Forecasting Sales Using Genetic Algorithms
La prévision des ventes en utilisant les Algorithmes Génétiques

Omar MOUFFOK[1*], Youcef SOUAR[2]

[1] *Doctor, MIFMA Laboratory – Tlemcen, Algeria*, mo-om22@hotmail.fr
[2] *Professor, Taher Moulay University – Saida, Algeria*, syoucef12@yahoo.fr

Abstract:

Genetic Algorithms is an Artificial intelligence technique that can be used as a quantitative method to solve various types of problems including forecasting. Usually, we forecast phenomena behaviors through time series analyses using Box-Jenkins methodology. Thus, this study aims to make a combination between these methods and apply it on a time series of plastic manufacturer monthly sales. The results obtained showed that Genetic Algorithms is an efficient method for sales forecasting; besides it has several advantages concerning characteristics and application.

Keywords: genetic algorithms, sales forecasting, quantitative methods, time series, optimization.
(JEL) Classification : C01, C32, C53, C61.

Résumé :

Les Algorithmes Génétiques est une technique d'intelligence artificielle que peut être utilisé comme méthode quantitative dans plusieurs domaines. Les prévisions des comportements des phénomènes s'effectuent habituellement par la méthodologie de Box-Jenkins en analysant les séries temporelles, le but de cette étude est de faire une combinaison entre ces méthodes et l'appliquer sur une série temporelle des ventes mensuelles d'un fabriquant de plastique. Les résultats obtenues ont montré que les Algorithmes Génétiques est une méthode efficace pour la prévision des ventes, ainsi qu'elle a plusieurs avantages concernant les caractéristiques et l'application.

Mot-clé : algorithmes génétique, prévision des ventes, méthodes quantitatives, séries temporelles, optimisation.
(JEL) Classification : C01, C32, C53, C61.

---

[*]Corresponding author: **Omar MOUFFOK**, Email: mo-om22@hotmail.fr

1. INTRODUCTION :

Due to various changes in the several economical phenomena that affect management and decision making, it is very essential to study these changes and to have some ideas about possible future events. Thus, managers always need to try to forecast the behaviors related to their activities.

Sales forecasting is one of the most important processes for company success, because sales management represents the revenues and have a huge financial impact on the other management activities. With the help of future sales estimation, managers can determine production policies and strategies, as well as attain the optimization of resources exploitation.

Many factors influence sales management, the time is a basic factor. So, time series that contains values overtime during a determined while could be analyzed by mathematical techniques and specific methodology to forecast future values.

There are many methods that analyze time series to find the best forecasting model, among these methods Box-Jenkins methodology is considered as the most common one, which is based on some mathematical and statistical aspects through several analytical processes under specific conditions.

Even so, sometimes these methods are not enough in some complicated cases, and then we need techniques that are more efficient. The development of computing science and artificial intelligence has discovered new methods like those that simulate nature like Genetic Algorithms.

Genetic Algorithms use some natural selection principals and produce new search points depending on artificial intelligence using statistical tools. It has been applied to solve various type of problem in several domains including econometrics and forecasting.

1.1. Research Problematic:

The main question to answer in this study is: how could be possible to use Genetic Algorithms for forecasting sales?

1.2. Research Aims:

We aims through this study to:

- Recognize Box-Jenkins methodology as the usual forecasting method.
- Recognize Genetic Algorithms and their concepts.
- Apply Genetic Algorithms in a practical study about monthly sales forecasting of plastic manufacturer.
- Know how much is Genetic algorithms efficient in sales forecasting.

2. BOX-JENKINS METHODOLOGY:

Box-Jenkins methodology contains specific statistical aspects used for studying time series through several analytical operations, which allow to determine the best ARIMA model for a given phenomenon forecasting.

2.1. Stationarity Conditions:

- The mean must be independent from time and constant: $E(y_t) = E(y_{t+m}) = \mu$.
- The variance must be independent from time and limited: $var(y_t) < \infty$.
- The covariance must be independent from time: $cov(y_t, y_{t+k}) = E[(y_t-\mu)((y_{t+k}-\mu)] = \gamma_k$.

Then stationarity could be determined using the following tests:

2.1.1. Autocorrelation Function Analyses (ρ_k):(Bourbonnais, 2011)

We test the following hypothesis:
$$\begin{cases} H0 : \rho_k = 0 \\ H1 : \rho_k \neq 0 \end{cases}$$

If $\rho_k = \frac{cov(y_t y_{t-k})}{\sigma_{y_t}\sigma_{y_{t-k}}}$ is significantly unequal to 0 ($\rho_k = 0 \pm t^{\alpha/2}\frac{1}{\sqrt{n}}$), there is a stationarity.

These tests could be easier by using correlograms.

