LAPPEENRANTA UNIVERSITY OF TECHNOLOGY
Faculty of Technology
Degree Programme in Technomathematics and Technical Physics

Emelyne Umunoza Gasana

COMPUTATIONAL MARKET DYNAMICS STUDIES OF SUGAR MARKETS

Examiners: docent Ph.D. Tuomo Kauranne
D.Sc. Matylda Jabłońska
Abstract

Lappeenranta University of Technology
Faculty of Technology
Degree Programme in Technomathematics and Technical Physics

Emelyne Umunoza Gasana

Computational Market Dynamics Studies of Sugar Markets

Master's thesis

2013

52 pages, 34 figures, 5 tables

Examiners: docent Ph.D. Tuomo Kauranne
D.Sc. Matylda Jabłońska

Keywords: Sugar prices, Stochastic modelling, Ensemble modelling, Animal spirits, Market momentum.

Commodity markets have stimulated the curiosity of numerous researchers in different aspects for decades. Several models of simulating but also predicting prices have been developed in financial econometrics, but with little success. In this study, computational market dynamics studies of the US sugar market are applied. Two ensemble coupled mean reversion models have been used for this purpose, that is, the Jabłońska-Capasso-Morale (JCM) and Kalman Dynamics (KD) models. Both models were fitted to the data using the Maximum Likelihood method. These models are applied on a set of monthly US sugar price observations from January, 1960 to February, 2012. The two models were observed to be fairly able to follow the movement of the given series of sugar prices, even though they presented some non-significant differences.
Acknowledgements


Emelyne UMUNOZA GASANA
List of Symbols and Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACF</td>
<td>Autocorrelation Function</td>
</tr>
<tr>
<td>BC</td>
<td>Before Christ</td>
</tr>
<tr>
<td>EU</td>
<td>European Union</td>
</tr>
<tr>
<td>FAS</td>
<td>Foreign Agricultural Service</td>
</tr>
<tr>
<td>GARCH-M</td>
<td>Generalized Autoregressive Conditional Heteroscedastic in Mean</td>
</tr>
<tr>
<td>ISO</td>
<td>International Organization for Standardization</td>
</tr>
<tr>
<td>JCM</td>
<td>Jabłońska-Capasso-Morale</td>
</tr>
<tr>
<td>KD</td>
<td>Kalman Dynamics</td>
</tr>
<tr>
<td>MLEM</td>
<td>Mixture of Local Expert Model</td>
</tr>
<tr>
<td>MMT</td>
<td>Million Metric Tons</td>
</tr>
<tr>
<td>MRCPM</td>
<td>Mean-Reverting Commodity Price Model</td>
</tr>
<tr>
<td>ODE</td>
<td>Ordinary Differential Equation</td>
</tr>
<tr>
<td>PACF</td>
<td>Partial Autocorrelation Function</td>
</tr>
<tr>
<td>RMSD</td>
<td>Root Mean Square Deviation</td>
</tr>
<tr>
<td>RMSE</td>
<td>Root Mean Square Error</td>
</tr>
<tr>
<td>SDE</td>
<td>Stochastic Differential Equation</td>
</tr>
<tr>
<td>SMW</td>
<td>Sherman-Morrison-Woodbury</td>
</tr>
<tr>
<td>US</td>
<td>United States</td>
</tr>
<tr>
<td>USA</td>
<td>United States of America</td>
</tr>
<tr>
<td>USD</td>
<td>United States dollar</td>
</tr>
<tr>
<td>USDA</td>
<td>United States Department of Agriculture</td>
</tr>
<tr>
<td>VEnKF</td>
<td>Variational Ensemble Kalman Filtering</td>
</tr>
<tr>
<td>VECM</td>
<td>Vector Error Correction Model</td>
</tr>
<tr>
<td>WTO</td>
<td>World Trade Organization</td>
</tr>
</tbody>
</table>
1 INTRODUCTION

Sugar is a sweet-flavored substance efficiently extracted in sufficient concentrations from sugarcane. The cultivation of sugarcane, first discovered in New Guinea [1], spread all over the world and Brazil is currently the world’s dominant sugar producer, followed by the USA. The production of sugarcane, from which sugar is extracted, is predicted to rise in years to come, mainly resulting from the recent discovery of sugar-based ethanol, extracted from sugarcane as well [19]. The market of sugar, the second most-traded commodity in the world after oil [2], has elicited the curiosity of several researchers. In order to capture the dynamics of commodity prices, one might use a model able to simulate the behavior of its price.

The analysis of financial and commodity market behavior has been of a great interest to researchers for decades. Therefore, various models able to simulate commodity prices have been developed in financial econometrics with the aim to use them for other purposes such as prediction and risk minimization. Researchers have tried to build models explaining the prices in financial markets using econometric models, in order to be able to forecast future prices, but have failed in predicting the emergence and scale of financial crises in these markets. That is due to the high volatility of prices, which is one of the major sources of risk in commodity markets. The primary reasons for those failures were found to be irrational actions of traders described as trading biases (of the human psychology), referred to as animal spirits in this study. In financial markets, these animal spirits stem mainly from worry and greed of the traders. This pulls them to behavioral patterns such as herding, overconfidence or short-term thinking.

In addition, those failures may also result from the presence of momentum in financial markets, it means, the fact that the prices deviate from the global tendency due to the traders’ pursuit of gaining high profits [15].

Sugar market is a deregulated market showing sudden price changes, including price spikes that express high volatility in the prices. Sugar prices appear as mean-reverting prices. Due to the specificity of this type of data it seems appropriate to model them using stochastic approaches. Traders’ behavior and the presence of momentum in financial markets are among different other major reasons explaining these irregularities. Several researchers in their simulation models seem to ignore these actions whose effects are too big to be taken as
normal irregularities. In this study, US sugar prices are modeled using two models that take into account the irrational actions of traders, also called animal spirits and the momentum effect of financial markets.

The aim of this work is to simulate the US monthly sugar prices, from January, 1960 to February, 2012, applying the computational market dynamics approaches. In this study, two models which make use of stochastic processes are used to capture the dynamics of the sugar prices, namely, the Jabłońska-Capasso-Morale (JCM) and the Kalman Dynamics (KD) models which take into account the trading biases known as animal spirits, and the momentum phenomenon of the financial markets. Furthermore, the Maximum Likelihood approach is used to fit both models to the original US sugar data. Graphical representations and numerical values of basic statistical properties are used to analyze the performances of the two models. In addition, the closeness of the simulated prices to the original will be measured using the Root Mean Square Errors (RMSE) for models’ comparison.

The format of this study work is divided into 7 sections. Section 2 reviews the sugar market, including its description and history, its functioning and reviews some models applied in the past on the sugar market. Section 3 describes the data set of the sugar prices used in this work. In Section 4, a full description of the two models, the JCM and KD models, applied in this work, is provided with an introduction to stochastic differential equations and modelling, widely used by the two models. Section 5 presents the modelling results obtained using the MATLAB software with a comparison of the two models’ performances. In Section 6, a summary of the obtained results and their discussion are provided. And finally, Section 7 provides the conclusion of the study.

