Using Genetic Algorithms for Forecasting Financial Markets Volatility

Omar MOUFFOK*, Youcef SOUAR**

Received:06/06/2018
Accepted: 21/07/2018

Abstract

This study aims to use Genetic Algorithms, as a developed artificial intelligent technique to forecast volatility of financial markets according to Econometrics principals. Therefore, we try to apply it on three stock markets depending on their indexes time series: Tunindex, Madex and Dow Jones. Using Evolver software, we succeeded to obtain the optimal forecasting models, and then we make a comparison with Econometrics methods. From the results, we conclude that it is possible to use Genetic Algorithms efficiently in financial markets volatility forecasting, in addition it has some advantages concerning analytical characteristics comparing to the other methods.

Keywords: *genetic algorithms, volatility, financial markets, quantitative methods, forecasting, time series, optimization.*

JEL Classification : C01, C32, C58, G17.

ملخص:

تعتبر الخوارزميات الجينية من التقنيات الذكاء الاصطناعي المتطورة التي تستخدم في حل المسائل في مختلف المجالات، و عليه تهدف هذه الدراسة إلى استخدامها في التنبؤ بتقلبات الأسواق المالية وفقا لمبادئ طرق الاقتصاد القياسي. بالاعتماد على برنامج Evolver قمنا بتطبيقها على ثلاث أسواق مالية من خلال تحليل السلاسل الزمنية لمؤشرات البورصة الخاصة بكل منها و هي: Tunindex، Madex و Dow Jones. بذلك تمكنا من الحصول على النموذج الأمثل للتنبؤ لكل مؤشر ثم قمنا بالمقارنة مع الطرق القياسية. استنتجنا من خلال النتائج أنّ الخوارزميات الجينية يمكن أن تستخدم بفعالية في التنبؤ بتقلبات الأسواق المالية، كما أنها تتميز عن الطرق الأخرى بعدة خصائص تحليلية.

الكلمات المفتاحية : خوارزميات جينية، تقلبات، أسواق مالية، طرق كمية، تنبؤ، سلاسل زمنية، أمثلية.

* Abou Bekr Belkaid University – Tlemcen, Algeria, mo-om22@hotmail.fr
** Taher Moulay University – Saida, Algeria, syoucef12@yahoo.fr

1.Introduction

It is very important to forecast economic and financial behaviors, one of the most forecasting issues targeted by researchers is the volatility of financial markets, that because of its importance in risk measurement that helps in the investing decision.

The time is one of the basic factors that influence financial markets behaviors and their volatility. So, time series that include values overtime could be analyzed by mathematical techniques and specific methodology to forecast future values.

There are many methods to analyze time series. The most common are Econometrics methods including Box-Jenkins method, ARIMA and ARCH models; which are based on some mathematical and statistical tests through specific steps under determined conditions and criterions.

