to guarantee that the model performs satisfactorily.

Although, in the Brazilian study, more than 82 months of spot market data exist, it would not be possible to obtain such data for the sugar futures market, which is relatively new in Brazil. In addition, the use of vehicles powered by biofuel only has become prominent since 2004, rendering the

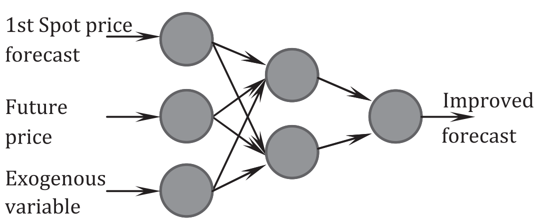

Figure 2 Structure of the neural network.

increase in demand for ethanol in earlier years irrelevant. Therefore, even with a greater likelihood of overfitting, only 82 months were taken into consideration. In the Indian study, the data set is even more reduced, because there are no data about sugar spot prices before 2004.

The variance of the error, the maximum number of periods and cross-validation by loss generalisation were used as the stopping criteria for the Levenberg–Marquardt training algorithm. The variance of the error should be less than that which the previous model has produced.

The maximum number of periods consists of a full scan of the test data and can guarantee an early stop as well as a reduction in processing time. The algorithm used already relies on cross-validation by loss generalisation to reduce the effect of overfitting, guaranteeing an early stop in case the error begins rising instead of falling.

4. Results

4.1. Brazilian study

Table 5 summarises the results for each variable in comparison with the Kalman filter output.

The results demonstrate that two variables, sugar cane production (SCP) and sugar production (SPB), do not improve the MAE relative to the results forecast by the Kalman filter model, perhaps because they are annual variables that must be estimated monthly.

Monthly agricultural GDP also failed to improve the results in comparison with the first model, possibly because it takes into account products that are unrelated to sugar cane output. The interest rate SELIC (SLC) did reduce the MAE but not the MAPE in relation to the Kalman filter model's output.

Oil price, biofuel vehicle sales and the exchange rate were the variables with the most significant impacts. They reduced the values of all error measurements, with the exchange rate and the oil price having the greatest effects.

Figure 3 shows the output of the Brazilian neural network model (by month) using the exchange rate (ERR) as its variable and compares this result

Table 5 Results of the neural network model for each exogenous variable (Brazilian Data)

Error	Kalman filter	Exchange rate	Sales of biofuel vehicles	Oil price per barrel	Interest rate (SELIC)	Agricultural GDP	Sugar cane production	Sugar production
MAE	0.0778	0.0280	0.0363	0.0105	0.0562	0.0679	0.1227	0.4143
MSE	0.7539	0.6026	0.6562	0.6044	0.6835	0.6588	0.6813	0.9047
MAPE (%)	5.30	5.07	5.19	5.03	5.30	5.00	5.17	6.41
Variance	0.7572	0.6093	0.6630	0.6119	0.6888	0.6624	0.6746	0.7422

both to actual prices (in Brazilian Real) and to those forecast by the Kalman filter model.

4.2. Indian study

The results for each variable in comparison with the Kalman filter output are presented in Table 6. The exchange rate and oil prices showed similar results; neither of them reduced the MAE, but they reduced the other error measures. The small sample size used to forecast prices, only 64 months, may have contributed to these results, as pointed by Zhang (2003).

5. Conclusions

This paper used information from the two major sugar cane-producing countries to examine the potential of a hybrid model to predict prices. The Kalman filter was used to describe prices as stochastic processes and also to include future prices in the forecasting. Exogenous variables whose influence in prices cannot be explicitly described were included in the model through an ANN. The Kalman filter had good performance in prediction, and the neural networks were responsible for the improvement of the results. In both cases studied, the proposed model showed high potential to predict the sugar price 1 month ahead.