2 SUGAR MARKET

2.1 Sugar market description and history

Sugar is the generalized name for a class of sweet-flavored substances used as food. It is a carbohydrate that provides a source of energy. It can take many forms including white, raw or brown sugar, honey or corn syrup [4]. It is efficiently extracted in sufficient concentrations from sugarcane.
The discovery of sugarcane is considered to have happened in New Guinea, and thereafter was spread to Southeast Asia and India. The process of sugar production by pressing out the juice and then boiling it into crystals was developed in India around 500 B.C. [1].

Many years later, the cultivation of sugarcane continued to spread along to Europe, in the middle-ages, brought by Arabs to Spain. Later it began in the United States in the 18th Century, in the Southern climate of New Orleans. However, after several attempts, it was only a while before 1872, that a factory built in California, was finally able to produce sugar profitably. And thus, more than thirty factories were operating in the U.S. at the end of that century [2].

The sugar industry has a large impact on a nation’s burgeoning economy, especially in the major producers of sugar cane like Asia, South America, North America, and those of sugar beet like Germany, France and the USA, where for instance, the American sugar industry creates more than 420000 jobs in 42 states and will contribute twenty six billion dollars-plus in economic activity annually [1].

However, the long history of sugar is twisted with that of trade, religion, colonialism, capitalism, industry and technology, which makes it bittersweet. The labor-intensive nature of sugar cultivation and processing has meant that much of the history of the sugar industry has had associations with large-scale slavery. "Nowhere in the world did the sugar business have a more destructive effect than in Africa", stated Abbott [2].

### 2.2 Sugar market functioning

Sugar has being widely traded for its precious value since many centuries ago. Sugarcane offers production substitutes to food, such as feed, fibre and energy such as biofuels (sugar-based ethanol) and/or co-generation of electricity [19]. Sugar is, after oil, the second most-traded commodity in the world [2]. Global sugar production for the 2012/13 marketing year is forecast at 174 million metric tons (MMT) raw value, up 2% from last year, as shown in Figure 1. It is forecast to rise for most of the largest producers, says the Foreign Agricultural Service (FAS)/United States Department of Agriculture (USDA) Office of the Global Analysis and approved by the World Agricultural Outlook Board/USDA [26]. This rise primarily results from the relationships between
sugar as the main ethanol feedstock in Brazil, the major worldwide producer and exporter of both sugar and ethanol [19]. In addition, the competitiveness between the corn ethanol produced in the US and sugarcane ethanol produced in Brazil, these two countries being the world’s leading sources of biofuel, may explain, this rise since the cost of sugarcane ethanol production in Brazil is lower than that of corn ethanol in the US [9]. However, this general perception that sugarcane ethanol presents lower costs is not always valid and depends on the prevailing exchange rate as well as the price of feedstocks, including the cost of transportation from Brazil to the US [9].

Sugar market is one of the highly protected markets in World Agricultural Trade (WTO). The international sugar trade is defined by preferential trade agreements in such a way that producing countries have the benefit of higher priced domestic markets of the EU or USA through preferential access which, consequently, makes the trade very important to the sugar sectors of many developing countries [19]. "The current policy of allowing efficient U.S. sugar growers to compete with foreign subsidies and trades provides a reliable supply of sugar at reasonable prices to American consumers. This program operates at no cost to American consumers" [1].

Specifications for sugar differ corresponding to the specific applications of the end user, with trade volumes and prices broadly categorized as raw or refined types of sugar [19].

Brazil, as the largest producing and exporting country in the world, is the
dominant player of the global trade in raw and refined sugars, accounting for 51% of global export trade in 2005, according to ISO data [14]. The Russian Federation, EU-25, USA, South Korea and Japan are the largest importing nations in the world, even if India also emerged as an important importer of raw sugar in 2004 and 2005 [19].

2.3 Past modelling efforts of the sugar market—literature review

The market for sugar, as a commodity, has aroused the curiosity of several researchers. This section reviews some of them, showing different facets of sugar market models.

Melo et al. in their research paper [5] applied a Mixture of Local Expert Models (MLEM) to forecast the daily and monthly prices of the Sugar No. 14 contract in the New York Board of Trade. It was a forecasting technique that performed data exploratory analysis and mathematical modeling simultaneously. In this article, given the data set points, they first divided them into clusters of points, next using several modeling techniques, they constructed models for each cluster and then the best model for each cluster was selected and called the Local Expert Model. Finally, they combined the outputs of all local expert models using a neural network called Gating Network. Furthermore, for comparison purposes, the same modeling techniques were also evaluated when acting as Global Experts, in other words, when using the entire data set without any clustering. However efficient this system of the MLEM technique might be, costs and results when one uses global experts based on a single technique should also be considered.

In their research paper, Rapsomanikis and Hallam [22] explored the possibility of nonlinear dynamic adjustment in the sugar-ethanol-oil link in Brazil, using the threshold vector error correction models (VECM) for testing for linearity in the adjustment of prices of each of those two commodities. The tests in that paper demonstrate the long run behaviour of sugar prices to be determined, not by ethanol prices but by oil prices. In addition, it appears that there was no evidence for threshold cointegration between sugar and ethanol prices; they were linearly cointegrated, with sugar being the dominant market. On the other hand, the results indicated that sugar price reacted in a successive
manner to oil changes but not through ethanol prices.

Similar conclusions regarding sugar-ethanol-oil were made in an other paper by Balcombe and Rapsomanikis [3] from which they examined the relationships between Brazilian sugar and ethanol prices and international oil prices using a Bayesian methodology. The existing VECM models were extended to all for adjustment with respect to a nonlinear transformation of disequilibrium errors. The Bayesian model averaging was used to estimate the elasticities of transmission from oil to sugar and ethanol. The model suggested that the long-run drivers of Brazilian sugar prices were oil prices and that there existed nonlinearities in the adjustment processes of sugar and ethanol to oil, but a linear adjustment between ethanol and sugar. In addition, it was found that there was an asymmetric adjustment between sugar and oil.

Given that China is a member of the WTO and is fast becoming a major sugar consumer, in their paper Wei Si and Xiu-Qing Wang [24] examined the extent to which the Chinese and the world sugar markets are integrated and how price fluctuations in the world market may affect China’s domestic sugar market. Using the Johansen co-integration method, the study examined if there exists a long-run co-integration relationship between main China’s domestic sugar markets and international sugar markets. Furthermore, the study used the error correction models (ECM) to analyze the likely short-run impacts of international sugar price fluctuations on China’s domestic sugar market. The study showed that there was long-run co-integration relationship between China’s domestic sugar markets, and between world sugar spot market and China’s domestic sugar market where the world sugar market price tends to lead the price changes in domestic sugar market. On the other hand, in the short run, changes in the world sugar price did not seem to have a direct effect on the sugar price in China’s domestic market.

