THE INTERNATIONAL JOURNAL OF SCIENCE & TECHNOLEDGE

Empirical Comparisons of Univariate Exponential Smoothing and Moving Average Methods for Forecasting Crime Rate

G.O. Daramola
Head, Training and Research Unit, Computer Resource Centre,
Federal University of Technology, Akure, Nigeria

Abstract:
The purpose of this research paper is to carry out comparative analysis and evaluation of some univariate time series models suitable for forecasting crime rate, and determine the most appropriate forecasting method. Moving average methods and Exponential smoothing methods of forecasting were used on monthly crime data obtained in Ekiti State of Nigeria from January, 2004 to December, 2012. The accuracy of the forecasting methods was measured using Mean Forecast Error (MFE), Mean Absolute Deviation (MAD), Mean Square Error (MSE) and Root Mean Square Error (RMSE). The result showed that 3-month moving average method produced the most accurate forecasting based on the data obtained in the ranking table.

Keywords: *Crime, forecasting, exponential smoothing, moving average, accuracy, time series*

1. Introduction

Crime is any act or offence which violates the law and socially acceptable norms of the people or the government and thus punishable under the law. It is however necessary to know that there are different categories of crime. For instance, crimes committed by those who are under the age of maturity i.e. by people under age eighteen is regarded as juvenile delinquency. Moreso, there are crimes against persons and property, which involves inflicting suffering on the victims i.e. crimes like murder, manslaughter, theft, armed robbery, rape, indecent assaults etc. Other categories of crimes are, organized crime, white collar crime etc. Forecasting crime is thus, a challenge for law enforcement agents in the field of criminal intelligence.

Time series is a chronological sequence of observations on a particular variable. Usually the observations are taken at reg ular intervals; days, months, years, etc. the sampling could be irregular. Common examples of time series are the Gross Domestic Product, Unemployment rate, etc. Time series analysis is a procedure that can be used to analyze historical information, build models and predict trends. A time series analysis therefore basically consists of two steps:

(1) Building a model that represents a time series, and
(2) Using the model to predict (forecast) future values.

Forecasting is thus regarded as a technique for predicting what will occur in future. It is a process of predicting the future based on past and present data. Forecasting provides pertinent information about potential future events and their consequences for any society. Forecasting crime helps in policy and planning so as to enhance tactical deployment of limited police resources.

2. Literature Survey

There has been a great deal of research endeavours on the application of various forecasting models for forecasting various issues. e.g. Wilpen et al. (2003) in the research work titled “Short-term forecasting of crime”, investigated whether it is possible to accurately forecast selected crimes one month ahead in small areas such as police precincts. He was able to establish that average crime count by precinct is the major determinant of forecast accuracy. He also observed that Holt exponential smoothing with seasonality estimated using city-wide data is the most accurate forecast model for precinct level crime series.

Several other time series forecasting techniques such as naive model, moving average, double moving average, exponential smoothing, semi average etc. have equally being applied to forecasting. For example, Pradeep and Rajesh (2013) analyzed various forecasting models such as naive model, moving average, double moving average, simple exponential smoothing and semi-average method. The accuracy of the forecasting methods was measured using Mean Forecast Error (MFE), Mean Absolute Deviation (MAD), Mean Square Error (MSE), Root Mean Square Error (RMSE). The study identified the most appropriate forecast method for sales of salted butter milk in Chhattisgarh, India based on accuracy and simplicity. The result showed that Naive Method obtained the best accuracy, and was selected as the most appropriate forecasting method for forecasting sales of salted butter milk in Chhattisgarh.

Ryu and Sanchez (2003) analyzed forecasting methods for institutional food service facility by identifying the most appropriate forecasting method for forecasting meal count for an institutional food service facility. The naive model 1,2 and 3, moving average method, double moving average method, exponential smoothing method, double exponential smoothing method, Winter method, Holt's method, simple linear regression and multiple linear regression methods were used. The accuracy of the forecasting methods was measured using mean absolute deviation, mean squared error, mean percentage error, mean absolute percentage error method, root mean squared error and Theil's U-statistic. The result of the study showed that multiple regressions was the most accurate forecasting method but naïve method 2 was selected as the most appropriate method because of its simplicity and high level of accuracy.