For the Brazilian data, the variables that had significant impact in forecasting sugar prices were the oil price, sales of biofuel vehicles in Brazil and the exchange rate. These variables reduced all the error measures examined in the

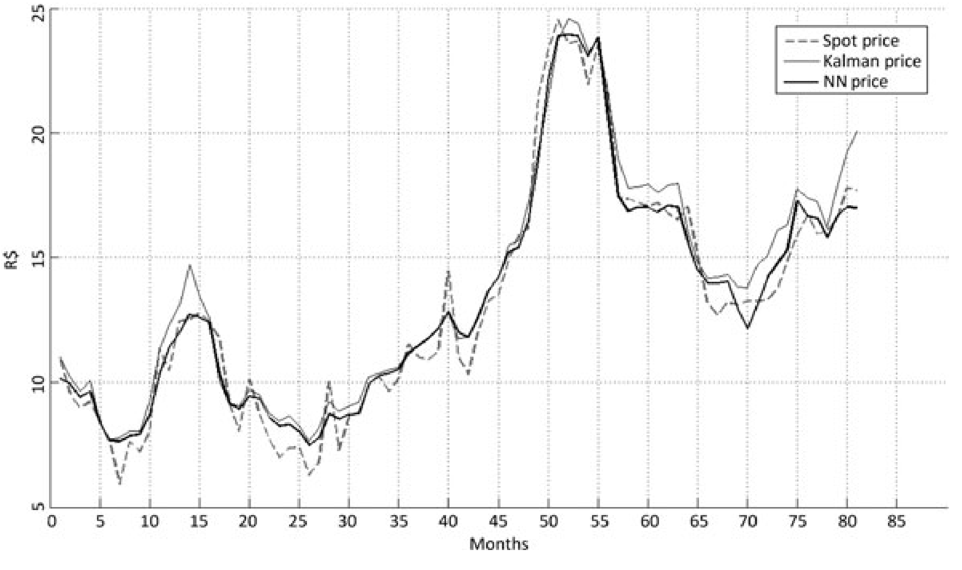

Figure 3 Results of the neural network model with the exchange rate as input (Brazilian Data).

Table 6 Results of the neural network model for each exogenous variable (Indian Data)

Error	Kalman filter	Exchange rate	Oil price per barrel
MAE	0.0423	0.1378	0.1351
MSE	0.6661	0.5941	0.6082
MAPE (%)	3.51	3.21	3.23
Variance	0.6749	0.5842	0.5993

study. For the Indian data, the exchange rate and oil prices improved the Kalman filter model, even though they could not reduce the MSE. Although there were more data available about Brazilian production, information about commodity production was not relevant to determining Brazilian prices. It may have happened because of the lack of monthly data because annual production is the only available information. Monthly data related to sugar cane production and stock levels would make a deeper analysis easier.

There exist several potential extensions of this research. The proposed hierarchical model could be applied to other agricultural commodities. The Kalman filter's structure is adequate for describing any stochastic process that includes a convenience yield and therefore is appropriate for use across commodities markets. By extension, the hybrid model could be used to forecast prices simultaneously in several commodities markets. In terms of the mathematical model, other structures besides the feedforward network naturally could be applied. The use of robust optimisation techniques to reduce uncertainties in the model's parameters is another approach to be considered.

References

Ajanovic, A. (2010). Biofuels versus food production: does biofuels production increase food prices? *Energy*, in press, doi:10.1016/j.energy.2010.05.019.

Akram, Q.F. (2009). Commodity prices, interest rates and the dollar, *Energy Economics*, in press, doi:10.1016/j, eneco.2009.05.016.

Allen, P.G. (1994). Economic forecast in agriculture, *International Journal of Forecasting* 10, 81–135.

Brennan, M.J. and Schwartz, E.S. (1985). Evaluating natural resource investments, *The Journal of Business* 58, 135–157.

Chen, A., Leung, M.T. and Daouk, H. (2003). Application of neural networks to an emerging financial market, Forecasting and trading the Taiwan Stock Index, *Computers and Operations Research* 30, 901–923.

Chen, Y., Rogoff, K. and Rossi, B. (2008). Can exchange rates forecast commodity prices?, National Bureau of Economic Research, Working Paper No. 13901.