The estimation of the demand for sugar encounters problems if the measurements of the price data for the substitute sweeteners for sugar are inaccurate. Considering the demand for beverage sugar and nonbeverage sugar in the US, the price data for substitute sweeteners for sugar contained measurement error. Therefore, Uri in his paper [27] elaborated two diagnostics to assess the effect on the estimated coefficients of the sugar demand relationship caused by this measurement error. These are, the Regression Coefficients Bounds and the Bias Correction Factor. Using the regression coefficients bounds diagnostic, Uri indicated a range which spanned the true price responsiveness of
consumers to changes in the price of sugar substitutes. In addition, the evaluation of the magnitude of the overestimation of the responsiveness of quantity of sugar substitutes demanded for changing the price of sugar was done by computing the bias correction factor. The results obtained suggested that, in the presence of measurement error in the data for the price of the sugar substitute, any conclusions or policy recommendations based on the estimated sugar demand relationships must be qualified.

In their paper, Angus Deaton and Guy Laroque [10] developed an idea to model the long-run behavior of the prices of primary commodities from Sir Arthur Lewis’ account in which the price of tropical production was held down by the existence of unlimited supplies of labor in poor countries. In their study paper, they assumed commodity supply to be infinitely elastic in the long run and the demand was linked to the level of world income and to the price of the commodity to make a price stationary around its supply price. In addition, commodity supply and world income were cointegrated. The model was fitted to long-run data for 6 commodities, sugar among them, over the years 1900 to 1987. They found out that despite the enormous growth of world income in the 20th century, the increases in the real commodity prices must wait for the elimination of poverty in the tropics, as predicted by Lewis. In contrast, this study was not able to provide any crucial evidence that would convert a skeptic to the Lewis story. Several difficulties occurred, namely, the reversion of prices to their long run base was very slow, such that however numerous was the data, they would not be able to provide clear results on either stationarity of prices or the cointegration of production and output, and the empirical evidence is never very clear in rejecting alternatives. Except from two commodities (cocoa and coffee) for which the Lewis model was most obviously consistent with the data, it was found that the model lacked statistical significance and the models’ fit came mainly from the univariate time-series representations that was not of big interest in this study context.

Another research paper by Pinto et al. [21] reviews modelling switching options using Mean-Reverting Commodity Price Models (MRCPM). By censored probability lattice, they implemented a precise and flexible framework for modeling a one factor mean-reverting process and thus, using a bivariate lattice, they extended that approach to a two-variable mean-reverting process and hence used the latter to value the switching option available to producers of two commodities, namely, sugar (a food commodity) and ethanol (an energy commodity), which can be chosen as output from one basic source, that is
sugarcane. The modelling results in that paper indicated that the switching option had significant value for the producer. The prices of these two commodities were approximated by a mean-reverting model and it appears that sugarcane-based ethanol producers benefit from a natural hedge based on the sugar market.

In [18], Nijman and Beetsma tested the empirical implications for commodity futures for the marginal process of prices of sugar futures of a simple pricing model. According to the pricing model, the futures price bias depends linearly on the conditional variance. Using the GARCH-M model from monthly as well as daily data, a significant impact of the conditional variance on the change in futures prices was obtained. These estimates implied contango in the futures market and a net hedging demand on the long side of the futures market. In contrast, the results suggested that the simple pricing model focuses on at least one important aspect of the pricing of Sugar futures: the risk premia depended on the time varying volatility.

To capture the volatility in the global food commodity prices, sugar prices among others, Onour and Sergi, in their paper [20], employed two competing models, the thin tailed normal distribution and the fat-tailed Student $t$-distribution models. Results were based on monthly data covering the period from October 1984 to September 2009. The performance of each was assessed with the use of predictive power of the volatility forecast and other goodness of fit measures. It was found that the $t$-distribution model outperforms the normal distribution model, revealing the evidence of leptokurtosis in the volatility of food commodity prices. This result implies that if such leptokurtic behavior was ignored when estimating the conditional volatility, then the standard option pricing formula of Black and Scholes used could lead into unreliable results when pricing the future option contracts, since it depended on the expected volatility parameter. In addition, this paper indicated that the volatility of the future commodity prices was mean reverting and exhibited the short memory behavior for many commodities, sugar among others. That is, the effect of shocks on the future conditional variance persisted only for short periods.

Several models, such as the ones discussed above, have been developed in order to simulate commodity prices, sugar prices among others, for different purposes such as forecasting, risk minimization or seeking for co-integration evidence with other commodities. Volatility of prices is one of the major sources of risks in commodity markets. This volatility can be modeled by stochastic models
of various forms. The main reason behind this volatility is the deregulation of many markets. Among different other reasons, traders' behavior and the presence of momentum may cause sudden price changes in financial markets. Like several others, the models discussed above seem to ignore these actions whose effects are too big to be taken as normal irregularities. In this study, US sugar prices are modeled using two models that take into account the irrational actions of traders, also called animal spirits and the moment effect of financial markets.

3 SUGAR PRICES

The data used in this paper represent a set of monthly US sugar prices observations from January, 1960 to February, 2012. Sugar prices, among other commodities, were found on the World Bank Commodity Price Data (Pink Sheet), in cents/kg, with monthly prices on the March 5, 2012 US dollars. These data can easily be accessed from the World Bank (Bloomberg) website (http://databank.worldbank.org).

This set of monthly observations constitutes a time series. Classical analysis of time series needs stationary data. Consequently, graphical representations such as time line, autocorrelation function, and partial autocorrelation function plots and histograms and some basic statistics such as mean, standard deviation, skewness and kurtosis have been used for exploratory analysis.

A time series is considered to be nonstationary if it has one or more of the following features: trend component, which defines a long-term movement in a time series, cyclical component, which describes any regular fluctuations, in other words, it is a non-seasonal component which varies in a recognizable cycle and seasonal component which refers to variations in time series representing fluctuations that are more or less stable period after period. After all these three components have been accounted for, what left is the irregular component. These features can be checked out using the time line plot (see Figure 2) of the time series. Therefore, for further analysis, any observed feature must be removed. However, since we use the concept of moving mean reversion in this study, stationarity in the presented models is not strictly required, and thus the data does not need to be transformed before applying the model.
Figure 2: Time line plot of the sugar price.

Figure 2 pictures the time line plot of sugar prices data set. There are enough observations in the data for further analysis. Obviously, the series is not stationary; both its mean value and variance change over time (some high spikes are observed), for instance there was a very high peak in November 1974 (in that time of the year, the world sugar prices peaked at a record 66 cents per pound [8]).

Next, the autocorrelation function (ACF) and partial autocorrelation function (PACF) are plotted (see Figure 3) to examine whether the series is random or if there are seasonalities in the series.

The ACF plot in Figure 3(a) shows a significant and slowly decaying autocorrelation within the series, hence, that reveals that the series is not stationary. And the PACF plot in Figure 3(b) shows that the majority of the coefficients lie within the confidence limits, with only few significant autocorrelations (note: the spike at lag 0 is not considered since it is just the autocorrelation of the series with itself and it is always equal to 1).