Fig. (01): Example of a correlogram

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	-0.023	-0.023	0.1187	0.730
		2	0.008	0.008	0.1334	0.935
		3	0.035	0.035	0.4059	0.939
		4	-0.040	-0.038	0.7625	0.943
		5	-0.024	-0.027	0.8957	0.971
		6	-0.096	-0.098	2.9818	0.811
		7	0.011	0.009	3.0085	0.884
		8	-0.048	-0.047	3.5457	0.896
		9	0.035	0.038	3.8266	0.922
		10	-0.018	-0.026	3.9039	0.952
		11	0.062	0.061	4.7971	0.941
		12	0.061	0.049	5.6712	0.932
		13	0.122	0.132	9.2099	0.757
		14	-0.126	-0.139	12.965	0.529
		15	0.006	0.010	12.972	0.604
		16	0.075	0.067	14.330	0.574
		17	-0.085	-0.045	16.087	0.518
		18	-0.018	-0.025	16.166	0.581
		19	-0.020	0.001	16.261	0.640
		20	0.018	0.000	16.337	0.695

Source: By the researchers.

2.1.2. Box-Pierce Statistic :(Box & Pierce, 1970)

We test the following hypothesis: $\begin{cases} \rho_1 = \rho_2 = \cdots = \rho_h = 0 \\ the\ is\ \text{at least } one\ \rho_k\ \text{unequal to 0} \end{cases}$

There is a Stationarity if: $Q = n\sum_{k=1}^{h} \rho_k^2$ is higher than χ^2 (h).

2.2. The No-Stationarity Case: (Bourbonnais, 2011)

In case of no stationarity there are two processes to study the series: Trend Stationary (TS) and Deferency Stationary (DS).

2.2.1. Trend Stationary:

Where: $x_t = f_t + \varepsilon_t$

Linear model: $x_t = a_0 + a_1 t + \varepsilon_t$

- $E(x_t) = a_0 + a_1 t + E(\varepsilon_t) = a_0 + a_1 t$
- $var(x_t) = 0 + var(\varepsilon_t) = \sigma_\varepsilon^2$
- $cov(x_t, x_{t'}) = 0$

E (x_t) depends on time.

2.2.2. Defrency Stationary:

Where: $(1 - D)x_t = \beta + \varepsilon_t \Leftrightarrow x_t = x_{t-1} + \beta + \varepsilon_t$

- If $\beta = 0$: DS process without a constant.

$$x_t = x_{t-1} + \varepsilon_t \Leftrightarrow x_t = x_0 + \sum_{i=1}^{t} \varepsilon_i$$

- $E(x_t) = x_t$
- $var(x_t) = t\sigma_\varepsilon^2$
- $cov(x_t, x_{t'}) = \sigma_\varepsilon^2 \times \min(t, t')$

Var (x_t) depends on time.

- If $\beta \neq 0$: DS process with a constant.

$$x_t = x_{t-1} + \beta + \varepsilon_t \Leftrightarrow x_t = x_0 + \beta t + \sum_{i=1}^{t} \varepsilon_i$$

- $E(x_t) = x_t + \beta t$
- $var(x_t) = t\sigma_\varepsilon^2$
- $cov(x_t, x_{t'}) = \sigma_\varepsilon^2 \times \min(t, t')$

Both mean and variance depend on time.

2.3. Unit root Tests:

To determine the type of no-stationarity process we must use one of Unit Root tests as follow:
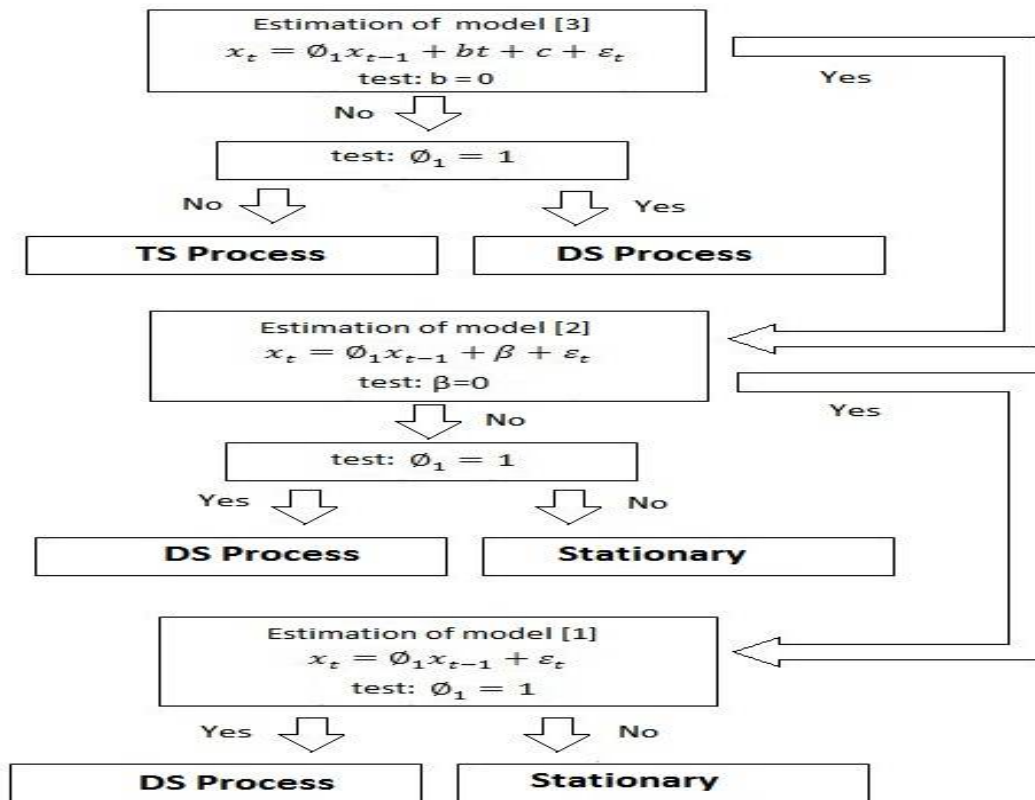
2.3.1. Dickey-Fuller Tests: (Dickey & Fuller, 1979)