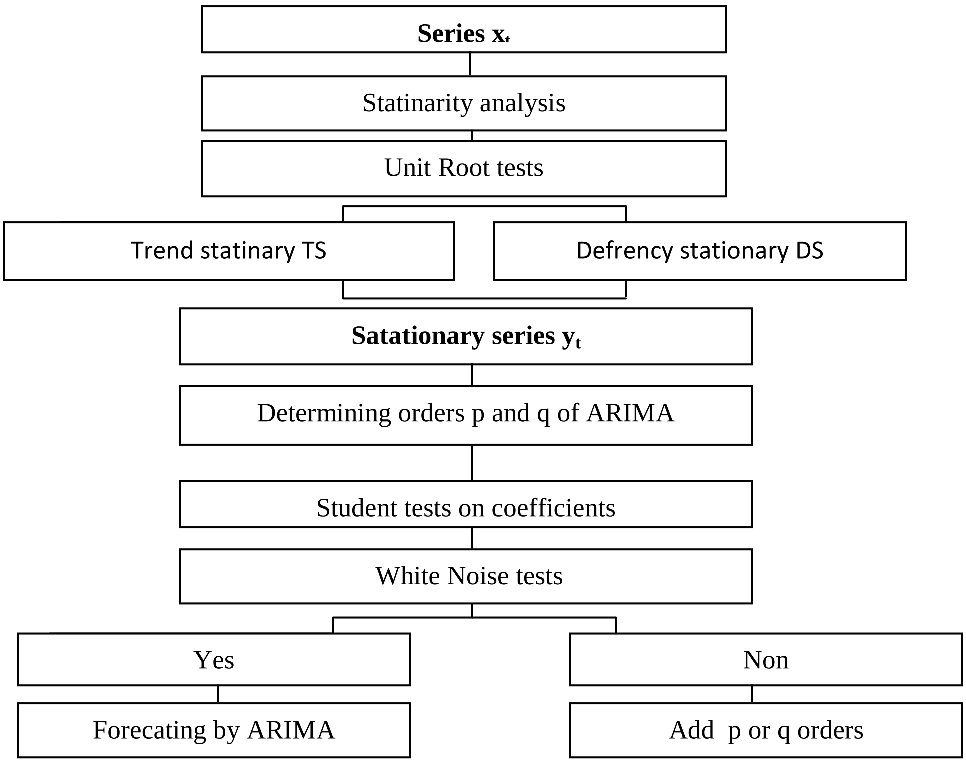
There were many studies about volatility forecasting, the most remarkable are Engle studies who was the first who introduce ARCH models in 1982 and won Nobel Prize for his contributions in 2003. Bollerslev introduced GARCH models in 1986 then there was more many studies on various financial markets.

Due to the development of networks and artificial intelligence, new methods such as Genetic Algorithms has been discovered, that method simulates scientific explanations in genetics and natural evolution for getting an optimal solution population, it has a wide scope of application in various kinds of domain including forecasting.

Thus, how it could be possible to use Genetic Algorithms for forecasting financial markets volatility?

2.Econometrics methods for volatility forecasting

Box-Jenkins approach is a method used for studying time series in order to determine the most adapted ARIMA model of a given phenomenon. The steps of Box-Jenkins method (Box & Jenkins, 1976) are as follows:



Source : (Bourbonnais, 2011, p 261)

Figure (1): Box-Jenkins methodology steps

2.1.Model identification: (Bourbonnais, 2011)

There are three conditions of the statinarity of time series:

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

Then stationarity is analyzed using Autocorrelation function (ρ_k) analysis. In case of no stationarity there are two processes to study the time series: Trend Stationary (TS) and Deferency Stationary (DS). To determine the type of process we use Unit Root tests: Dickey-Fuller (1979), Augmented Dickey-Fuller (1981), Phillips-Perron (1988).

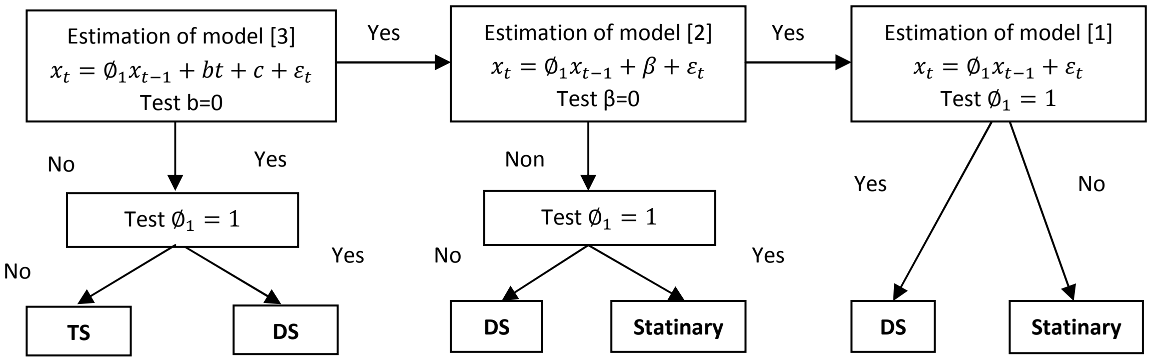
First we estimate the following models using Ordinary Least Squares (OLS):

[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 the following tests using Dickey-Fuller tables:



Source: (Bourbonnais, 2011, p 249)

Figure (2): Unit Root tests

One of ARIMA models family could fix the non-stationarity problem:

- Auto Regressive models (AR):

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

- Moving Average models (MA):

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

- ARMA models**:**

ARMA(p,q):

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

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

- Integrated ARMA:

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

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

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

2.2.Model estimation:

After determining p and q orders we estimate models coefficients using mathematical methods, AR(p) coefficients can be estimated using OLS, but MA(q) estimation requires more complicated methods like Likelihood maximization methods. (Tsay, 2002)

2.3.Model diagnostic checking:

We check that all coefficients are significantly different to zero, furthermore we perform White noise tests, among the several p and q orders we choose the best model according to Akaike (1973) or Schwartz (1978) criteria.

