

REGIONAL CRIME ANALYSIS OF TURKEY: A TIME SERIES APPROACH

Yunus Emre KARAMANOĞLU¹ , Yılmaz AKDİ²

ABSTRACT

In this study, the most common types of crime in 7 different geographical regions of Turkey have been analyzed via time-series approach. The purpose of the study is; to put forth the most appropriate time series models for each geographical region and make forecasts for the crime rates. In the analysis; the crime data obtained from official sources covering the years 2001-2010 have been used. The stationarity of 7 regions has been tested and the most appropriate time series model for each region has been determined. The validity of the obtained time series models has been tested and the official data for 2011 have been compared with the forecasts made. As a result, importance and effectiveness of the forecasts obtained from the time series models were given. The availability of the time series approach to crime forecasting was shown for planning security.

Keywords: Time Series Analysis, Turkey's Crime Rates, Crime Prevention, Security, Forecast.

TÜRKİYE'NİN BÖLGESEL SUÇ ANALİZİ: ZAMAN SERİLERİ YAKLAŞIMI

ÖZ

Bu çalışmada, Türkiye'nin 7 farklı coğrafi bölgesindeki en yaygın suç türleri zaman serisi yaklaşımı ile analiz edilmiştir. Çalışmanın amacı; her coğrafi bölge için en uygun zaman serisi modelini ortaya koymak ve suç oranları için tahminler yapmaktır. Analizde; 2001-2010 yıllarını kapsayan resmi kaynaklardan elde edilen suç verileri kullanılmıştır. 7 bölge verisinin öncelikli olarak durağanlıkları test edilerek, her bölge için en uygun zaman serisi modeli tespit edilmiştir. Elde edilen zaman serisi modellerinin geçerliliği test edilmiş ve modellerin verdiği sonuçlar 2011 yılı resmi verileriyle karşılaştırılmıştır. Zaman serisi modellerinden elde edilen tahminlerin sonuçlarının önemi ve etkinliği ortaya konulmuştur. Zaman serileri yaklaşımının suçların öngörülebilirliğinde ve güvenlik planlamalarında kullanılabilirliği gösterilmiştir.

Anahtar Kelimeler: Zaman Serileri Analizi, Türkiye'nin Suç Oranları, Suç Önleme, Güvenlik, Öngörü.

¹ Lit.Col. (Ph.D.), Provincial Gendarmerie Command, Bitlis, Email:eyunus@bilkent.edu.tr (Corresponding Author).

² Ph.D., Department of Statistics, Ankara University, Tandoğan, Email:akdi@science.ankara.edu.tr.

1. INTRODUCTION

One of the essential tasks creating of a state is to ensure people to live in safety by preventing crimes and create a more peaceful community structure. To this end, the state should eliminate possible threats against the safety of citizens and enable the healthy development of the society (Bal, 2003). All of the countries give weight to the crime prevention strategies because of the damage crimes inflict on societies. Costs related with prevention measures taken by the countries reaches very high levels. However, despite crime prevention efforts the rates of crime is increasing day by day throughout the world (İçli, 2007).

The main point of the study is to analyze the crimes occurring in different regions of Turkey via time series analyzing methods. The reason why we need to study such a topic is the importance of crime prevention for Turkey. Peace and security is very important for the community and to ensure them, crime prevention strategies are one of the essential means (Kaygısız and Sever, 2006).

In an environment where safety and security is not provided; since the basic security needs of the society are not being provided, to achieve or to sustain the development progress will be difficult. Prevention of incidents before they occur or prediction of the events makes the law enforcement be ready for the duties and in terms of citizen means to live in a peaceful society. This idea is the motivation of this study.

The aim of this study is to give regional crime models by taking into account the creation process of historical data. By the help of obtained regional crime time series models, consistent forecasting for the future can be done. Another aim of the study is to bring a different point of view to policy makers for planning their security strategies in the future. In this study; time series models are going to be put forward for problems encountered in different geographical regions of Turkey by using time series analysis methods in the field of crime prevention.