Normality in the series is also required in most classical time series approaches. Even though the models presented in this work do not require normality, we study the data distribution as the histogram will be one of the comparison measures of the model performance. A histogram (see Figure 4) can be used to check if the series is normally distributed. The histogram below shows that
the series is clearly not normally distributed.

There exist several ways of transforming a non-stationary series into a stationary one. The differencing or logarithmic differencing transformation can be applied to induce stationarity with respect to the mean value. In other
words, defining the price returns and logarithmic price returns, respectively, may transform the observed non-stationary series into a stationary one (with respect to the mean). Therefore, the respective time line plots of these transformations (see Figure 5), are used to check whether the series is stationary or not.

Figure 5: (a) Time line plot of the sugar price returns. (b) Time line plot of sugar price logarithmic returns.

Figure 5, presents a series moving around mean zero but still the variances are not constant and change over time.

In every way, the transformed data are less correlated than the original data, thus, they are likely to be used for further analysis. Therefore, similar features must be checked for returns and log returns to see whether it will be better to use the transformed data. Hence, Figures 6 and 7 present the ACF and PACF plots of the sugar price and log returns.

The ACF and PACF plots of the two transformations, that is the sugar price returns (see Figure 6) and the sugar price log returns (Figure 7), both present autocorrelations of significant but not very high level. The series seems more autocorrelated at lag 1, but not very significantly. The histograms below, of the sugar price returns (see Figure 8) and the sugar price log returns (Figure 9), both show that the data in the series are not normally distributed. They are closer to Gaussian distribution than the original data but they present a
Figure 6: (a) ACF plot of sugar price returns. (b) PACF plot of sugar price returns.

Figure 7: (a) ACF of sugar price logarithmic returns. (b) PACF of sugar price logarithmic returns.

very long peak above the normal curve.

The basic statistics of the data have been summarized in Table 1 which includes distribution parameters such as mean, standard deviation, skewness and
Figure 8: (a) Histogram of sugar price returns. (b) Normal distribution histogram of sugar price returns.

Figure 9: (a) Histogram of sugar price logarithmic returns. (b) Normal distribution histogram of sugar price logarithmic returns.

kurtosis, to give a better understanding on the behavior of the series distribution.

Table 1 indicates that the mean value is closer to 0 in the price log returns and
the standard deviation shows that the series in the logarithmic returns are less deviated. Moreover, when considering the skewness and kurtosis, one can see that the original sugar prices present less flat tail and shorter peak. On the other hand, the sugar price log returns present better distribution since the skewness is closer to 0 (i.e. closer to symmetry) and the kurtosis closer to 3 (i.e. less extreme observations) than the price returns. The skewness values are all greater than zero and positive, this indicates that the distributions are, in all the cases, asymmetric and right skewed.

4 MODELLING BACKGROUND

This section reviews mathematical models to be applied in this paper on the US sugar prices previously described. Starting by introducing stochastic differential equations and stochastic modelling, this section continues with a short description of the Jabłońska-Capasso-Morale and the Kalman Dynamics models.

4.1 Introduction to stochastic differential equations (SDEs) and stochastic modelling

Ordinary differential equations (ODEs) are considered a standard approach in modelling several engineering problems. In many applications, however, the experimentally measured trajectories of systems modeled by ODE do not in fact behave as predicted when there is uncertainty [11]. Hence it seems reasonable to modify ODE in a way that would allow to include the possibility of random effects disturbing the system. Therefore, these systems are often described in terms of the probability and must be described by means of a
stochastic model [16].

Let \((\Omega, U, P)\) be a probability space and let \(W(\cdot)\) be an \(m\)-dimensional Brownian motion and \(X_0\) an \(n\)-dimensional random variable which is independent of \(W(\cdot)\). Take

\[
F(t) := U(X_0, W(s)(0 \leq s \leq t))(t \geq 0),
\]

the \(\sigma\)-algebra generated by \(X_0\) and the history of the Wiener process up to (and including) time \(t\). Suppose \(T>0\) is given and

\[
b : \mathbb{R}^n \times [0, T] \to \mathbb{R}^n,
\]
\[
B : \mathbb{R}^n \times [0, T] \to M_{n \times m}
\]

are given functions. (These are not random variables.) We display the components of these functions by writing

\[
b = (b^1, b^2, \ldots, b^n), \quad B = \begin{pmatrix}
b^{11} & \cdots & b^{1m} \\
\vdots & \ddots & \vdots \\
b^{n1} & \cdots & b^{nm}
\end{pmatrix}
\]

We say that an \(\mathbb{R}^n\)-valued stochastic process \(X(\cdot)\) is a solution of the Itô stochastic differential equation (SDE)

\[
dX = b(X, t)dt + B(X, t)dW
\]
\[
X(0) = X_0
\]

for \(0 \leq t \leq T\), provided

(i) \(X(\cdot)\) is progressively measurable with respect to \(F(\cdot)\),
(ii) \(F := b(X, t) \in L^1_n(0, T)\),
(iii) \(G := B(X, t) \in L^2_{n \times m}(0, T)\),

and

(iv) \(X(t) = X_0 + \int_0^t b(X(s), s)ds + \int_0^t b(X(s), s)dW\) for all \(0 \leq t \leq T\).
4 MODELLING BACKGROUND

4.2 Jabłońska-Capasso-Morale (JCM) model

For decades researchers tried and failed in predicting the emergence and scale of financial crises in markets while trying to describe a model for the price in financial markets, using econometric models, in order to be capable of forecasting future prices [15]. For instance, finance scientists failed to identify any rational econometric crisis pioneers to the crash of September 2008 after the worldwide economic crisis. Later on, the main reasons of those failures, were found to be in irrational actions as what was described as trading biases also known as animal spirits; term proposed by John Maynard Keynes in 1936. These animal spirits, in financial markets, mean mainly worry and greed of the traders, which cause such reactions as herding, overconfidence or short-term thinking [15].

Any commodity price formation is always a process involving, not a single trader but, a group of traders. Consequently, an intuitive analogy would be to treat traders as individuals of a bigger population. This may be done by treating the price as the measure of their distance, that is, the range of prices can be considered as their space and the differences between the individuals’ bids as their distance [15].

A Capasso-Biandri system of stochastic differential equations, used for modelling animal populations or price herding, in which the movement of each particle $k$ in the total population of $N$ individuals is founded on the location of each individual with respect to the whole population $f(X_k^t)$, as well as on its local interaction with the nearest neighbors $h(k, X_i)$, is in general form:

$$dX_k^t(t) = [f(X_i^t) + h(k, X_i)]dt + \sigma dW^k(t), \quad (1)$$

for $k = 1, \ldots, N$.

Moreover, the target population density may be a cause of external environment, for instance, terrain type. Bigger distances may be observed on a perfectly flat surface where there is no risk of losing sight of neighbors whereas closer distances may be kept by the individuals on an irregular terrain. Therefore, the above equation (1) needed to be revised considering the environmental influence.