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

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

2.4.Forecasting:

We take the model and forecast future values depending on past observations of the time series.

Even so, these models are not good enough for volatility forecasting concerning financial time series, where there is mostly a Heteroscedasticity problem.

2.5.Autoregressive Conditional Heteroscedasticity models:

The Heteroscedasticity problem of a model is solved through Autoregressive Conditional Heteroscedasticity models and their derives (ARCH, GARCH…).

2.5.1.**ARCH models:** (Engle, 1982)

ARCH(p) : $\sigma^2 = a_0 + a_1\varepsilon^2_{t-1} + a_2\varepsilon^2_{t-2} + \cdots + a_p\varepsilon^2_{t-p}$

In condition of :

$a_0 > 0\,; (a_i = a_1; a_2; \dots; a_p) \geq 0$

2.5.2.GARCH models: (Bollerslev, 1986)

GARCH(p,q):

$$\sigma^2 = a_0 + a_1\varepsilon^2_{t-1} + a_2\varepsilon^2_{t-2} + \cdots + a_p\varepsilon^2_{t-p} + \beta_1\sigma^2_{t-1} + \beta_2\sigma^2_{t-2} + \cdots + \beta_q\sigma^2_{t-q}$$

In condition of:

$$a_0 > 0\,; \left(a_i = a_1; a_2; \dots; a_p\right) \geq 0$$

$$(\beta_j = \beta_1; \beta_2; \dots; \beta_q) \geq 0$$

$$\left(\sum_{i=1}^{p} a_i + \sum_{j=1}^{q} \beta_j\right) > 1$$

These models cannot be estimated using normal mathematic methods, it depends mainly on the maximization of the Log-Likelihood function, and thus it requires complicated algorithms.

$$LL = -\frac{n}{2}\ln(2\pi) - \frac{1}{2}\sum_{t=0}^{n} \ln\sigma^2_t - \frac{1}{2}\sum_{t=0}^{n} \ln\frac{\varepsilon^2_t}{\sigma^2_t}$$

3.Genetic Algorithms

John Holland has integrated biological aspects about genetics and natural selection in computing, so he was the first who created Genetic Algorithms and founded its theoretical basis in 1975 (Holland, 1975). De Jong’s works about function optimization (De Jong, 1980) and David Goldberg’s book (Goldberg, 1989) and his work about pipeline operations (Goldberg, 1981) have introduced the efficiency Genetic Algorithm and have made it more famous.

Therefore, it has been used to solve difficult problems in various domains like computing, programming, artificial intelligent, biology, engineering, planning, decision making…. It is proven to be used in Operational Research, data mining, econometrics, forecasting and time series. (Alander, 2012)

Genetic Algorithms is an artificial intelligent technique and a heuristic method uses genetics and natural selection principals to give the optimum solution among all the possible solutions, according to specific fitness function it can be applied to any kind of problems (Coley, 1999). Genetic Algorithms is characterized by: the randomness as a principal of all its operations, it gives a population of solutions; performance elasticity. (Sivanandam & Deepa 2008)

3.1.Basic elements

Genetic Algorithms is based on the following elements: population, individual, encoding and fitness; these elements depend to an environment is the search space. Population consists of individuals subject to Genetic Algorithms operations. A chromosome in form of an encoded string represents each individual, and each chromosome represents a possible solution evaluated according the fitness function that can take any mathematical form. (Rothlauf, 2006)

All chromosomes containing genetic information in form of genes; each gene represents a solution variable. (Sivanandam & Deepa 2008)