We estimate the following models using Ordinary Least Squares (OLS):

[1] $x_t = \emptyset_1 x_{t-1} + \varepsilon_t$ Autoregressive model order 1

[2] $x_t = \emptyset_1 x_{t-1} + \beta + \varepsilon_t$ Autoregressive model with a constant

[3] $x_t = \emptyset_1 x_{t-1} + bt + c + \varepsilon_t$ Autoregressive model with a trend

Then we determine the process type by testing the several coefficients significance comparing to Dickey-Fuller tables values according to figure (2).

Fig. (02): Unit Root tests

Source: Bourbonnais, 2011, p. 249.

2.3.2. Augmented Dickey-Fuller Tests: (*Dickey & Fuller, 1981*)

This test takes auto regression of errors into account, so with the same principles of simple Dickey-Fuller test, we use the following models instead:

[4] $\Delta x_t = \rho x_{t-1} - \sum_{j=2}^{p} \emptyset_j x_{t-j} + \varepsilon_t$

[5] $\Delta x_t = \rho x_{t-1} - \sum_{j=2}^{p} \emptyset_j x_{t-j} + \beta + \varepsilon_t$

[6] $\Delta x_t = \rho x_{t-1} - \sum_{j=2}^{p} \emptyset_j x_{t-j} + bt + c + \varepsilon_t$

2.3.3. Phillips-Perron Tests: (*Phillips & Perron, 1988*)

Those tests takeheteroce dasticity of errors into account by following these steps:

1) Estimating Dickey-Fuller models by OLS.

2) Estimating the variance: $\hat{\sigma}^2 = \frac{1}{n}\sum_{t=1}^{n} e_t^2$

3) Estimating a corrective factor :

$$s_t^2 = \hat{\sigma}^2 + 2\sum_{i=1}^{l}(1 - \frac{i}{l+1})\frac{1}{n}\sum_{t=i+1}^{n} e_t\, e_{t-i} l \approx 4(n/100)^{2/9}$$

4) Calculate the statistic: $t_{\emptyset_1}^s = \sqrt{k} \times \frac{(\hat{\emptyset}_1 - 1)}{\hat{\sigma}_{\emptyset_1}} + \frac{n(k-1)\hat{\sigma}_{\hat{\emptyset}_1}}{\sqrt{k}}$ with: $k = \frac{\hat{\sigma}^2}{s_t^2}$

We compare this statistic to critical McKinnon table values.

2.4. ARIMA Models :(*Bourbonnais, 2011*)

One of the various ARIMA models could fix the no-stationarity of the series.

2.4.1. Auto Regressive Models (AR):

AR (1): $x_t = \theta_1 x_{t-1} + \varepsilon_t$

AR (2): $x_t = \theta_1 y_{t-1} + \theta_2 x_{t-2} + \varepsilon_t$

AR (p): $x_t = \theta_1 x_{t-1} + \theta_2 x_{t-2} + \cdots + \theta_p x_{t-p} + \varepsilon_t$

$\rho_k = \rho^k$

2.4.2. Moving Average Models (MA):

MA(1) : $x_t = \varepsilon_t - \alpha_1 \varepsilon_{t-1}$

MA(2) : $x_t = \varepsilon_t - \alpha_1 \varepsilon_{t-1} - \alpha_2 \varepsilon_{t-2}$

MA(q) : $x_t = \varepsilon_t - \alpha_1 \varepsilon_{t-1} - \alpha_2 \varepsilon_{t-2} - \cdots - \alpha_q \varepsilon_{t-q}$

$$\rho_k = \frac{\sum_{i=0}^{i=q-k} \alpha_i \alpha_{i+k}}{\sum_{i=0}^{i=q} \alpha_i^2}$$

2.4.3. ARMA Models:

ARMA (p.q.) :

$$x_t = \theta_1 x_{t-1} + \theta_2 x_{t-2} + \cdots + \theta_p x_{t-p} + \varepsilon_t - \alpha_1 \varepsilon_{t-1} - \alpha_2 \varepsilon_{t-2} - \cdots - \alpha_q \varepsilon_{t-q}$$