Fildes (1989) examined a specific area of time series forecasting and described what was learnt from forecasting competitions and compared the result with the expected result obtained from statistical theory. Markridakis (1989) examined the implications of forecast accuracy and proposed guidelines for implementation. Bunn (1991) carried out a judgmental and statistical study on the quality of judgment in forecasting and the possible structure available for facilitating interaction with existing statistical models. Cacatto et al. (2012) discussed the various forecasting methods that have been used by food industries in Brazil and how these forecasting methods were used.

3. Data Collection

The crime data used in this study was obtained on monthly basis and it is shown in table 1 below. The data shows the rate of crime committed in Ekiti State of Nigeria from January 2004 to December 2012. Fig1 shows the time plot of the crime data in table1. The time plot showcases the original crime figure plotted against time. The time plot reveals that there is decline in crime rate over the years.

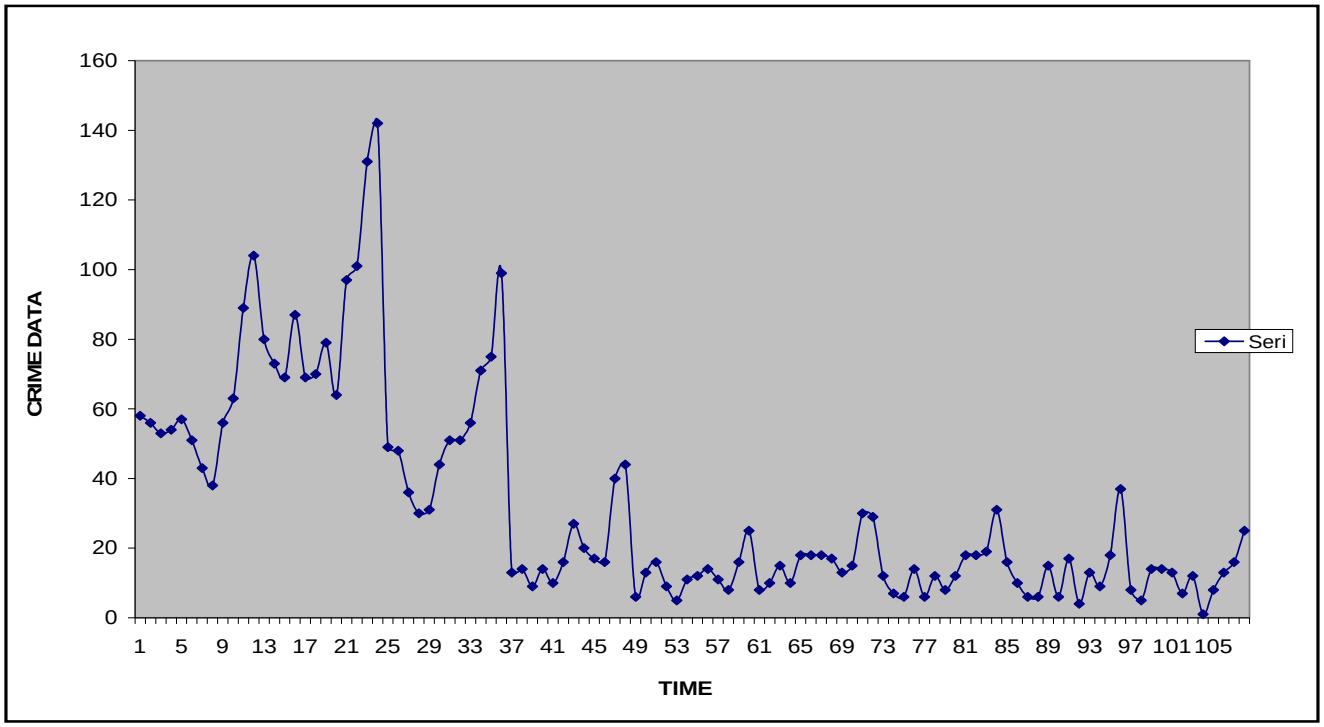

Figure 1: Time Plot of Crime Data

Month No	Crime Rate	Month No	Crime Rate	Month No	Crime Rate	Month No	Crime Rate	Month No	Crime Rate
1	58	23	131	45	17	67	18	89	15
2	56	24	142	46	16	68	17	90	6
3	53	25	49	47	40	69	13	91	17
4	54	26	48	48	44	70	15	92	4
5	57	27	36	49	6	71	30	93	13
6	51	28	30	50	13	72	29	94	9
7	43	29	31	51	16	73	12	95	18
8	38	30	44	52	9	74	7	96	37
9	56	31	51	53	5	75	6	97	8
10	63	32	51	54	11	76	14	98	5
11	89	33	56	55	12	77	6	99	14