Encoding is the process of transforming the real values of variables into string of codes adapted to Genetic Algorithms work. The binary encoding is the most common way, where a binary string (0 and 1) represents each chromosome (Adeli & Sarma 2006), For example: *Chromosome 1: 0010100110; Chromosome 2: 1010110011…*

3.2.Genetic Algorithms operations and steps

Genetic Algorithms handles a population of possible solutions. First, we form an initial population generation of individual with a specific size (n), we evaluate them according to the fitness function, and then we reproduce new individual through a repeated loop of operations: selection, reproduction (crossover), mutation and replacement. This loop continues until attaining termination criterions. There is always a possibility to have a better solution in the next generation. (Sivanandam & Deepa, 2008)

Selection is the process of selecting parents from the population for reproduction according to natural selection. There are several ways of selection: using Roulette wheel, taking the higher values individuals as parents, using a tournament between individuals or picking parents randomly regardless to their fitness.

After selection, parents reproduce new offspring, so it exchange parts of the parent’s strings in one or more point to have new strings, the genes get emerged by crossover between chromosomes which allows Genotypes diversity.

After reproduction, we move to mutation in a specific site of the string, which gives new Genotypes and covers some lost genes and prevent to fall in tight space.

Replacement is the last part of Genetic Algorithms operation loop; we nominate the new offspring to enter the population. According to fitness, we eliminate the weakest individuals and replace them by the best offspring (“Survival of the Fittest”).

The Genetic Algorithms stops when the terminations criterions are attained, then we take the best individual to be the optimal solution, so we can stop the algorithm after a specific time, specific number of generations, and no change in the fitness or attaining a convergence value.

4.The practical study

After addressing the theoretical aspects about Econometrics models of volatility forecasting and Genetic Algorithms, we try to make a combination between the both. So we apply it on three financial time series that contain ten years of observations.

We have chosen three time series of stock market indexes, two local (North Africa): Tunindex of Tunis, Madex of Casablanca and universal one: Dow Jones of New York.

We collect the data of the time series and indexes values from ***investing.com*** website, the seires of Tunindex contain 2482 observations from 02/01/2008 to 29/12/2017. The seires of Madex contain 1900 observations from 19/05/2008 to 29/12/2017. The seires of Dow Jones contain 2517 observations from 02/01/2008 to 29/12/2017.

The main goal of this study is to obtain the optimal model for volatility forecasting of each index, therefore we try to find the maximum likelihood using Genetic Algorithms according to ARIMA and GARCH principals.

First, we find models through Econometrics method using EViews 9 software. Then we create an adapted model of Genetic Algorithms using Evolver software. At last, we make comparison of results and test the efficiency of our proposed method.

Evolver is a Palisade is one of Palisade Decision Tools software pack that could be used in decision support. Evolver can solve various problems using Genetic Algorithms. After shaping the model in Excel that includes: fitness function, variables and constraints; then determining the model in Evolver, we start the task of applying the repeated loop of Genetic Algorithms operations: selection, reproduction, mutation and replacement. The termination could be whether automatic according to the chosen parameters or manual according to our estimation.

The adapted Genetic Algorithm global model used for this study is as follows:

$$fitness = \max LL$$

$$LL = \sum_{t=1}^{n} l_t$$

$$l_t = ln\left(\frac{1}{\sqrt{2\pi\hat{\sigma}_t^2}}\right)e^{-0,5\frac{\varepsilon_t^2}{\hat{\sigma}_t^2}}$$

$$\hat{\sigma}_t^2 = a_0 + a_1\varepsilon_{t-1}^2 + a_2\varepsilon_{t-2}^2 + \beta_1\sigma_{t-1}^2 + \beta_2\sigma_{t-2}^2$$

$$\begin{cases} (a_0; a_1; a_2; \beta_1; \beta_2) > 0 \\ (a_1 + a_2 + \beta_1 + \beta_2) < 1 \end{cases}$$

$$\Delta\hat{x}_t = \theta_1\Delta x_{t-1} + \theta_2\Delta x_{t-2} + \theta_3\Delta x_{t-3} + \alpha_1\varepsilon_{t-1} + \alpha_2\varepsilon_{t-2} + \alpha_3\varepsilon_{t-3} + bt + c$$

$$\varepsilon_t = \hat{x}_t - x_t$$

The fitness function is the Log-Likelihood function; we assume that is preferable to use the Quasi-ML that follows Normal Distribution (Wooldridge & Bollerslev, 1992). Individuals are the possible models and genes are model coefficient.