In the countries; control and supervision of crime are given to criminal justice and law enforcement institutions. This task includes; preventing crime, arresting and prosecution of criminals quickly, execution of the sentence, social reintegration of the offender back to society (Sutherland and Cressey, 1966). Factors affecting the internal security of the countries include terrorism, organized crime, smuggling, attacks on virtual environment, drugs, illegal migration, human trafficking and traffic offenses (Williams and McShane, 1999).

Although crime and crime preventive strategies are the subjects of the social sciences, in recent years, different disciplines have begun to give importance to this field. For this reason, many important scientific research has been made in the area of crime prevention strategies.

To find the right solution to the problems encountered, accurate modeling and analysis needs to be done. Crime prevention strategies for modeling are also important. In accordance with past data to reveal the structures and the creation of models, time series analysis of the data comes to the forefront.

Time series analysis is a source of information and method (Göktaş, 2005). Time series analysis is one of the most widely used methods and may contribute to solution of the problems encountered in social sciences. By looking at problems from different disciplinary perspectives, routines can be demolished and creative and more effective solutions can be found for the existing problems. The creation of new models for the crime and making crime forecasts by these models develops as the time series analysis studies gain popularity.

The steps taken to prevent crimes are important. Threats to effective crime prevention planning in the regions should be studied well. Revealing the correct data structure of the past has a crucial property for efficient forecasts. By means of accurate forecasts, security planning will be successful.

Desired result of this study is to forecast rates of crime in different regions and to assist decision-makers for their crime preventing strategies. Crime rates have been studied in 7 different geographical regions of Turkey. For each region; crimes (*Most common crimes are included in the study. These are "Murder", "Intentional injury", "Coercion and blackmail", "Maltreatment", "Theft/Burglary", "Offence against property", "Injury/kill with traffic accident".* The data including the crime rates (2001-2011) is obtained from Turkish Statistical Institute and Turkish General Command of Gendarmerie) occurred between the years 2001-2010 have been analyzed by time series analysis on a monthly basis. In the analysis, the appropriate time series models are given for each region, the validations of the models are tested and forecasts have been done. Consistency of the forecasts obtained from the time series models has been compared and checked by the official data of 2011.

In the second chapter, the basic concepts of time series are given. In the third chapter, a literature review is presented. In the fourth chapter, a short review of the applied method; in the fifth section, time series graphs and correlograms of the regions are given. The appropriate time series models for each regions, the projections obtained from these models and their comparisons with the official data are given in the sixth section.

2. TIME SERIES ANALYSIS

Time series concept, although has been growth in 1970, mainly has emerged as a method after 1940 with the book of Box-Jenkins named "Time Series Analysis: Forecasting and Control".

Time series are analyzed for various purposes. One of the purposes of the analysis of series is making predictions about the future and this is the most used purpose of the time series (Gujarati, 1999). Other aims can be considered as;

- Revealing the major features of the series,
- Explaining the relationship between the series,

- Revealing the functioning mechanism of the event that creates the series or ensuring the improvement and control of the system in scheduled direction by using the information obtained from the past events of the system (Göktaş, 2005).

A time series is a sequence of data points, measured typically at successive points in time spaced at uniform time intervals. Examples of time series are; amount of product exported monthly from a factory, number of accidents that occur weekly on a highway, monthly precipitation of a city (Akdi, 2010). According to another definition; a time series is the series that created by the arrangement of raw data according to time units such as day, month, year. (Çilingirtürk, 2011).

Time series are used in statistics, signal processing, pattern recognition, econometrics, mathematical finance, weather forecasting and earthquake prediction.

Time series analysis comprises methods for analyzing time series data in order to extract meaningful statistics and other characteristics of the data. While being used in the applications, it is necessary that the time series have a trusted extension into the past (Göktaş, 2005).