Month No	Crime Rate	Month No	Crime Rate	Month No	Crime Rate	Month No	Crime Rate	Month No	Crime Rate
12	104	34	71	56	14	78	12	100	14
13	80	35	75	57	11	79	8	101	13
14	73	36	99	58	8	80	12	102	7
15	69	37	13	59	16	81	18	103	12
16	87	38	14	60	25	82	18	104	1
17	69	39	9	61	8	83	19	105	8
18	70	40	14	62	10	84	31	106	13
19	79	41	10	63	15	85	16	107	16
20	64	42	16	64	10	86	10	108	25
21	97	43	27	65	18	87	6		
22	101	44	20	66	18	88	6		

Table 1: Crime Data obtained in Ekiti State of Nigeria

4. Methodology

This study considered different forecasting model using Crime Data obtained from January 2004 to December, 2012 in Ekiti State of Nigeria. The forecast model used in the analysis included the Single Exponential smoothing method (α=0.3, α=0.5, α=0.7, α=0.9). Moving Average method (3-month, 5-month, 7-month, 9-month), the most suitable forecasting method was determined based on accuracy. Accuracy measures such as Mean Forecast Error, Mean Absolute Deviation (MAD), Mean Square Error (MSE), and Root Mean Square Error (RMSE) are the most common indicators often used.

5. Forecasting Methods

Forecasting method is the category of forecasting which is being applied there are a number of forecasting methods but the moving Average Method and the exponential smoothing method are employed in this work.

5.1. Moving Average Method

The method of obtaining a time series trend involves calculating a set of averages, each one corresponding to a trend (t) value for a time point of the series. These are known as moving averages, since each average is calculated by moving from one overlapping set of values to the next. The number of values in each set is always the same and is known as the period of the moving average. The procedure involves computing the average of observations and then using the computed average as the predictor for the next period. The moving average method depends on the number of terms selected for constructing the average. It is represented using the following equation.

$$F_{t+1} = \frac{Y_t + Y_{t-1} + Y_{t-2} + \cdots + Y_{t-n+1}}{n} \quad (1)$$

Where:
n = Number of terms in the moving average
Y_t= the exact value at period t
F_{t+1} = Forecast value for the next period

5.2. Exponential Smoothing Method

Exponential smoothing has been a popular forecast method for over half a century. It is commonly accepted that the method dates back to 1944 when R.G Brown used it to model the trajectories of bombs fired at submarines in World War II (Gardner, 2006). However, Brown's work did not appear in print until 1959. (Brown, 1959). At this time, the method was also being used independently to model series containing seasonal components. (Holt, 1957 and Winters, 1960)
Algorithm for univariate exponential method was initiated by Hyndman et al (2008) and was first introduced to literature by Athanasopoulos et al (2008). The algorithm provides a means of forecasting time series data by selecting from an extensive range of innovations state space models, which have been shown to generate optimal forecasts for all exponential smoothing methods.

The Exponential Smoothing technique uses weighted moving average of past data as the basis for the forecast. The basic Exponential Smoothing formular can be shown as follows:
New Forecast = Last Periods Forecasts + α(Last Period's actual value – Last Periods Forecast)
It can be represented mathematically as follows:

$$F_t = F_{t-1} + \alpha (A_{t-1} - F_{t-1}) \quad (2)$$

Where:
F_t = New Forecast Value
F_{t-1} = Previous Forecast Value
α = Smoothing Constant $(0 \leq \alpha \leq 1)$
A_{t-1} = Previous Period's actual value.

The accuracy of simple exponential smoothing method strongly depended on the optimal value of α. The preferred range for α in this study is from 0.3 to 0.9

6. Measuring Forecasting Error

According to (Hanke & Reitsch, 1998) the forecast error is the difference between an actual value and its forecast value. i.e. Forecast Error = Actual Value – Forecast Value. Accuracy is the criterion that determines the best forecasting method because there is no consensus among researcher as to which measure is best for determining the most appropriate forecasting method (Levine et al., 1999). Accuracy is thus the most important concern in evaluating the superiority of a forecast. The overall goal of the forecast is to minimize error. The most common indicators used to evaluate accuracy of a forecast are: Mean Forecast Error (MFE), Mean Absolute Deviation (MAD), Mean Square Error (MSE), Root Mean Square Error (RMSE).