Cigizoglu, H.K. and Kişi, O. (2005). Flow prediction by three back propagation techniques using k-fold partitioning of neural network training data, *Nordic Hydrology* 36, 49–64.

Dooley, G. and Lenihan, H. (2005). An assessment of time series methods in metal price forecasting, *Resources Policy* 30, 208–217.

El-Bakry, M.Y. (2003). Feed forward neural networks modeling for K–P interactions, *Chaos, Solutons & Fractals*, 18, 995–1000.

Elliott, R.J. and Hyndman, C.B. (2007). Parameter estimation in commodity markets: a filtering approach, *Journal of Economic Dynamics and Control* 31, 2350–2373.

Engel, C. and West, D. (2005). Exchange rates and fundamentals, *The Journal of Political Economy* 113, 485–517.

Farinelli, B., Carter, C.A., Lin, C.C.Y. and Sumner, D.A. (2009). Import demand for Brazilian ethanol: a cross-country analysis, *Journal of Cleaner Production*, in press, doi: 10.1016/, jclepro.2009.05.008.

Frechette, D.L. and Fackler, P.L. (1999). What causes commodity price backwardation?, *American Journal of Agricultural Economics* 81, 761–771.

Funahashi, K. (1989). On the approximate realization of continuous mappings by neural networks, *Neural Networks* 2, 183–192.

Geman, H. (2005). *Commodities and Commodity Derivatives*, 1st edn. John Wiley & Sons Ltd, Chichester.

Geman, H. and Ohana, S. (2009). Forward curves, scarcity and price volatility in oil and natural gas markets, *Energy Economics* 31, 576–585.

Geman, S., Bienenstock, E. and Doursat, R. (1992). Neural networks and the bias/variance dilemma, *Neural Computation* 4, 1–58.

Gibson, R. and Schwartz, E.S. (1990). Stochastic convenience yield and the pricing of oil contingent claims, *Journal of Finance* 45, 959–976.

Goldemberg, J. and Moreira, J.R. (1999). The alcohol program. *Energy Policy* 27, 229–245.

Hagan, M.T. and Menhaj, M.B. (1994). Training feed forward networks with the Marquardt algorithm, *IEEE Transactions Neural Networks* 6, 861–867.

Hamid, S.A. and Iqbal, Z. (2004). Using neural networks for forecasting volatility of S&P500 Index Futures Prices, *Journal of Business Research* 57, 1116–1125.

Harvey, A.C. (1989). *Forecasting, Structural Time Series Models and the Kalman Filter*, 1st edn. Cambridge University Press, Cambridge.

Hira, A. (2010). Sugar rush: prospects for a global ethanol market. *Energy Policy*, in press, doi:10.1016/j.enpol.2010.05.035.

Hornik, K., Stinchcombe, M. and White, H. (1990). Universal approximation of an unknown mapping and its derivatives using multilayer neural networks, *Neural Networks* 3, 551–560.

Hu, M.Y. and Tsoukalas, C. (1999). Combining conditional volatility forecasts using neural networks: an application to the Ems exchange rates, *Journal of International Financial Markets* 9, 407–422.

Jain, A. and Kumar, A.M. (2007). Hybrid neural network models for hydrologic time series forecasting, *Applied Soft Computing* 7, 585–592.

Kisi, O. (2004). Multi-layer perceptrons with Levenberg-Marquardt training algorithm for suspended sediment concentration prediction and estimation, *Hydrological Sciences–Journal–des Sciences Hydrologiques* 49, 1025–1040.

Levenberg, K. (1944). A method for the solution of certain problems in least squares, *The Quarterly of Applied Mathematics* 2, 164–168.

Luenberger, D.G. (1998). *Investment Science*, 1st edn. Oxford University Press, New York.

Manoliu, M. and Tompaidis, S. (2002). Energy futures prices: term structure models with Kalman filter estimation, *Journal of Applied Mathematical Finance* 9, 21–43.