Time series forecasting is the use of a model to predict future values based on previously observed values. While regression analysis is often employed in such a way as to test theories that the current value of one time series affects the current value of another time series, this type of analysis of time series is not called "time series analysis" (Akdi, 2010).

a. Stationarity

One of the important concepts in time series is the stationarity. Prior to the statistical analysis of series; it is necessary to put out the process that makes up that series whether constant over time or not. Stationarity of series is usually seen as an assumption. Almost all the statistical inference is done under the assumption of stationarity of the series. Focusing of the stationarity of the time series can be explained simply as follows: If you have a stationary

stochastic process, the behavior of the series is valid only for the period under consideration; it cannot be made a generalization for the series about other periods. A stationary process has the property that the mean, variance and autocorrelation structure do not change over time.

Before carrying out a statistical analysis of time series, the stationarity of the process that the series are created should be investigated whether it is constant over time or not. When an analysis is conducted with non-stationary time series, traditional t and F tests and R^2 value can give biased results (Göktaş, 2005).

If a time series is nonstationary, we can study its behavior only for the time period under consideration. Each set of time series data will therefore be for a particular episode. As a consequence, it is not possible to generalize it to other time periods. Therefore, for the purpose of forecasting, such (nonstationary) time series may be of little practical value (Gujarati, 1999).

A stochastic process is said to be stationary if its mean and variance are constant over time and the value of the covariance between the two time periods depends only on the distance or gap or lag between the two time periods and not the actual time at which the covariance is computed. In the time series literature, such a stochastic process is known as a weakly stationary, or covariance stationary, or second-order stationary, or wide sense, stochastic process (Gujarati, 1999).

A time series $\{x_t : t \in T\}$ is said to be stationary if it satisfies the following conditions:

$$E(X_t) \text{ is free of time,}$$

$$\text{Cov}(X_t, X_s) \text{ is only a function of } |t - s|.$$

If a time series is not stationary in the sense just defined, it is called a nonstationary time series (weak stationarity). In other words, a nonstationary time series will have a time varying mean or a time-varying variance or both.

b. Unit Root

A test of stationarity (or nonstationarity) that has become widely popular over the past several years is the unit root test. Systematically investigating the existence of unit root test in a series has been put forward by the Dickey and Fuller (1979). Dickey and Fuller developed unit root test to test for stationarity of time series in their (1979) and (1981) studies. Unit root series are the series with AR component. The most commonly used unit root tests of Dickey-Fuller indicates whether the parameters of time series can be expressed with autoregressive variables (AR) or not.

Let;
$$X_t = \alpha + \beta X_{t-1} + u_t \tag{1}$$

is given.

If we subtract X_{t-1} from both sides of the equation (1), new model can be seen from (2) and (3).

$$X_t - X_{t-1} = \alpha + \beta X_{t-1} - X_{t-1} + u_t \tag{2}$$

$$(X_t - X_{t-1}) = \alpha + (\beta - 1)X_{t-1} + u_t \tag{3}$$

Testing hypotheses is determined as follows;

H_0 : There is a unit root ($\beta = 1$) series is not stationary.

H_1 : There is no unit root ($\beta < 1$) series is stationary.

If the equation (1) is an auto regressive process (AR) and coefficient of X_{t-1} is equal to one ($|\beta|=1$) it shows the presence of unit root.

A stochastic process is purely random if it has zero mean, constant variance σ^2 , and is serially uncorrelated. u_t is assumed to be a white noise process which is independently and identically distributed as a normal distribution with zero mean and constant variance.

Dickey-Fuller uses three different regression equations for the existence of unit root.

$$\Delta X_t = \gamma X_{t-1} + u_t \quad (4)$$

$$\Delta X_t = \alpha_0 + \gamma X_{t-1} + u_t \quad (5)$$

$$\Delta X_t = \alpha_0 + \gamma X_{t-1} + \beta t + u_t \quad (6)$$

In the equations where t is the time or trend variable; (4) is the random walk; equation (5) is the random walk with drift; equation (6) shows the state with drift around a stochastic trend.