Following most of past studies, we assume number 3 of maximum p and q orders for ARIMA and number 2 for maximum GARCH orders. Our proposed method based on launching several algorithms that include a model for each, every model is adapted to specific order of ARIMA and GARCH.

After following all Box-Jenkins methodology steps and testing several orders of ARIMA and GARCH, the best models obtained according to Akaike criterion of each index using EViews are as follows:

Table (1): Tunindex results using EVews

Variable	Coefficient	Std. Error	z-Statistic	Prob.
C	1.660181	0.704742	2.355727	0.0185
AR(1)	1.159010	0.053952	21.48223	0.0000
AR(2)	-0.206517	0.030867	-6.690505	0.0000
MA(1)	-0.912757	0.046204	-19.75492	0.0000
Variance Equation				
C	99.34773	8.958687	11.08954	0.0000
RESID(-1)^2	0.287542	0.017618	16.32089	0.0000
GARCH(-1)	0.536581	0.026965	19.89910	0.0000
R-squared	0.064065	Mean dependent var		1.475179
Adjusted R-squared	0.062931	S.D. dependent var		25.40433
S.E. of regression	24.59198	Akaike info criterion		8.901844
Sum squared resid	1498004.	Schwarz criterion		8.918254
Log likelihood	-11035.74	Hannan-Quinn criter.		8.907804
Durbin-Watson stat	2.002009			

Table (2): Madex results using EVews

Variable	Coefficient	Std. Error	z-Statistic	Prob.
AR(1)	0.101494	0.025307	4.010432	0.0001
Variance Equation				
C	178.4331	30.05033	5.937810	0.0000
RESID(-1)^2	0.173362	0.019200	9.029418	0.0000
GARCH(-1)	0.336476	0.091758	3.666974	0.0002
GARCH(-2)	0.433845	0.078790	5.506360	0.0000
R-squared	0.026676	Mean dependent var		-0.035334
Adjusted R-squared	0.026676	S.D. dependent var		56.78360
S.E. of regression	56.02110	Akaike info criterion		10.69624
Sum squared resid	5956614.	Schwarz criterion		10.71085
Log likelihood	-10151.08	Hannan-Quinn criter.		10.70162
Durbin-Watson stat	1.838492			

Table (3): Dow Jones results using EVews

Variable	Coefficient	Std. Error	z-Statistic	Prob.
@TREND	0.007064	0.001359	5.198424	0.0000
AR(1)	-0.061919	0.020972	-2.952433	0.0032
Variance Equation				
C	821.9607	127.7832	6.432461	0.0000
RESID(-1)^2	0.072922	0.013173	5.535894	0.0000
RESID(-2)^2	0.067177	0.019632	3.421782	0.0006
GARCH(-1)	0.812649	0.019724	41.20048	0.0000
R-squared	0.006632	Mean dependent var		4.640405
Adjusted R-squared	0.006237	S.D. dependent var		138.3357
S.E. of regression	137.9037	Akaike info criterion		12.42076
Sum squared resid	47809793	Schwarz criterion		12.43466
Log likelihood	-15619.31	Hannan-Quinn criter.		12.42580
Durbin-Watson stat	2.034506			

As we see in the tables all coefficients are significantly different from zero (z-Statistic >1.96) statistics, as well as GARCH coefficients are all positives.