In financial market analogy, traders observe the prices of the market and, hence, tend to follow a general path. On the other hand, there is a physical
impossibility of two market participants to buy the same asset. Thus, that is how came up a motivation to use models proposed by Morale et al. [17] in mathematical biology into financial market modelling, by taking as a particle, each individual price path simulated from the model, whose movement may be driven by external information coming from the environment.

A fusion between mathematical biology and financial time series modeling has been proposed [6] [7], by studying a phenomenon called *price herding*. The proposed model was estimated and analyzed for 8 car brands over the years 1991 – 1999. The estimation was done with the use of maximum likelihood and maximum a posteriori methods. The obtained prediction results were very promising. Moreover, the authors admitted that the fitted model was not able to produce jumps in the process which are an important characteristic of the real prices. The fitted model was reproducing spikes, but not in the right time instances. Therefore, spikes were generated through separate jump processes dependent on the price level [15].

On the other hand, researchers have continuously been facing Efficient Market Hypothesis failure due to the presence of so called *momentum* in financial markets. The momentum effect is consistently too strong to be considered as a simple market anomaly. For instance, clients are attracted to invest more money where managers hold the most popular and valuable stocks resulting from the fact that they have been rewarded for their good performance and beating the market. That is, investors simply buy stocks just because their price has risen. Thus, the money again goes into the same investments and additionally boosts shares that are performing well already. In other words, the momentum effect consists of the fact that the individual prices deviate from the global tendency due to the traders pursuit of an extra profit [15].

In addition, physical analogy to the momentum phenomenon can be found in fluid dynamics; the Burgers' equation (see equation (2)), is hence applied.

\[
\frac{\partial u}{\partial t} + \alpha \frac{\partial u}{\partial x} + \beta \frac{\partial^2 u}{\partial x^2} = f(x, t) \tag{2}
\]

In terms of market dynamics, the following analogies came up:

- *u* specifies the price,
- \( f(x, t) \) stands for the fundamentals (of a periodic character),
- \( \alpha \frac{\partial^2 u}{\partial x^2} \) is the diffusion term related to the fact that the market tends to reach an equilibrium price,
MODELLING BACKGROUND

- $u_x$ describes the spread between any given day’s average and most common bids,
- $uu_x$ is the momentum term expressing traders’ movement towards higher prices.

Furthermore, the jump processes have been eliminated and the global interaction as Burgers’ type momentum component $h(k, X_t)$ was introduced in the model and thus, the model became

$$dX^k_t = [\gamma_t(X^*_t - X^k_t) + \theta_t(h(k, X_t) - X^k_t)]dt + \sigma_t dW^k_t, \quad (3)$$

Where $h(k, X_t) = M(X_t) \cdot [E(X_t) - M(X_t)]$ and $M(X)$ stands for the mode of a random variable $X$. And $\theta_t$ represents the strength of that global interaction at time $t$ and is allowed to be different from the mean field force $\gamma$. And,

- $X^k_t$ is the price of trader $k$ at time $t$,
- $X^*_t$ is the global price reversion level at time $t$,
- $\gamma_t$ is the mean reversion rate at time $t$,
- $X_t$ is the vector of all traders’ prices at time $t$,
- $W^k_t$ is the Wiener process value for trader $k$ at time $t$,
- $\sigma_t$ is the standard deviation for Wiener increment at time $t$.

Finally, the model was extended, with the goal of accounting for the main components of the Capasso-Bianchi population dynamics model, into a model that preserves all the observed trading biases.

$$dX^k_t = [\gamma_t(X^*_t - X^k_t) + \theta_t(h(k, X_t) - X^k_t) + \xi_t(g(k, X_t) - X^k_t)]dt + \sigma_t dW^k_t, \quad (4)$$

Where the new component $g(k, X_t)$ represents the maximally distant member of $k$ th trader’s neighborhood, formed by the closest $p\%$ of the population.

The previous model (see equation (4)) is known as the Jabłońska-Capasso-Morale (JCM) model and will be used to fit the data in this work.
4 MODELLING BACKGROUND

4.3 Kalman Dynamics (KD) model

The KD model is based on Jabłońska-Capasso-Morale (JCM) model and Variational Ensemble Kalman filter (VEnKF) algorithm, that incorporates the sample covariance calculated using the state estimate evolved from the previous time as the expectation [25]. Such an attempt was motivated by the fact that extreme volatility of electricity prices enforces market participants to hedge against sharp price changes [23], hence, "the designed models will make it possible to develop bidding strategies and negotiation skills to hedge producers and wholesale consumers and, simultaneously, to maximize their profit" [23].

A simulation procedure considers subsequent application of evolution and filtration operators to the state vector. The evolution operator is based on the JCM model for a group of traders in the market and the state vector stands for a set of bids from market participants that constitute a distribution [13]. The objective of synthetic model development is to simulate the real price, combining information from model output with information about correlations between decisions of the market participants accumulated in the covariance matrix of the state vector.

The KD model requires a set of preliminary detrended and deseasonalized prices together with the same set of data with removed spikes. Therefore, in this study, the KD model was fitted to preliminary detrended and deseasonalized original monthly data. The pure trading original prices are taken to estimate the $H$ days moving average and the standard deviation of the KD simulated price. To initialize the model we replicate the first value of the real price $N$ times, hence, we obtain the vector $u \in \mathbb{R}^{[N,1]}$, where $N$ is the number of market participants. Initial particles $s = (s_1, \ldots, s_{N_p})$ constitute the set of $N_p$ replicates of the main trajectory $u$.

Kalman Dynamics based algorithm

The algorithm is presented as in [23].
1. Estimate a mean reversion level from original data via Least Squares:

\[
\begin{pmatrix}
1 & x_{n-1} \\
1 & x_{n-2} \\
\vdots & \vdots \\
1 & x_{n-H}
\end{pmatrix}
\begin{pmatrix}
p_1 \\
p_2 \\
\vdots \\
p_H
\end{pmatrix} =
\begin{pmatrix}
dx_{n-1} \\
dx_{n-2} \\
\vdots \\
dx_{n-H}
\end{pmatrix}
\]  
(5)

where \(x_k, k = n - H, n - 1\) is the original pure trading price, \(dx_n = x_{n+1} - x_n\) is the price return. Equivalently, in vector-matrix representation the system (5) can be rewritten in the following way:

\[Xp = dx.\]  
(6)

We also compute the standard deviation for \(H\) days assimilation window. The latter two values (\(H\) days moving average and standard deviation) are the parameters required for JCM evolutionary operator.