In each case, the null hypothesis is that $\gamma = 0$; that is, there is a unit root—the time series is nonstationary. The alternative hypothesis is that γ is less than zero; that is, the time series is stationary. If the null hypothesis is rejected, it means that X_t is a stationary time series with zero mean in the case of (4), that X_t is stationary with a nonzero mean in the case of (5), and that X_t is stationary around a deterministic trend in (6).

Dickey and Fuller have shown that under the null hypothesis that $\gamma = 0$ the estimated t value of the coefficient of X_{t-1} in (4) follows the τ (tau) statistic. These authors have computed the critical values of the tau statistic on the basis of Monte Carlo simulations.

c. Determining the Appropriate Time Series Model

Autocorrelation coefficients of a time series refer to the strength of the relationship between the different periods of observation. If the coefficients of the variables are high that means previous period is dependent on each other; if they are low that indicates they are random variables. Autocorrelation function (ρ_k) shows any correlation between the observation values of the series; namely it means the dependence of each other.

The partial correlation between two variables is the correlation that remains if possible impact of all other random variables has been eliminated. To define the partial autocorrelation coefficient, another notation is being used;

$$x_t = \phi_{k1}x_{t-1} + \phi_{k2}x_{t-2} + \dots + \phi_{kk}x_{t-k} + u_t.$$

where ϕ_{ki} is the coefficient of the variable with lag "i" if the process order "k". The coefficients ϕ_{kk} are the partial autocorrelation coefficients (of order k), $k=1,2,\dots$. The partial autocorrelation measures the correlation between x_t and x_{t-k} which remains the influences of $x_{t-1}, x_{t-2}, \dots, x_{t-k+1}$ on x_t and x_{t-k} have been eliminated (Kirchgassner and Wolters, 2007).

The analysis of the correlogram of the autocorrelation and partial autocorrelation functions is used in determining the Autoregressive (AR) and Moving Average (MA) models. Because autoregressive process explains a variable with its lagged values, partial autocorrelation function is used to determine the order of the process. Autocorrelation coefficients are used in the determination of the order of the moving average model (Akdi, 2010).

An autocorrelation and partial autocorrelation functions of a stationary time series may approach to zero value slowly. In this case, it can be said that time series model have both AR and MA process at the same time (Akdi, 2010).

3. LITERATURE REVIEW

There are studies in which time series analysis, panel data studies and panel cointegration analysis are applied to different type of crime data. As a general evaluation, studies were largely carried out in order to contribute to econometrics literature. The vast majority of the studies are related to identify the relations between the society's economic, social, demographic, justice and security features with various crimes. Questions whose answers investigated are; "What is the relationship between social structure and crimes? How the effects of economic development to a particular crime can be modeled? What are the effects of the demographic structure to the crime types? How does unemployment rate affect crime amounts? What are the effects of income inequality to crime rates?".

Looking at the general framework, common property of the econometric models given in this problem field aims to give a formula for understanding the problems encountered. In this study, we aimed to give rather than an econometric model; we prefer to give a time-series model. Economic modeling approach to crime has gained momentum after the 1960s. The model put forward in the article ["Crime and Punishment: An Economic Approach" (Crime and Punishment: an Economic Approach)] published in 1968 by Gary Becker is accepted as the first model in the economic literature. Following the work of Becker, a large literature has emerged examining potential relationships between crime rates and a number of economic variables.

Identifying and modeling the scientific elements of the crime helps to produce more effective policies to combat with crime. Although modeling crime as an economic model is a new field of study; recently, many economists have developed "econometric models" and in these models especially the relation between crime and econometrics is kept at the forefront. In reviewing the literature on the economic models of crime, important models are; The Traditional Economic Model, The Econometric Model, The Expectations Theory Model and The Expected Utility Model (Akkuş, 2003).