We launched Genetic Algorithms of several models using Evolver, and we let the iterations of the processes loop for hours, the fitness has evaluated generations after generations and the optimal models for forecasting for each index are as follows:

- Tunindex:

$$\Delta\hat{x}_t = 1.1534\Delta x_{t-1} - 0.2042\Delta x_{t-2} - 0.9074\varepsilon_{t-1} + 0.0824$$
$$\hat{\sigma}_t^2 = 98.8441 + 0.2876\varepsilon_{t-1}^2 + 0.5379\sigma_{t-1}^2$$
$$LL = -11035.69$$
$$AIC = 8.9018$$

- Madex:

$$\Delta\hat{x}_t = 0.1015\Delta x_{t-1}$$
$$\hat{\sigma}_t^2 = 178.6502 + 0.1726\varepsilon_{t-1}^2 + 0.3348\sigma_{t-1}^2 + 0.4356\sigma_{t-2}^2$$
$$LL = -10151.08$$
$$AIC = 10.6962$$

- Dow Jones:

$$\Delta\hat{x}_t = -0.0619\Delta x_{t-1} + 0.0074$$
$$\hat{\sigma}_t^2 = 818.3099 + 0.0736\varepsilon_{t-1}^2 + 0.0659\varepsilon_{t-2}^2 + 0.8133\sigma_{t-1}^2$$
$$LL = -15619.43$$
$$AIC = 12.4208$$

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

Table (4): Real Values vs. Estimated Values in the last month (December 2017)

Date	Tunindex		Madex		Dow Jones	
	Real values	Estimated values	Real values	Estimated values	Real values	Estimated values
01/12/2017	6219,47	6229,64635	10172,69	10253,0138	24231,59	24270,5394
04/12/2017	6228,98	6220,8385	10136,73	10163,7385	24290,05	24252,8484
05/12/2017	6226,16	6233,22164	10021,92	10133,0796	24180,64	24305,172
06/12/2017	6216,28	6227,4552	10083,53	10010,2652	24140,91	24206,1643
07/12/2017	6206,1	6215,6839	10102,53	10089,7843	24211,48	24162,1271
08/12/2017	6178,31	6205,15615	10054,18	10104,4588	24329,16	24225,8747
11/12/2017	6169,74	6172,78051	10167,26	10049,2718	24386,03	24340,645
12/12/2017	6156,99	6168,3742	10126,33	10178,7392	24504,8	24401,288
13/12/2017	6141,8	6154,44791	10152,3	10122,175	24585,43	24516,2325
14/12/2017	6149,26	6138,44428	10114,44	10154,9363	24508,66	24599,2317
15/12/2017	6142,25	6151,23542	10062,88	10110,5967	24651,74	24532,2157
18/12/2017	6138,68	6140,8769	10047,77	10057,6459	24792,2	24661,6897
19/12/2017	6121,36	6138,07049	9990,25	10046,2361	24754,75	24802,3194
20/12/2017	6134,32	6117,3587	9952,34	9984,41092	24726,65	24775,8934
21/12/2017	6142,15	6137,49751	9920,95	9948,4916	24782,29	24747,2219
22/12/2017	6186,37	6144,39417	10001,81	9917,76348	24754,06	24797,6841
26/12/2017	6213,75	6197,76573	10025,26	10010,0184	24746,21	24774,655
27/12/2017	6229,58	6221,87416	10040,2	10027,6405	24774,3	24765,5505
28/12/2017	6262,43	6235,33484	10063,98	10041,7166	24837,51	24791,4225
29/12/2017	6281,83	6272,58064	10100,32	10066,394	24719,22	24852,4654