Marquardt, D. (1963). An algorithm for least-squares estimation of nonlinear parameters, *SIAM Journal of Applied Mathematics* 11, 431–441.

Moraes, M.A.F.D. (2007). As profundas mudanças institucionais ao longo da história da agroindústria canavieira e os desafios atuais, *Economia Aplicada* 11, 555–557.

Mussatto, S.I., Dragone, G., Guimarães, P.M.R., Silva, J.P.A., Carneiro, L.M., Roberto, I.C., Vicente, A., Domingues, L. and Teixeira, J.A. (2010) Technological trends, global market, and challenges of bio-ethanol production-R1, *Biotechnology Advances*, in press, doi:10.1016/j.biotechadv. 2010.06.003.

Pereira, L.M. (2009). Modelo de formação de preços de commodities agrícolas aplicado ao mercado de açúcar e álcool, PhD, São Paulo University, Brazil.

Prechelt, L. (1998). Automatic early stopping using cross validation: quantifying the criteria, *Neural Networks* 11, 761–767.

RIRDC – Rural Industries Research and Development Corporation (2010). Overview of bioenergy in Australia, Bioenergy Australia.

Schwartz, E.S. (1997). The stochastic behaviour of commodity prices: implications for valuation and hedging, *Journal of Finance* 52, 923–973.

Schwartz, E.S. (1998). Valuing long-term commodity assets, *Financial Management* 27, 57–66.

Schwartz, E.S. and Smith, J.E. (2000). Short-term variations and long-term dynamics in commodity prices, *Management Science*, 46, 893–911.

Sumner, D.A. (2000). Domestic support and the WTO negotiations, *The Australian Journal of Agricultural and Resource Economics*, 44, 457–474.

Thompson, W., Meyer, S. and Westhoff, P. (2009). How does petroleum price and corn yield volatility affect ethanol markets with and without an ethanol use mandate?, *Energy Policy* 37, 745–749.

Tseng, C.H., Chenga, S.T., Wangb, Y.H. and Pengc, J.T. (2008). Artificial neural network model of the hybrid EGARCH volatility of the Taiwan stock index option prices, *Physica A: Statistical Mechanics and its Applications* 387, 3192–3200.

Tu, J.V. (1996). Advantages and disadvantages of using artificial neural networks versus logistic regression for predicting medical outcomes. *Journal of Clinical Epidemiology* 49, 1225–1231.

UNICA. União da Indústria de Cana-de-açúcar (2009). Metade da frota nacional em 2012 será de bicombustíveis, Available from URL: http://g1.globo.com/Noticias/Carros/0,,MUL1090989-9658,00-METADE + DA + FROTA + NACIONAL + EM + SERA + DE + BICOMBUSTIVEIS + DIZ + UNICA.html [accessed 02 October 2009].

Wang, Y.H. (2009). Nonlinear neural network forecasting model for stock index option price: hybrid GJR–GARCH approach, *Expert Systems with Applications* 36, 564–570.

Wang, Z., Massimo, C.D., Tham, M.Z. and Morris, S.J. (1994). A procedure for determining the topology of multilayer feedforward neural networks, *Neural Networks* 7, 291–300.

Welch, G. and Bishop, G. (2001). *An Introduction to the Kalman Filter*, 1st edn. Siggraph, Chapel Hill.

Zhang, G. (2003). Time series forecasting using a hybrid ARIMA and neural network model, *Neurocomputing* 50, 159–175.

Zhang, G. and Qi, M. (2005). Neural network forecasting for seasonal trend series, *European Journal of Operational Research* 160, 501–514.

Zhang, G., Lohr, L., Escalante, C. and Wetzstein, M. (2010). Food versus fuel: what do prices tell us? *Energy Policy* 38, 445–451.

Zou, H.F., Xia, G., Yang, F.T. and Wang, H.Y. (2007). An investigation and comparison of artificial neural networks and time series models for Chinese food grain prices, *Neurocomputing* 70, 2913–2923.