Sanlı (1998), studied the structure of criminality in Turkey according to socio-economic factors and grouped the provinces according to crime regions. Study was based on the socio-economic indicators showing the socio-ecological indicators of provinces and indicators of criminal offenders in prisons; the relation between multivariate statistical analysis and two variables data set was determined.

Beki, Zeelenberg, and Montfort (1999) in their study for the Netherlands, analyzed the tendencies of various crimes by using the data including years 1950-1993. Deadman's (2000) study for England; econometric and time series analysis has been used to estimate the trend of burglary from houses.

Andrienko (2001), for Russia, using the data between the period 1990-1998 belong to 70 settlement aimed to find out the factors which affect the crime trends. Sookram et al. (2010), using the time series data obtained from Trinidad and Tobago, examined the major crimes rates and long-term cointegration relationship between the various socio-economic indicators.

Cook and Cook (2011), in their study of "Are US Crime Rates Really Unit Root Processes?"; showed the unit root hypothesis to be rejected for all classifications of criminal activity examined over the period 1960-2007.

Aslan and Ocal (2012), have done a study for the continuity of crime rates in Turkey, in their study they used the unit root test for heterogeneous panels, the continuity of the crime rates for the provinces have been analyzed and the convergence of crimes were studied.

4. APPLIED METHOD

Since the stationary of a time series is important, the entire series is tested whether they are stationary or not. Firstly; model and model parameters are estimated, secondly; tests have been done for the given models. As a last step we have used the models for forecasting the crime rates. Stationary test is applied the equations (4), (5) and (6) respectively. If the significance levels of the critical values of the ADF test values of variables are greater than absolute value of the series, the series is stationary and there is no unit root. If the absolute value is less than the critical value, the series has a unit root and it is not stationary. In this case, the first difference of the series on the stationary test is applied again. Process continues until the series is stationary.

5. RESULTS

Testing stationary of time series EViews-7 and obtaining time series model and for predictions SAS 9.1 programs were used. After the stationary test; models for each region were obtained.

In modeling of time series, autocorrelations and partial autocorrelations are the most important tools. After a certain delay, if the autocorrelations have the value around zero, modeling this kind of data as MA (Moving Average) is suitable. After a certain delay, if partial autocorrelations have the value of zero then modeling this kind of data as AR (Auto Regressive) is suitable. If autocorrelation follows a certain periodicity, the seasonal time series model could be proposed for this type of time series data. If the autocorrelations of any time series are different from the given above, it would be appropriate for modeling the data as an ARMA model. For appropriate time series models especially (AR) models were taken into account because (AR) models are more sensitive for forecasts (Akdi, 2010).

The time-series graphs, correlograms of autocorrelation and partial autocorrelation of the series are given in Figure-1, 2, 3, 4, 5, 6 and 7. Data between years 2001 and 2010 for Region-1 has been given in Figure-1. In each year, crime amounts increase in months June, July and August. ACF graph has a periodical characteristic. According to partial ACF graph of Region-1; first, second, seventh, eighth, tenth, eleventh and thirteenth lags are outside of the borders. These lags were taken into account while obtaining appropriate time series (AR) model.