In case of model with constant, we add: $\mu = E(x_t) \times (1 - \theta_1 - \theta_2 - \cdots - \theta_p)$

2.4.4. Integrated ARMA Models (ARIMA):

ARIMA (p,d,q) : $\Delta^d x_t = x_t - x_{t-1} - x_{t-2} - \cdots - x_{t-d}$

ARIMA (p, 1,q): $x_t = x_{t-1} + \Delta x_t$

$$\Delta x_t = \theta_1 \Delta x_{t-1} + \theta_2 \Delta x_{t-2} + \cdots + \theta_p \Delta x_{t-p} + \varepsilon_t - \alpha_1 \varepsilon_{t-1} - \cdots - \alpha_q \varepsilon_{t-q}$$

2.5. Box-Jenkins Methodology Steps :(*Box & Jenkins, 1976*)

Box-Jenkins methodology goes through several analytical steps to find the most suitable model of forecasting.

2.5.1. Model Identification

We analyze the series characteristics and determine whether it is: stationary, TS or DS. TS series should be fixed using ARMA (p, q) models, whereas DS series should be fixed using ARIMA (p,d,q). Thus, we try to determine p and q orders with the help of correlograms.

2.5.2. Model Estimation:

Then we estimate several models coefficients using mathematical and statistical techniques such as: OLS for AR (p) and Likelihood maximization algorithms for MA (q). (*Tsay, 2002*)

2.5.3. Model Diagnostic Checking:

After models estimation we check its quality through coefficients significance, White Noise, Normality and Homoscedasticity tests. Then, we choose the best model among the several p and q orders according to specific criteria: Sum Squared Errors (SSE), Akaike (AIC) (1973) or Schwartz (SC) (1978).

$$AIC = \ln \hat{\sigma}_{\varepsilon t}^2 + \frac{2(p+q)}{n}$$

$$SC = n \ln \hat{\sigma}_{\varepsilon t}^2 + \frac{(p+q)\ln n}{n}$$

2.5.4. Forecasting:

At last, we take the model and use the time series observations to forecast future values.

3. GENETIC ALGORITHMS:

Holland from Michigan University has integrated biological concepts about genetics and natural selection in computing, so he was the first who created Genetic Algorithms and founded its theoretical basis in 1975.De Jong's works (*De Jong, 1980*) concerning function optimization were considered as first efforts to determine the application parameters. Then Goldberg's contributions has introduced Genetic Algorithms efficiency and made it more famous (*Goldberg, 1982, 1989*).

Genetic Algorithms is an artificial intelligence and heuristic technique that can be applied as a quantitative method on various types of problem. It finds the optimum solution by achieving the highest fitness value according to genetics and natural selection principals (*Coley, 1999*).

Genetic Algorithms has been used to solve many difficult problems in various technical domains like: computing, programming, artificial intelligent, biology, engineering, planning, industries… as well as social and economic sciences where it is applied in: economy, finance, management, marketing….where it's used as a tool that provides data needed to support a decision attached to quantitative model. Therefore, it is proven to be used even in econometrics, forecasting and time series (*Alander, 2012*).

3.1. Basic Elements:

Genetic Algorithms is based on elements, which are: population, individual, encoding and fitness; these elements depend to an environment which is the search space. Population consists of individuals, a chromosome in form of an encoded string represents each individual, and each chromosome is evaluated according the fitness function. Taking the basic principal of natural selection "survival of the fittest", the basic criteria in Genetic Algorithms is the fitness of population individuals to the environment.

3.1.1. Population:

The population is a group of individuals subject to Genetic Algorithms operations; it belongs to the search space where it lives in the studied environment. There are two important concepts of population: the initial generation and the size (n) (*Sivanandam & Deepa 2008*).

3.1.2. Individuals:

An individual is one of the possible solutions represented by a chromosome divided to genes, each gene represents a solution variable, and the chromosome must contain all the information needed about the represented solution. The coordination between genetic information and real values is done through the encoding that allows transforming the solution to suitable form to Genetic Algorithms (***Sivanandam & Deepa 2008***).

3.1.3. Encoding:

Encoding is the process of transforming the real values of variables or genes into string of codes (bits, numbers, letters….) adapted to Genetic Algorithms work, the most common way is the binary encoding, where each chromosome is represented by a binary string of bits (0 and 1), so each string represents a possible solution (***Adeli & Sarma, 2006***). For example:

Chromosome 1:0010110110

Chromosome 2:1000110011

3.1.4. Fitness:

The individual's fitness in Genetic Algorithms is the value of its goal function; it is the basic criterion of the evaluation and all operations. It can take any possible function that could be evaluated according to problem nature, so it can be: linear, non-linear, logarithmic, exponential… (***Rothlauf, 2006***)

3.2. Genetic Algorithms Operations: (***Sivanandam & Deepa 2008***)

Inspiring from natural selection phenomenon concepts that explain genes transmission among species through generations, Genetic algorithms run through several operations such as: selection, reproduction (crossover), mutation and replacement. These operations allow obtaining the fittest individual under specific environmental condition according to "Survival of the fittest" principal.