2. Define the prior of the price increment \(\delta u^p = (\delta u^p_1, \ldots, \delta u^p_N)\), where \(N\) is the number of participants, from JCM model, which is employed as an evolutionary operator;

3. Propagate the particles \(\delta s = (\delta s_1, \ldots, \delta s_N)\) via the JCM operator, where each of the particles \(\delta s_i, i = 1, N_p\) represents price increments \(\delta u^p\) supplied with a small perturbation;

4. Perform lbfgs-minimization of the quadratic cost function with respect to the vector of price increments \(\delta u^*:\)

\[
l(\delta u^*|\delta u_{obs}) = \frac{1}{2}(\delta u^* - \delta u^p)^T C_p^{-1}(\delta u^* - \delta u^p) + \frac{1}{2}(\delta u_{obs} - K\delta u^p)^T C_{\varepsilon o}^{-1}(\delta u_{obs} - K\delta u^p),
\]  
(7)

where components of artificial observational vector \(\delta u_{obs}\) are composed by the particles \(\delta s\) in the following way:

\[(\delta u_{obs})_i = \sum_{j=1}^{N_p} \omega_{ij} \delta s_{ij}, \quad i = 1, N.\]  
(9)

Here \(\delta s_{ij}\) denotes \(j - \text{th}\) component of the \(i - \text{th}\) vector-particle \(\delta s_i\), coefficients \(\omega_{ij}\) are specified as Gaussian weights \(\omega_{ij} = f(s_{ij}|u_i, \sigma)\), standard deviation \(\sigma\) is model parameter, \(f\) stands for the normal distribution probability density function; observational operator \(K\) is an identity matrix, the inverse of observational vector error covariance \(C_{\varepsilon o}^{-1} = c_o I_N\) is
represented by a diagonal matrix with scalar model parameter $c_0$ as a
diagonal entity.

The inverse of the prior state vector $\delta u^*$ covariance is obtained from the
Sherman-Morrison-Woodbury (SMW) formula:

$$
C_p^{-1} = \left(\bar{X}X^T + C_{\varepsilon_m}\right)^{-1} = 
C_{\varepsilon_m}^{-1} - C_{\varepsilon_m}^{-1}\bar{X}\left(I + \bar{X}C_{\varepsilon_m}^{-1}X^T\right)^{-1}X^TC_{\varepsilon_m}^{-1}
$$

where model error covariance $C_{\varepsilon_m}^{-1} = c_m I_N$ is diagonal, and $c_m$ is the KD
model parameter. We calculate sample covariance $\bar{X}$ in the following
way (see [25]):

$$
\bar{X} = ((\delta s_1 - \delta u^p), (\delta s_2 - \delta u^p), \ldots, (\delta s_{N_p} - \delta u^p)) / \sqrt{N_p - 1}.
$$

5. Sample new ensemble $\delta s^* \sim N(\delta u^*, C_{est})$ using low-storage represen-
tation of the covariance estimate $C_{est}$;

6. Increment the main trajectory $u^* = u + \delta u^*$ and the particles $s^*_i = s_i + \delta s^*_i, i = 1, N_p$;

7. $s \rightarrow s^*, u \rightarrow u^*$ and go to step 1.

Where,

- $(\gamma, \theta, \xi)$ – JCM evolutionary operator parameters;
- $N_p$ – number of particles employed for filtration;
- $H$ – the length of assimilation window for computing moving average
and standard deviation;

- $c_0, c_m$ – the inverse values of observation and model error variances, cor-
respondingly. These coefficients define how strongly we trust the observ-
ations and the model. When the model error variances are small in
comparison to the observational error variances, decisions of the traders
are affected by the changed expectations of the surrounding agents more
than by the rational reasons based on the previous price dynamics.

- $\sigma$ – observational operator parameter; the bigger $\sigma$ is, the more deviant
are the expectations, taken into account by market agents.
5 MODELLING RESULTS

In this section, the model fitting results are provided. The two models, the JCM and the KD models, described in the previous section, are applied in order to simulate the given set of data discussed in Section 3. Each model is first applied separately, comparing the simulated to the original data and then, a comparison of the two simulations is done. In addition, ACFs, PACFs, histograms and basic statistical properties, such as the mean, standard deviation, skewness and kurtosis for both models and from the 5th to the 7th moments for the JCM model, are presented to acquire the models’ performance.

5.1 JCM model’s results

Starting with the JCM model fitting, the following results are obtained. Figure 10 captures the time line plot of the original data (in blue) vs. simulated data (in red) for comparison. Obviously, the simulated prices trajectory tries to follow the original data but fail at some points in the observations. Despite upward spikes observed in the simulated prices which are not signified in original prices, the variations of the two series are comparable even though they seem to be higher in the original data.

Figure 11 presents the original prices’ histogram (in blue) plotted against the simulated prices’ curve (in red) for a better view of how the two series of sugar prices (original and simulated prices) are distributed. It is observed that they both present the same features of non-normality but the original prices show higher peaks and thus are more positively skewed than the simulated prices. These observations may also be supported by the statistical values presented in Table 2.

The autocorrelogram of the original sugar prices (see Figure 12) and of the simulated prices (see Figure 13) are plotted to capture the resemblance of their corresponding ACFs. Both the ACFs illustrate positive as well as negative autocorrelations within the data.

The PACFs plots presented in Figures 14 and 15 of the original and simulated prices, respectively, show a highly significant spike at lag 1 and the values at most of the other lags lie within the confidence limits, with some non-significant
5 MODELLING RESULTS

Figure 10: Time line plot for original and simulated sugar prices with Jabłońska-Capasso-Morale (JCM) model.

![Time line plot for original and simulated sugar prices with JCM model.](image)

Figure 11: Histogram of original vs simulated sugar prices with JCM model.

![Histogram of original vs simulated sugar prices with JCM model.](image)

values below the limits for both prices, but with some non-significant values above the limits for the original prices.

Basic statistical properties of both the original and simulated sugar prices are summarized in Table 2 for statistical comparison. Clearly, the JCM model
simulated prices give generally lower statistical values than the original prices but with no apparent big difference except for the 7th moment.

Resulting from the reasonable resemblance observed in the above figures and the numerical values of the basic statistics, one may say that the computed JCM model passably captures the dynamics of the presented US sugar prices.
5.2 KD model’s results

In this section, the results of the Kalman Dynamics (KD) model fitting are as follows. Figure 16 presents the time line plot of the original data (in green) versus the simulated data (in blue) with the KD model for comparison. Clearly, the simulated prices trajectory tries to follow the original data. Both prices
present jumps and significant upward spikes. Although they are higher in the original prices, variations of the two series may be comparable.

Figure 16: Time line plot for original and simulated sugar prices (Kalman Dynamics model).

The histogram in Figure 17 of the original prices’ histogram (in blue) plotted against the simulated prices’ curve (in red) gives a look at the distribution of the two series of sugar prices (original and simulated prices). Non-normality is observed in both series, the original prices show higher peak and are more

Table 2: Statistical properties of real prices and simulated sugar prices (JCM model)

<table>
<thead>
<tr>
<th></th>
<th>Real</th>
<th>JCM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>39.7225</td>
<td>39.1085</td>
</tr>
<tr>
<td>St.Dev</td>
<td>18.3771</td>
<td>17.3166</td>
</tr>
<tr>
<td>Skewness</td>
<td>0.3730</td>
<td>-0.0459</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>3.6614</td>
<td>2.4574</td>
</tr>
<tr>
<td>5th moment</td>
<td>0.3721</td>
<td>-0.0458</td>
</tr>
<tr>
<td>6th moment</td>
<td>3.6498</td>
<td>2.4495</td>
</tr>
<tr>
<td>7th moment</td>
<td>7.3765</td>
<td>1.6831</td>
</tr>
</tbody>
</table>
positively skewed than the simulated prices. Basic statistical values presented in Table 3 may support these observations.