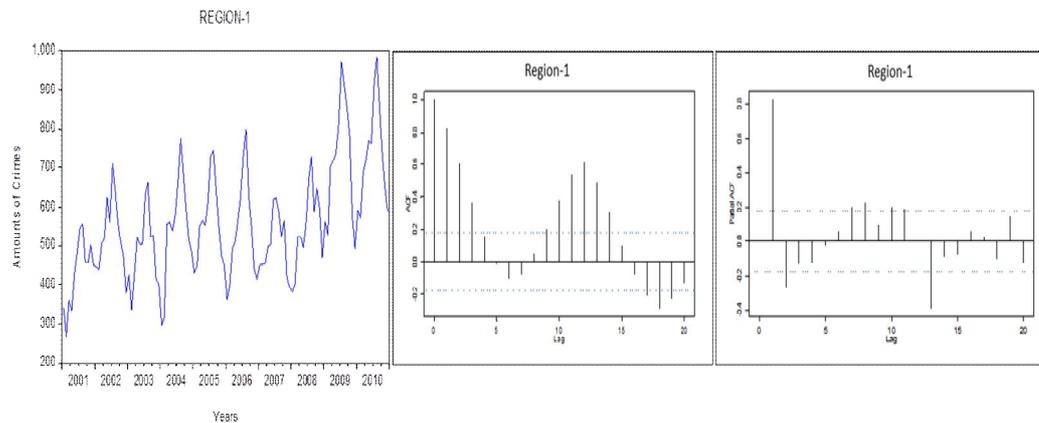


Figure-1: Time Series Graph and Correlogram of Region-1.

Data between years 2001 and 2010 for Region-2 has been given in Figure-2. As seen from ACF graph Region-2 has periodical characteristic. According to partial ACF graph of Region-2; first, second, third, fourth, seventh, ninth, tenth and eleventh lags are outside of the borders. These lags were taken into account while obtaining appropriate time series (AR) model.

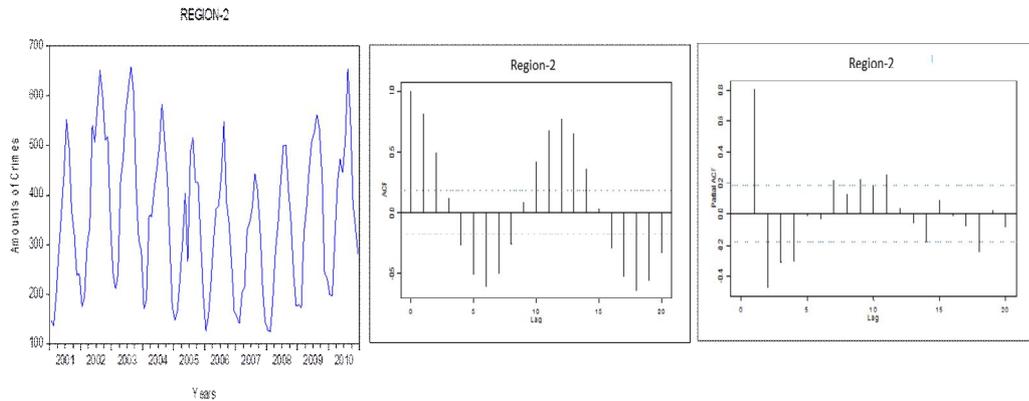


Figure-2: Time Series Graph and Correlogram of Region-2.

The other graphs of regions and data have been examined with the same manner. They have been given in Figure-3, 4, 5, 6 and 7.

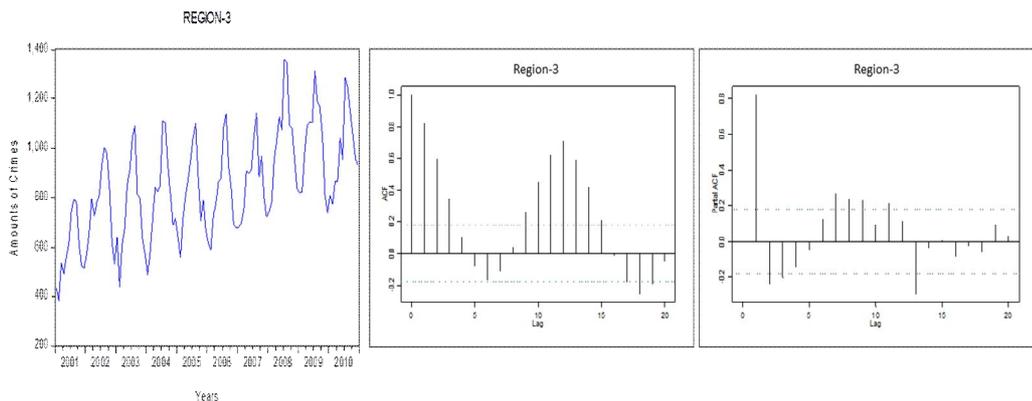


Figure-3: Time Series Graph and Correlogram of Region-3.