3.2.1. Selection:

It is the process of selecting parents from the population for reproduction according to natural selection, with hope of having a fitter individual. There are many ways of selection:

- Roulette wheel: a traditional way, where the selection is done using a wheel divided to selection probabilities according to fitness.
- Selecting the higher values individuals as parents.
- Selecting parents using a tournament between individuals according to fitness.
- Random selection by picking parents regardless to their fitness.

3.2.2. Reproduction:

Genes of the selected parents are emerged through crossover between chromosomes, which reproduce new offspring by exchanging parts in specific one or more points.
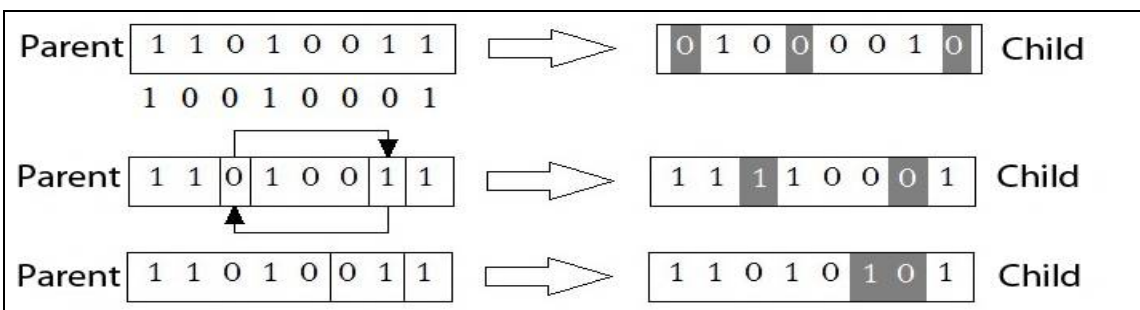
Fig. (03): Crossover types
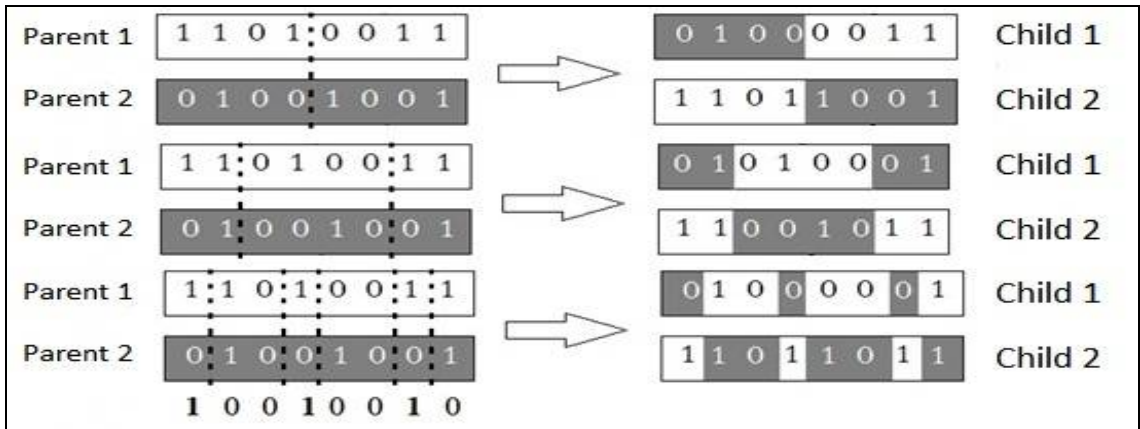
Source: Holland, 1992.

3.2.1. Mutation:

Mutation operation gives new Genotypes to cover some lost genes; the mutation is done by changing gene codes through: flipping codes, interchanging two specific points, or reversing codes in specific points.

Fig. (04): Crossover types

Source: Sivanandam, Deepa, 2008.

3.2.3. Replacement:

We nominate the new offspring individuals to enter the population, so we eliminate the weakest individuals from the current generation and replace them by better off spring according to fitness, while the population size stills always the same.

3.2.4. Termination:

The Genetic Algorithms stops when the terminations criteria are attained, so we can stop the algorithm after a specific time, specific number of generations or after no change in the fitness. We can also terminate according to following conditions:

- The best individual fitness overpasses a convergence value; this guarantees at least a good solution.
- The worst individual fitness overpasses a convergence value; this guarantees an entire good population, although the best individual may not be significantly better than the worst.
- The sum of fitness overpasses a convergence value considering the population size, this guarantees the population to be in a specific level of fitness, but the extreme values can affect this sum.
- The median fitness overpasses a convergence value; this guarantees at least half population of good solutions.