6.1. Mean Square Error (MSE)

According to Jarrett (1991) Mean Square Error is a technique for evaluating exponential smoothing and other methods. It is represented by the following equation;

$$\text{MSE} = \frac{1}{n}\sum_{t=1}^{n}(Y_t - F_t)^2 \quad (3)$$

Where:
Y_t = the actual value in time period t
F_t = the forecast value in time period t
n = the number of periods

6.2. Root Mean Square Error (RMSE)

According to Jarrett (1991) Root Mean Square Error is the square root of MSE. This measures error in term of units that are equal to the original value. It is represented by the following equation.

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{t=1}^{n}(Y_t - F_t)^2} \quad (4)$$

Where:
Y_t = the actual value in time period t
F_t = the forecast value in time period t
n = the number of periods

6.3. Mean Forecast Error

Mean Forecast Error (MFE) is the mean of the deviation of the forecast demands from the actual demands. It is represented by the following equation.

$$\text{MFE} = \frac{\sum_{t=1}^{n}(Y_t - F_t)}{n} \quad (5)$$

Where:
Y_t = the actual value in time period t
F_t = the forecast value in time period t
n = the number of periods

6.4. Mean Absolute Deviation

A renowned method for measuring overall forecast error is the Mean Absolute Deviation (MAD). Heizer and Render (2001) stated that this value is computed by dividing the sum of the absolute values of the individual forecast error by the sample size (the number of forecast periods). It is represented by the following equation.

$$MAD = \frac{\sum_{t=1}^{n} |(Y_t - F_t)|}{n} \quad (6)$$

Where:
Y_t = the actual value in time period t
F_t = the forecast value in time period t
n = the number of periods

S/n	Method	MSE	RMSE	MFE	MAD
1	3-month Moving Average	131.761	11.479	-0.321	7.383
2	5-month Moving Average	235.144	15.334	-0.672	10.020
3	7-month Moving Average	294.947	17.174	-1.079	11.476
4	9-month Moving Average	314.880	17.745	-1.280	11.710
5	Single Exponential Smoothing (α=0.3)	306.938	17.520	-1.308	11.639
6	Single Exponential Smoothing (α=0.5)	278.391	16.685	-0.722	11.036
7	Single Exponential Smoothing (α=0.7)	263.283	16.226	-0.478	10.329
8	Single Exponential Smoothing (α=0.9)	262.662	16.207	-0.349	9.689

Table 2: Summary of Forecast Accuracy

7. Results and Discussion

The accuracy of the forecasting methods was measured using Mean Forecast Error (MFE), Mean Absolute Deviation (MAD), Mean Square Error (MSE) and Root Mean Square Error (RMSE). The smaller the forecast error, the more accurate the forecasting method. Table 2 shows the summary of the forecast accuracy for each of the forecasting methods while table 3 shows the overall ranking of the forecasting methods for the crime data.

S/n	Method	MSE	RMSE	MFE	MAD	Total Ranking	Overall Ranking
1	3-month Moving Average	1	1	1	1	4	1
2	5-month Moving Average	2	2	4	3	11	3
3	7-month Moving Average	6	8	6	6	24	6
4	9-month Moving Average	8	7	7	8	31	8
5	Single Exponential Smoothing (α=0.3)	7	5	8	7	29	7
6	Single Exponential Smoothing (α=0.5)	5	4	5	5	20	5
7	Single Exponential Smoothing (α=0.7)	4	3	3	4	15	4
8	Single Exponential Smoothing (α=0.9)	3	3	2	2	10	2

Table 3: Overall Ranking of Forecasting method for crime data

The result showed that the 3-month Moving Average was ranked first because it produced the most accurate forecasting based on the data obtained in table 3. i.e. It produced the smallest error i.e. (MSE=131.761, RMSE=11.479, MFE= -0.321, MAD=7.383) thus making it more efficient than others

The Single Exponential Smoothing (α=0.9) was ranked second, it produced the second smallest error i.e. (MSE=262.662, RMSE=16.207, MFE= -0.349, MAD=9.689) as shown in the ranking table 3.

The 5-month Moving average was ranked third because it produced the third smallest error i.e. (MSE=235.144, RMSE=15.334, MFE= -0.672, MAD=10.020) as shown in the ranking table 3.