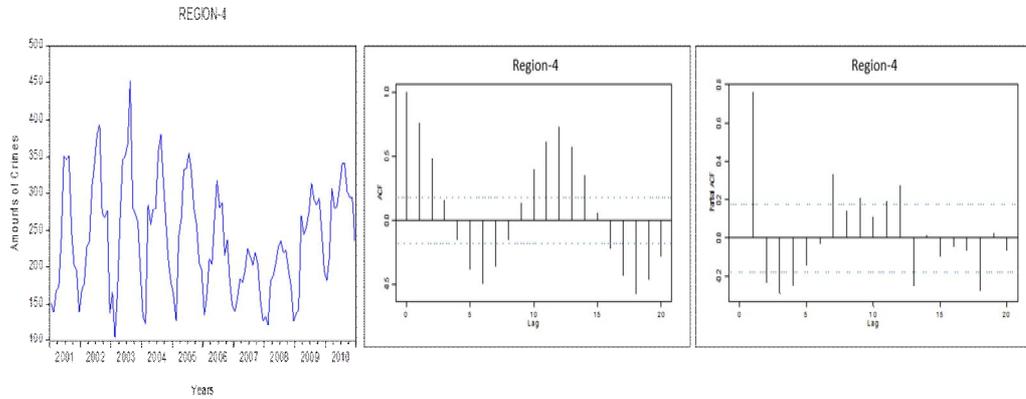


Figure-4: Time Series Graph and Correlogram of Region-4.

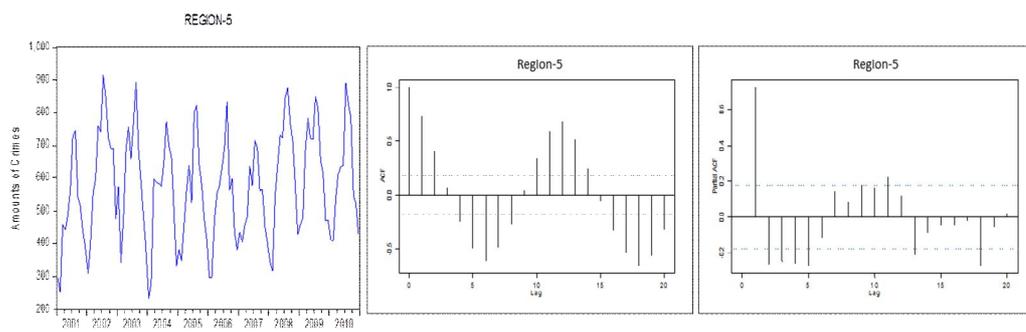


Figure-5: Time Series Graph and Correlogram of Region-5.

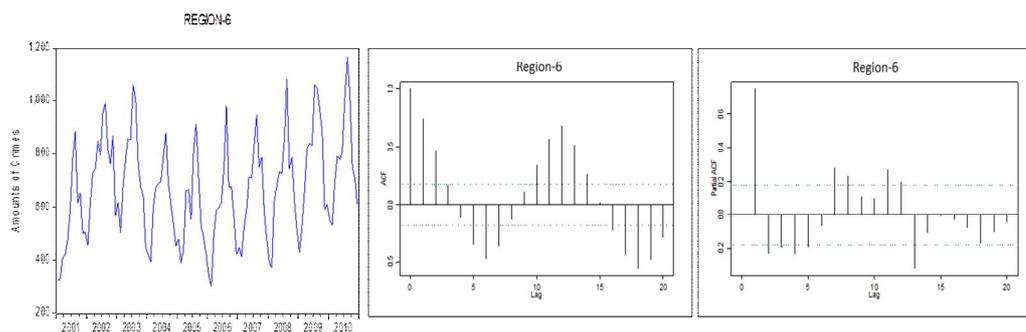


Figure-6: Time Series Graph and Correlogram of Region-6.