3.3. Genetic Algorithms Steps

Genetic Algorithms handles a population of possible solutions. First, we form an initial generation of individuals and evaluate them according to the fitness function, then we reproduce new individuals through a repeated loop of Genetic Algorithms operations: selection, reproduction, mutation and replacement. This loop continues until attaining termination criteria; we take the best individual to be the optimum solution. It is always possible to have a better solution in the next generation (***Haupt & Haupt, 2006***)

To clarify more, the figure (05) represents Genetic Algorithms steps and operations with a simulation of a natural example. We suppose a population of different mice, where the fitness is characterized by dark color and small size because cats have difficulties to find these kinds of mice (***Palisade Corporation, 2010***). To match between Genetic Algorithms and natural features we suppose that: each mouse represents a solution, the fittest mouse to this environment represents the optimal solution, natural factors represent problem constraints, and genes are the model variables.

Fig. (05): Genetic Algorithms Representation Simulating Natural Example
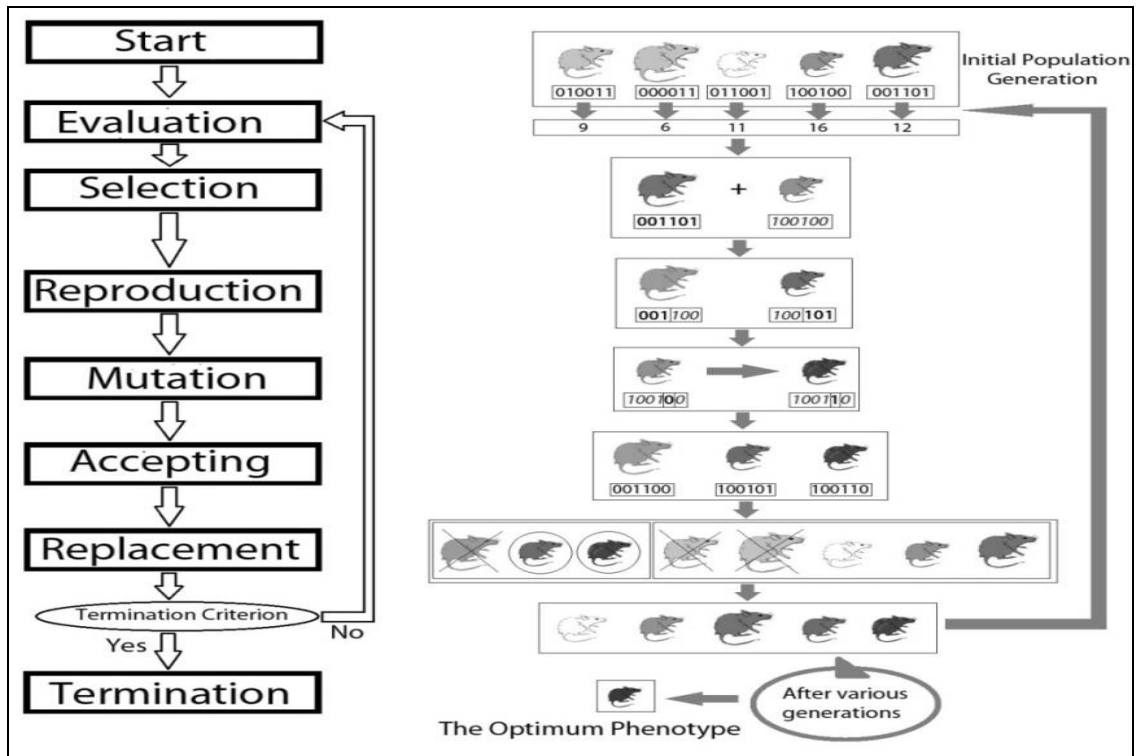
Source: by the researchers (inspired from: Haupt & Haupt, 2006, p. 28).

4. THE PRACTICAL STUDY:

After addressing the theoretical aspects about Box-Jenkins methodology and Genetic Algorithms, we try to make a practical study about forecasting sales with the help of these concepts. Therefore, we choose a time series that contains 60 observations of monthly sales of a plastic manufacturer.

This study aims to obtain the optimal model for sales forecasting using Genetic Algorithms according to ARIMA principals. Fora reason of testing this method efficiency; first, we must find the model through Box-Jenkins methodology using Eviews software, then we find it by Genetic Algorithms using Evolver software. At last, we make results comparison of both methods.

4.1. Eviews Results:

After following all Box-Jenkins methodology steps and testing several orders of ARIMA, the best model obtained according to quality criteria using Eviews is ARIMA (4; 1; 2) as follows:

Table (01): Eviews results

Variable	Coefficient	Std. Error	t-Statistic	Prob.
AR(1)	-1.016955	0.067857	-14.98670	0.0000
AR(2)	-0.976320	0.073701	-13.24697	0.0000
AR(3)	-0.969332	0.061143	-15.85349	0.0000
AR(4)	-0.922374	0.046775	-19.71936	0.0000
MA(1)	0.877750	0.116039	7.564240	0.0000
MA(2)	0.775165	0.103550	7.485865	0.0000
SIGMASQ	4053.783	875.0005	4.632892	0.0000
R-squared	0.894887	Meandependent var		4.593220
Adjusted R-squared	0.882759	S.D. dependent var		198.0682
S.E. of regression	67.81950	Akaike info criterion		11.54709
Sumsquaredresid	239173.2	Schwarz criterion		11.79358
Log likelihood	-333.6392	Hannan-Quinn criter.		11.64331
Durbin-Watson stat	1.456248			

Source: By the researchers.

$$\Delta \hat{x}_t = -1.016\Delta x_{t-1} - 0.976\Delta x_{t-2} - 0.969\Delta x_{t-3} - 0.922\Delta x_{t-4} + 0.877\varepsilon_{t-1} + 0.775\varepsilon_{t-2}$$

$$SSE = 239173.2$$

As we see in the table (01) all ARIMA (4;1;2) model coefficients are significantly different from zero (t-Statistic>1.96).