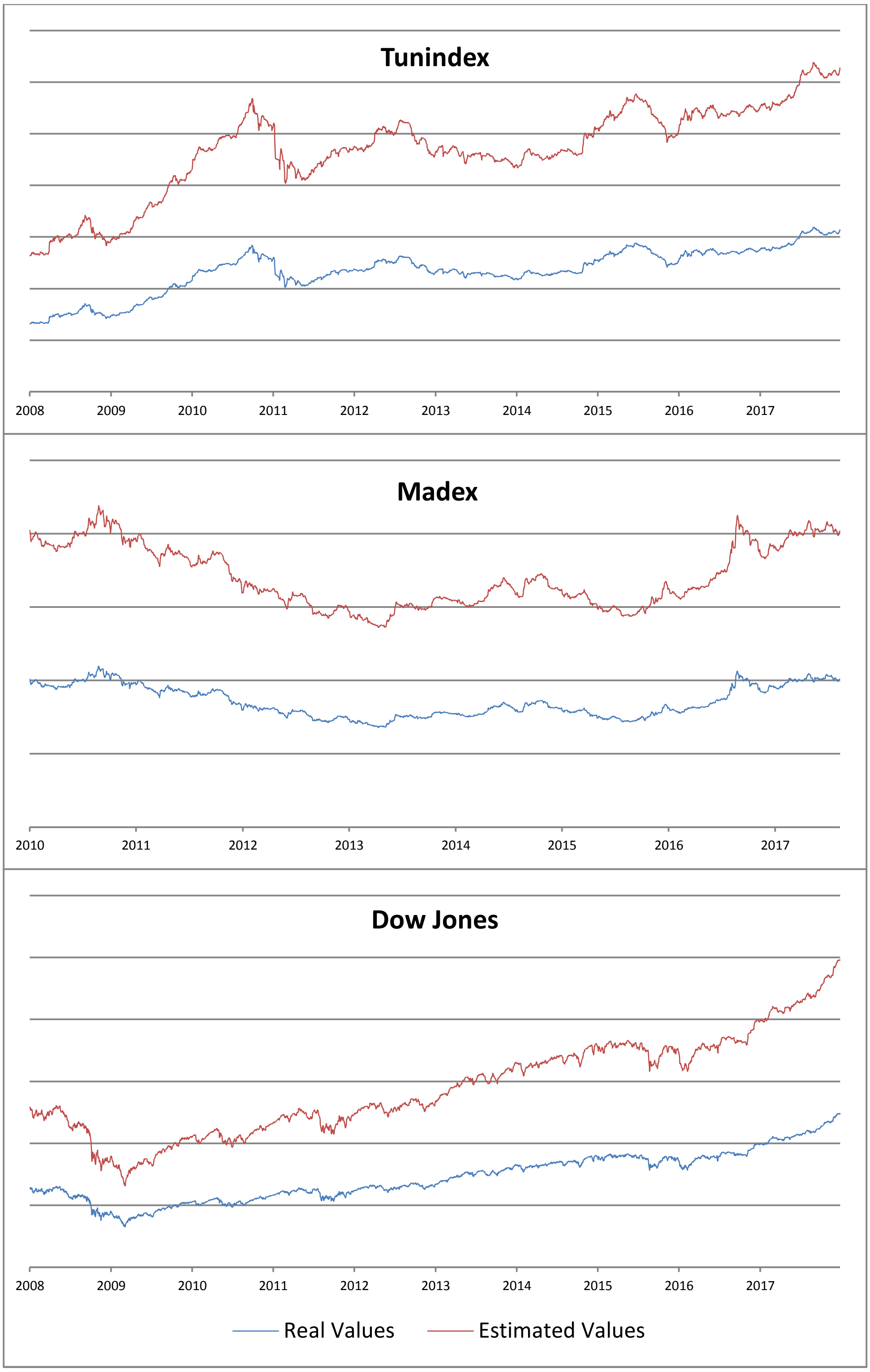


Figure (3): Real Values vs. Estimated Values

Both methods have given quite the same model for forecasting, which gives a confirmation of results. Genetic Algorithms gives a population of solutions, but the important in our case is only the optimal model.

In Evolver, the data entrance could be in any available form to Excel, besides outputs and model are displayed according to our choices and preferences including graphs and applications … which allow to the user more freedom and more facility to consult results and estimated values directly from one Excel page. While in EViews the entrance of data should be directly with numbers in specific tables and specific form, the user must follow specific analytical steps and outputs displays separately on demand, some could take some calculations like the constant coefficient of ARIMA, which requires adjusting.

Evolver also allows putting GARCH model constraints where all coefficients must be superior to zero and their sum must be inferior to zero, thus our Genetic Algorithms model respects these constraints, and when a coefficient goes negative it takes a value of zero automatically. While EViews makes the estimation and the user must check these constrains.

So, these are advantages for Genetic Algorithms, where the use is quite simple after creating the model on excel we launch the algorithm and wait for evolution.
There is also a disadvantage for Genetic Algorithms concerning the taken time, it is not constant because it depends to random operations, and this could create an uncertainty. In our case after experimenting several times with different computers, the taken time for less p and q orders (like the obtained for Madex and Dow Jones) was between 30 minutes and one hour, for more p and q orders (like the obtained for Tunindex) it took more than 6 hours for each model. The taken time depends also to quality of computer where it can take a short time with a high-speed processor.