The Single Exponential Smoothing (α=0.7) was ranked fourth, it produced the fourth smallest error i.e. (MSE=263.283, RMSE=16.226, MFE= -0.478, MAD=10.329) as shown in the ranking table 3.

The Single Exponential Smoothing (α=0.5) was ranked fifth, it produced the fifth smallest error i.e. (MSE=278.391, RMSE=16.685, MFE= -0.722, MAD=11.036) as shown in the ranking table 3.

The 7-month Moving average was ranked sixth because it produced the sixth smallest error i.e. (MSE=294.947, RMSE=17.174, MFE= -1.079, MAD=11.476) as shown in the ranking table 3.

The Single Exponential Smoothing (α=0.3) was ranked seventh because it produced the seventh smallest error i.e. (MSE=306.938, RMSE=17.520, MFE= -1.308, MAD=11.639) as shown in the ranking table 3.

The 9-month Moving average was ranked eight because it produced the eighth smallest error i.e. (MSE=314.880, RMSE=17.745, MFE= -1.280, MAD=11.710) as shown in the ranking table 3.

8. Conclusion and Recommendation

This research work uses moving average method and Exponential smoothing methods of forecasting on crime data and identified the most appropriate forecasting method based on accuracy and simplicity. The result showed that the 3-month Moving Average was ranked first and was therefore selected as the most appropriate forecasting method for the crime data obtained in this case study. It is therefore recommended for use in crime forecasting as a matter of policy consideration.

9. References

i. Athanasopoulos, G., Hyndman, R.J, Song, H. & Wu, D.C (2008), "The tourism forecasting competition, working paper. Department of Econometrics and Business Statistics, Monash University
ii. Brown R.G (1959), Statistical forecasting for inventory control, McGraw-Hill, New York.
iii. Bunn, D., & Wright, G. (1991) "Interaction of Judgmental and statistical forecasting method: issue and analysis " Management science, Vol. 37, Issue 5, pp 501-518.
iv. Cacatto, C., Belfiore, P., and Vieira, J.G.,(2012) "Forecasting Practices in Brazilian food industry, "Journal of logistics management: Vol. 1, No. 4, pp 24 – 36.
v. Fildes, R. (1989) Evaluation of aggregate and individual forecast method selection rules. Management Science, 35, 1056-1065.
vi. Gardner E.S (2006), 'Exponential smoothing: The state of the art- Part II', International journal of forecasting 22,637-666.
vii. Hanke, J. E., & Reitsch, A. G. (1998). Business forecasting (6th ed.). Upper Saddle River, NJ: Prentice-Hall.
viii. Heizer, J., & Render, B. (2001). Operations management (6th ed.). Upper Saddle River, NJ: Prentice-Hall.
ix. Holt, C.C (1957) Planning Production, Inventories and work force, Engle wood cliffs, New Jersey: Prentice Hall.
x. Hyndman R.J., Akram¸M. & Archibald B.C (2008) "The admissible parameter space for exponential smoothing models', Annals of the institute of statistical mathematics 60¸407 - 426
xi. Jarrett, J. (1991). Business forecasting methods (2nd ed.). Cambridge, MA: Basil Blackwell.
xii. Levine, D. M., Berenson, M. L., & Stephan, D. (1999). Statistics for managers using Microsoft@ Excel (znd ed.)Uvver Saddle River, NT: Prentice-Hall.
xiii. Markridakis S., et al., (1998) "Sliding simulation: a new approach to time series forecasting " management science, Vol. 89, Issue 51, pp 1 -12.
xiv. Pradeep K. and Rajesh K. (2013) "The Evaluation of Forecasting Methods for Sales of Salted Butter milk in Chhattisgarh, India", International Journal of Engineering Research & Technology, vol.2 issue 9.
xv. Ryu, K., and Sanchez, A., (2003) "The Evaluation of Forecasting Method at an Institutional Food Service Dining Facility, "Journal of Hospitality Financial Management: Vol. 11: issue 1, Article4.
xvi. Wilpen Gorr, Andreas Olligschlaeger, Yvonne Thompson (2003) "Short–term forecasting of Crime", International journal of forecasting 19(2003), pp 579-594.
xvii. Winters P.R (1960) 'Forecasting sales by exponentially weighted moving averages', management science 6, 324-342.