4.2. Evolver Results:

Evolver software is one of Palisade Decision Tools pack of decision support; it can be used to solve various problems using Genetic Algorithms. Therefore, we shape the model in Excel that includes: fitness function, variables and constraints; then we determine it in Evolver and start the repeated loop of Genetic Algorithms operations: selection, reproduction, mutation and replacement. At last, the termination is done according to chosen conditions.

ARIMA (p, 1, q) can generally represented by the following adapted Genetic Algorithm model:

$$Min\ SSE = \sum_{t=1}^{t} e_t^2$$

$$e_t^2 = (x_t - \hat{x}_t)^2$$

$$\hat{x}_t = x_{t-1} + \Delta\hat{x}_t$$

$$\Delta\hat{x}_t = \sum_{t=1}^{p} \theta_p \Delta x_{t-p} + \sum_{t=1}^{q} \alpha_q \varepsilon_{t-q} + c$$

The fitness function is the Sum Squared Errors (SSE) function; Individuals are the possible models and genes are model coefficients. Our proposed method based on launching several algorithms that include a model for each, every model is adapted to specific p and q orders of ARIMA.

We let the iterations of the operations loop for a while (as much time as possible); the fitness is evaluated generations after generations and the optimal forecasting model obtained is ARIMA (4; 1; 2) as follows:

$$\Delta\hat{x}_t = -1.037\Delta x_{t-1} - 0.994\Delta x_{t-2} - 0.982\Delta x_{t-3} - 0.937\Delta x_{t-4} + 0.869\varepsilon_{t-1} + 0.749\varepsilon_{t-2}$$

$$SSE = 232069.64$$

We use these models to calculate the estimated values and compare them with the real values to know how much similar they are.

Fig. (06): Real Values vs. Estimated Values
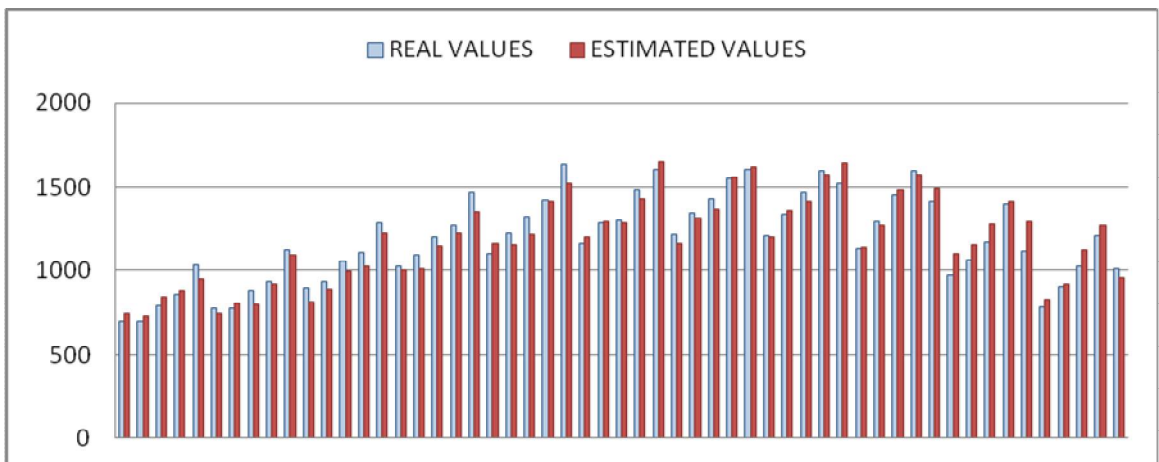
Source: By the researchers.

Table (02): Values estimating

	t	x_t	Δx_t	ε_t	$\hat{x}_t$
Past months	57	901	118	-16,66	917,66
	58	1023	122	-100,79	1123,79
	59	1209	186	-66,47	1275,47
	60	1013	-196	52,39	960,60
Future months	**61**	**-**	**-216,38**	**-**	**796,61**
	62	**-**	**161,34**	**-**	**957,96**
	63	**-**	**65,98**	**-**	**1023,95**
	64	**-**	**167,53**	**-**	**1191,49**

Source: By the researchers.

4.3. Results Discussion:

Forecasting models and Results of Genetic Algorithms was relatively better (SSE=232069.64<239173.2). In addition, Genetic Algorithms gives a population of solutions, but the important in our case is only the optimal model.