The equality of results between both methods 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 even if it takes long time, it can be helpful in many cases especially.

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

Nevertheless, there is a less similarity concerning Dow Jones, which means less forecasting ability and that because of lack of information efficiency of this market comparing to the other two markets. Therefore, the information is not enough and it doesn't encourage for investing while the estimated values don't represent the real values accurately. That requires more complicated models contain more explicative variables like the ARFIMA models family.

5.Conclusion

We have succeeded to obtain the optimal models to forecast volatility of three financial markets: Tunis (Tunindex), Casablanca (Madex) and New York (Dow Jones). After analyzing the results, we concluded that Genetic Algorithms could be used efficiently for forecasting according to Econometrics principals. Form the comparisons with Econometrics methods using EViews, Genetic algorithms using Evolver showed advantages concerning characteristics except the taken time. The estimated values of stock market indexes are similar to real values, which confirm the good quality of the obtained models especially Tunindex's and Madex's, so we can take them as good information for investing decision. However, the estimated values was less similar and less accurate concerning Dow Jones, which means that our model wasn't good enough and requires more complicated model like ARFIMA models; we hope to consider this in other researches.

References

- Adeli, H., Sarma, K.C. (2006). *Cost Optimization of Structures Fuzzy Logic, Genetic Algorithms and Parallel Computing*, John Wiley & Sons, England.
- 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.
- Alander, J.T. (2012). An Indexed Bibliography of Genetic Algorithms in Economics, *Report Series No. 94-1-ECO*, Finland.
- Bollerslev, T. (1986). Generalized Autoregressive Conditional Heteroskedasticity, *Journal of Econometrics*, North-Holland, 31, 307-327.
- Bourbonnais, R. (2011). *Econométrie : Manuel et exercices corrigés*, Dunod, Paris.
- Box, G., Jenkins, G. (1976). *Time Series Analysis Forecasting and Control*, Holden-day, Oakland, California.
- Coley, D.A. (1999). *An Introduction to Genetic Algorithms for Scientists and Engineers*, World Scientific, USA.
- 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.
- Dickey, D., Fuller W. (1981). Likelihood Ratio Statistics Estimators for Autoregressive Time Series with Unit Root, *Econometrica*, vol 49, N°4.
- Engle, R. (1982). Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation, *Econometrica*, Vol. 50, No. 4, pp. 987-1007.
- Goldberg, D.E. (1989). *Genetic Algorithms in Search, Optimization, and Machine Learning*, Addison-Wesley Publishing Company, USA.
- Goldberg, D.E. (1981). Robust Learning and Decision Algorithms for Pipeline Operations. Unpublished dissertation proposal, *University of Michigan*, Ann Arbor.
- Haupt, R.L., Haupt, S.E. (2006). *Practical Genetic Algorithms*, 2nd edition, John Wiley & Sons, USA.
- Holland, J.H. (1975). Adaptation in Natural and Artificial Systems, *University of Michigan Press*.
- Jong, K.D. (1980). Adaptive System Design: A Genetic Approach, IEEE Transactions on Systems, *Man and Cybernetics* 10, p 556–574.
- Palisade Corporation (2010). *Guide to Using Evolver the Genetic Algorithm Solver for Microsoft Excel Version 5.5*, January.
- Phillips, P., Perron, P. (1988). Testing for Unit Root in Time Series Regression, *Biometrika*, vol 75.
- Rothlauf, F. (2006). *Representations for Genetic and Evolutionary Algorithms*, 2nd edition, Springer, Netherlands.
- Sivanandam, S.N., Deepa, S.N. (2008). *Introduction to Genetic Algorithms*, Springer, USA.
- Schwartz G. (1978). Estimating the Dimention of a Model, *The annals of statistics*, vol 6.
- Tsay, R. (2002). *Analysis of Financial Time Series*, John Wiley & Sons, Canada.
- Wooldridge, J., Bollerslev, T. (1992). Quasi-Maximum Likelihood Estimation and Inference in Dynamic Models with Time-Varying Covariances, *Econometric Reviews*, 11(2), 143-172.