Figure 17: Histograms for original and simulated sugar prices (Kalman Dynamics model).

To seize the resemblance between the ACFs for the original and simulated sugar prices with the KD model, consider Figures 18 and 19, respectively. They both show highly significant positive autocorrelations but also non-negligible consistent negative autocorrelations.

The PACFs of the original and simulated prices are presented in Figures 20 and 21, respectively. They both show an evidence of a highly significant spike at the first lag and the values of the remaining lags lie within the confidence limits, with some less significant values below and above the confidence limits for both prices.

The basic statistics of both original and simulated sugar prices are summarized in Table 3 for statistical comparison. Obviously, the KD model simulated prices give remarkably lower statistical values than the original prices. The mean of the simulated prices is negative whereas the original prices give a positive mean, approximately equal to zero.

Trying to induce stationarity (with respect to the mean) on both the original as well as the simulated prices, price logarithmic returns are calculated. The
Figure 18: ACF for original sugar prices (Kalman Dynamics model).

Figure 19: ACF for simulated sugar prices (Kalman Dynamics model).

series oscillate around mean zero at both price log returns (see Figures 22 and 23).

The ACFs for both the real and simulated price log returns (see Figure 24) show no apparent significant autocorrelations.
The PACFs for both real and simulated price log returns presented in Figure 25 show no apparent significant values at any lag.

The Kalman Dynamics (KD) model observations show general resemblance with the real prices as shown in the above figures and the numerical values of the basic statistics. Hence, the computed KD model acceptably captures the
Figure 22: Time line plot for real sugar price returns (Kalman Dynamics model).

dynamics of the given US sugar prices.

5.3 Comparison of model performance

Finally, the two simulated prices are compared for model performance. A comparison of the results obtained from the two models (JCM and KD models) is performed and furthermore, their respective RMSE are compared as well.

Figure 26 shows the time line plots of both simulated prices (by KD and JCM models) compared to the real prices. It is observed that the two simulated models try to follow the original prices and it is not easy to tell at this stage which model captures the sugar dynamics better.

Table 3: Statistical properties of real and simulated prices (KD model)

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>St.Dev</th>
<th>Skewness</th>
<th>Kurtosis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Real</td>
<td>0.0011</td>
<td>12.0730</td>
<td>2.5592</td>
<td>14.7093</td>
</tr>
<tr>
<td>KD</td>
<td>-0.7741</td>
<td>10.0452</td>
<td>1.9609</td>
<td>10.1202</td>
</tr>
</tbody>
</table>
Figure 23: Time line plot for sugar price returns (Kalman Dynamics model).

Figure 24: ACFs for real and simulated sugar price returns (Kalman Dynamics model).

The histograms of the real and the two simulated prices are presented in Figure 27 for comparison. It is observed that they all present similar non-normality features. They all have a very high peak and a positive skewness. The numerical values of the basics statistics in Table 4 may support these observations. But, once again, it is difficult to distinguish which model is the best to capture...
Figure 25: PACFs for real and simulated sugar price returns (Kalman Dynamics model).

Figure 26: Time line plot for original and simulated prices (KD and JCM models).

the dynamics of the sugar prices between the KD and JCM models at this stage.

The ACFs the real and the two simulated prices are plotted in Figure 28. Ob-
Figure 27: Histograms for original and simulated prices (KD and JCM models).

Apparently, the simulated prices from the two models show the same features as the original prices. In addition to the presence of positively high autocorrelations as well as negative autocorrelations, all the models also have practically similar spikes disposition.

Figure 28: ACFs for original and simulated prices (KD and JCM models).

The PACFs of the real and the two simulated prices are presented in Figure
29. They all show a highly significant spike at lag 1, some other less significant spikes lying outside the significance limits and others lying within the limits.

A comparison of the basic statistical properties is presented in Table 4. All the series have very high positive skewness values and their kurtosis values are much greater than 3. These values support the earlier observations on histograms in Figure 27. The mean and the standard deviations of the simulated prices by the JCM model have closer numerical values to those of the original prices than the simulated prices by the KD model whereas skewness and kurtosis values of the simulated prices by the KD model are closest to the original prices.

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>St.Dev</th>
<th>Skewness</th>
<th>Kurtosis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Real</td>
<td>0.0011</td>
<td>12.0730</td>
<td>2.5592</td>
<td>14.7093</td>
</tr>
<tr>
<td>KD</td>
<td>-0.7741</td>
<td>10.0452</td>
<td>1.9609</td>
<td>10.1202</td>
</tr>
<tr>
<td>JCM</td>
<td>-0.7294</td>
<td>12.1745</td>
<td>1.8350</td>
<td>8.9309</td>
</tr>
</tbody>
</table>

The returns of the original prices and the simulated prices (with KD and JCM models) are computed. Their time line are hence plotted in Figure 30; real
price returns in green, simulated price returns by the KD model in red and simulated price returns by the JCM model in blue. All the series show similar features and it is difficult to distinguish which model is the best to capture the dynamics of the sugar prices between the KD and JCM models regarding these plots.

Figure 30: Time line plot for real and simulated return prices (KD and JCM model).

The ACFs of all the series in Figure 31 present same features, no evidence of significant autocorrelations in all the series and it is not possible to distinguish which of the simulated models has more similarities to the real prices.

The PACFs in Figure 32 show no apparent significant values at any lag. The PACFs of both simulated prices show no big differences when comparing them to the real prices.

In addition, the Root Mean Square Errors (RMSE), also called Root Mean Square Deviations (RMSD) are applied to the simulated prices by the two models in order to analyze their accuracy. In other words, the RMSE helps us to see how close or how far are the simulated prices to the real prices. Thus, the closer is the simulated prices to the real prices, the better is the model. The RMSE is the measure of the differences between the real prices and the
simulated prices and is defined as in follows:

$$RMSE = \sqrt{\frac{\sum_{k=1}^{n}(x_{k}^{mo} - x_{k}^{true})^2}{n}}$$ (12)

Therefore, the time lines of the RMSE of the simulated prices by both models are plotted in Figure 33. The two time line plots show no big difference, although the RMSE of the simulated prices by the JCM model seem to produce

Figure 31: ACFs for real and simulated return prices (KD and JCM model).

Figure 32: PACFs for real and simulated return prices (KD and JCM model).
a bit higher spikes than those of the simulated prices by the KD model, since JCM RMSE crosses over 4, whereas KD RMSE stays below 4.

Figure 33: RMSE time line plots for simulated prices (KD and JCM models).

The RMSE histograms of the simulated prices by both models (see Figure 32) show no apparent differences between them.

Figure 34: RMSE histograms for simulated prices (KD and JCM models).