In Evolver, the entrance of data and model conditions could be in any available form to Excel, while in EViews the entrance of data should be directly with numbers in specific tables and specific forms.

Outputs are displayed in Evolver according to our choices and preferences (graphs, applications …), which allows to the user more freedom and more facility to consult results including model coefficients and estimated values directly from one Excel page. Views user must follow specific analytical steps, outputs displays separately at will; it could even take some calculations like the constant coefficient of ARIMA, which requires adjusting.

Regarding the taken time, it is better with Eviews because it takes few seconds, whereas with Evolver takes some few minutes but it is not constant because it depends to random Genetic Algorithms operations, which could create an uncertainty. It is also related to computer's speed, we advise to let the running on as long as possible and to repeat several times.

The comparison between both methods results confirms the efficiency of Genetic Algorithms for forecasting, and its ability to give the optimal model. The comparison shows Genetic Algorithms superiority using Evolver concerning characteristics excepting

the taken time criterion; however, we can neglect that, because it doesn't a matter if it takes longer time.

The comparison between real and estimated values by the obtained models shows a good similarity, that prove a great ability for forecasting, therefore there is a high efficiency of information that we can take as quantitative input for decision support about production and sales policy.

5. CONCLUSION

After recognizing Box-Jenkins methodology as the usual forecasting method, as well as recognizing Genetic Algorithms and their theoretical aspects including their basic elements and operations. We succeeded to apply these methods in a practical study about monthly sales of a plastic manufacturer, and then we found the optimal ARIMA model for forecasting by analyzing time series using Genetic Algorithms with the help of Evolver software. Evolver results showed that the estimated values by the obtained ARIMA(4;1;2) model are similar to real values of the time series; in addition, the comparison between both methods showed outperformance of Genetic Algorithms excepting the taken time criterion.

Therefore, we concluded that Genetic Algorithms could be used efficiently for sales forecasting according to ARIMA principals, which make it very helpful for management and decision support concerning sales strategies

6. REFERENCES

1.   Adeli, H., Sarma, K.C. (2006). ***Cost Optimization of Structures Fuzzy Logic***, *Genetic Algorithms and Parallel Computing*, John Wiley & Sons, England.

2.   Akaike H. (1973). ***Information Theory and an Extension of the Maximum Likelihood Principle***, in B.N. Petrov and F. Csaki, eds, *2nd international symposium on information theory*, Akademia Kiado, Budapest.

3.   Alander, J.T. (2012). ***An Indexed Bibliography of Genetic Algorithms in Economics***, *Report Series No. 94-1-ECO*, Finland.

4.   Bourbonnais R. (2011). ***Econométrie : Manuel et exercices corrigés***, Dunod, Paris.

5.   Box G., Jenkins G. (1976).***Time Series Analysis Forecasting and Control***, Holden-day, Oakland, California.

6.   Coley, D.A. (1999). ***An Introduction to Genetic Algorithms for Scientists and Engineers***, World Scientific, USA.

7.   Dickey D., Fuller W. (1979). **Distribution of the Estimators for Autoregressive Time Series with Unit Root**, *Journal of American statistical association*, vol 74, N°366.

8.   Dickey D., Fuller W. (1981). ***Likelihood Ratio Statistics Estimators for Autoregressive Time Series with Unit Root***, *Econometrica*, vol 49, N°4.

9. Goldberg, D.E. (1989). ***Genetic Algorithms in Search, Optimization, and Machine Learning***, Addison-Wesley Publishing Company, USA.

10. Goldberg, D.E. (1981). ***Robust Learning and Decision Algorithms for Pipeline Operations***. Unpublished dissertation proposal, *University of Michigan*, Ann Arbor.

11. Greene W., ***Econometric Analysis***, seventh edition, Pearson. 2012.

12. Haupt, R.L., Haupt, S.E. (2006). ***Practical Genetic Algorithms***, 2nd edition, John Wiley & Sons, USA.

13. Holland, J.H. (1975). ***Adaptation in Natural and Artificial Systems***, *University of Michigan Press*.

14. Holland, J.H. (1992). *Genetic Algorithms*, *Scientific American*, p. 66-72.

15. Jong, K.D. (1980). ***Adaptive System Design: A Genetic Approach, IEEE Transactions on Systems,*** *Man and Cybernetics* 10, p 556–574.

16. Palisade Corporation (2010). ***Guide to Using Evolver the Genetic Algorithm Solver for Microsoft Excel Version 5.5***, January.

17. Phillips P., Perron P. (1988). ***Testing for Unit Root in Time Series Regression***, *Biometrika*, vol 75.

18. Rothlauf, F. (2006). ***Representations for Genetic and Evolutionary Algorithms***, 2nd edition, Springer, Netherlands.

19. Sivanandam, S.N., Deepa, S.N. (2008). ***Introduction to Genetic Algorithms***, Springer, USA.

20. Schwartz G. (1978). **Estimating the Dimension of a Model, *The annals of statistics***, vol 6.

21. Tsay R. (2002). ***Analysis of Financial Time Series***, John Wiley & Sons, Canada.