Furthermore, Table 5 provides numerical values of the basic statistics of the
RMSE of the simulated prices by the two models, such as, the mean, the standard deviation and the minimum and the maximum of the RMSE. Obviously, the simulated prices by the JCM model have a bit higher RMSE than the simulated prices by the KD model. Their mean and standard deviation are higher and they lie in a bigger interval since they have the smallest minimum and biggest maximum values.

Table 5: RMSE statistical properties for KD and JCM simulated prices.

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Min</th>
<th>Max</th>
<th>St.dev</th>
</tr>
</thead>
<tbody>
<tr>
<td>KD</td>
<td>0.1439</td>
<td>0.4241</td>
<td>3.8751</td>
<td>0.3489</td>
</tr>
<tr>
<td>JCM</td>
<td>0.1614</td>
<td>3.771e-5</td>
<td>4.0743</td>
<td>0.4126</td>
</tr>
</tbody>
</table>

6 RESULTS SUMMARY AND DISCUSSION

In this section, the modelling results obtained in section 5 are summarized and discussed. The results were obtained by executing the MATLAB codes of the two models, the JCM and the KD models, in producing figures such as, time line plots, ACFs and PACFs, histograms and summarized tables of the basic statistics.

First, playing with the JCM model parameters, the closest model to the original monthly sugar prices was obtained at a window of 12 and an ensemble size of 400. (Note: it was observed that the change in the ensemble size produces no difference in the model which is a desirable result). Figures 10-11 and Table 2 of basic statistical properties for the JCM model were provided to give a wide comparison of the resemblance of the real prices and the simulated prices behaviors, which lead to a conclusion that the JCM model was fairly able to capture the dynamics of the sugar prices.

Next, filtering the sugar price series, the KD model provided a model moderately able to follow the original sugar prices. The simulated model behavior was compared to the real prices model as shown by Figures 16-25 and Table 3 of basic statics to lead to such a conclusion.

Finally, the two models, the JCM and the KD models, were compared to analyze which model is better to acquire the dynamics of the sugar prices. Figures
7 CONCLUSION

26-34 and Tables 4 and 5 of the basic statistics are used for the comparison of the two models and their RMSE behaviors and resemblance. Resulting from the later comparison tools, we may conclude that the two models have approximately the same ability to capture the dynamics of sugar prices. However, the three series, the real sugar prices and the simulated prices by the two different models, present differences, no matter how small they seem and the RMSE statistical values, as measure of closeness to the original prices, indicate the simulated prices by the KD model seem to be slightly closer to the real prices than the simulated prices by the JCM model.

7 CONCLUSION

In this study, the Jabloński-Capasso-Morale (JCM) and the Kalman Dynamics (KD) models have been used to simulate the US monthly sugar prices in order to verify whether the ideas of the two models were true to seize the dynamics of the sugar prices.

At first, this study provided a wide introduction of the sugar market in general and particularly, the US sugar market and a deep description of the set of data used in this study was given through several graphical presentations and basic statistical properties. Since the series was found to be non-normally distributed, and hence non-stationary, an attempt on inducing stationary to it was done by applying differences and logarithmic differences to the sugar prices but, the later attempt failed. Moreover, the concept of moving mean reversion used in this study does not require the series to be stationary.

Furthermore, after presenting a background of the two models, they were applied to first simulate the sugar prices separately, comparing each model’s resemblance to the real prices and then the two models behavior have been compared for their performances. In summary, the two models, the JCM and the KD models, were acceptably able to follow the change in sugar prices with presence of non-significant differences. In addition, the KD model simulated sugar prices were observed to be closer to the original sugar prices than do the JCM sugar simulated prices.
References


REFERENCES


REFERENCES

the 11th International Conference on Real Options, Berkeley, CA, June 6-9, 2007. California, USA.


### List of Tables

1. Basic Statistics of Sugar Prices ........................................ 21
2. Statistical properties of real prices and simulated sugar prices (JCM model) ................................. 33
3. Statistical properties of real and simulated prices (KD model) ........................................ 37
4. Statistical properties of real and simulated prices (JCM and KD models) ........................................ 41
5. RMSE statistical properties for KD and JCM simulated prices ........................................ 45
List of Figures

1. Global sugar production vs. consumption. .................................. 10
2. Time line plot of the sugar price. .............................................. 16
3. (a) ACF and (b) PACF plots of sugar prices. ............................... 17
4. (a) Histogram of sugar price. (b) Normal distribution histogram of sugar price. ......................................................... 17
5. (a) Time line plot of the sugar price returns. (b) Time line plot of sugar price logarithmic returns. ................................................. 18
6. (a) ACF plot of sugar price returns. (b) PACF plot of sugar price returns. ................................................................. 19
7. (a) ACF of sugar price logarithmic returns. (b) PACF of sugar price logarithmic returns. ............................................................ 19
8. (a) Histogram of sugar price returns. (b) Normal distribution histogram of sugar price returns. ................................................. 20
9. (a) Histogram of sugar price logarithmic returns. (b) Normal distribution histogram of sugar price logarithmic returns. .............. 20
10. Time line plot for original and simulated sugar prices with Jabłońska-Capasso-Morale (JCM) model. ............................................ 30
11. Histogram of original vs simulated sugar prices with JCM model. 30
12. ACF for original sugar prices. ..................................................... 31
13. ACF for simulated sugar price with JCM model. ......................... 31
14. PACF for original sugar prices. .................................................. 32
15. PACF for simulated sugar prices with JCM model. ...................... 32
16. Time line plot for original and simulated sugar prices (Kalman Dynamics model). ............................................................... 33
LIST OF FIGURES

17 Histograms for original and simulated sugar prices (Kalman Dynamics model). ........................................ 34
18 ACF for original sugar prices (Kalman Dynamics model). .... 35
19 ACF for simulated sugar prices (Kalman Dynamics model). ... 35
20 PACF for original sugar prices (Kalman Dynamics model). .... 36
21 PACF for simulated sugar prices (Kalman Dynamics model). ... 36
22 Time line plot for real sugar price returns (Kalman Dynamics model). ..................................................... 37
23 Time line plot for sugar price returns (Kalman Dynamics model). 38
24 ACFs for real and simulated sugar price returns (Kalman Dynamics model). ............................................. 38
25 PACFs for real and simulated sugar price returns (Kalman Dynamics model). .............................................. 39
26 Time line plot for original and simulated prices (KD and JCM models). .................................................... 39
27 Histograms for original and simulated prices (KD and JCM models). ......................................................... 40
28 ACFs for original and simulated prices (KD and JCM models). ... 40
29 PACFs for original and simulated prices (KD and JCM models). 41
30 Time line plot for real and simulated return prices (KD and JCM model). .................................................... 42
31 ACFs for real and simulated return prices (KD and JCM model). 43
32 PACFs for real and simulated return prices (KD and JCM model). 43
33 RMSE time line plots for simulated prices (KD and JCM models). 44
34 RMSE histograms for simulated prices (KD and JCM models). ... 44