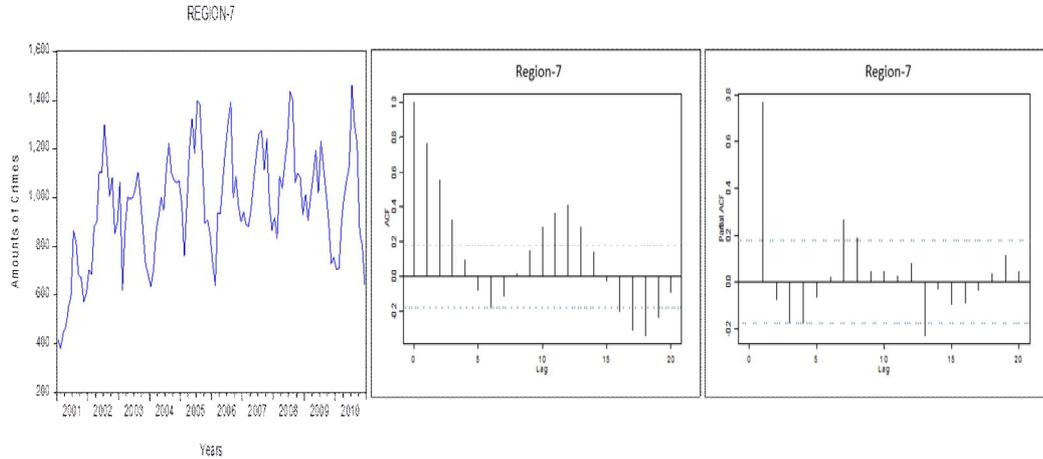


Figure-7: Time Series Graph and Correlogram of Region-7.

6. THE TIME SERIES MODELS FOR EACH REGION

According to the values of AIC (Akaike Info Criterion) and SBC (Schwarz Bayesian Criterion) statistics, the best model for the i^{th} region

$$X_{i,t} = \alpha_1 X_{i,t-1} + \alpha_2 X_{i,t-2} + \dots + \alpha_p X_{i,t-p} + e_t$$

is proposed.

In the proposed models; $e_t \sim WN(0, \sigma^2)$ and $X_{i,t}$ is the numbers of crime in the i^{th} region at the t^{th} month.

The plots of Autocorrelation Functions (ACF) and Partial Autocorrelation Functions (PACF) are also support the validity of the model. The model parameters have been estimated by *PROC ARIMA* in SAS:

a. Time Series Model for Region-1

$$\hat{X}_{i,t} = -0.83745X_{i,t-1} + 0.01874X_{i,t-2} + 0.1374X_{i,t-7} - 0.11687X_{i,t-8} - 0.02213X_{i,t-10} - 0.4169X_{i,t-11} + 0.23722X_{i,t-13}$$

After the determinations of the models and estimation procedures have been completed, we observed that for all regions the crime rates are stationary according to the ADF (Augmented Dickey Fuller) results. The values of the ADF test results have been tabulated for each region in Table A1 in the Appendix.

After obtaining the appropriate time series models for the regions, second part of the study has been done. Since the proposed model is stationary, it is meaning to make forecast for the future. We calculated the values of monthly forecasts for the 2011. The data of the year 2011 has been obtained from official sources at the completion stage of the study. Forecasts and the official data of the year 2011 are given in the tables and graphics below (Figure-8, 9, 10, 11, 12, 13 and 14).

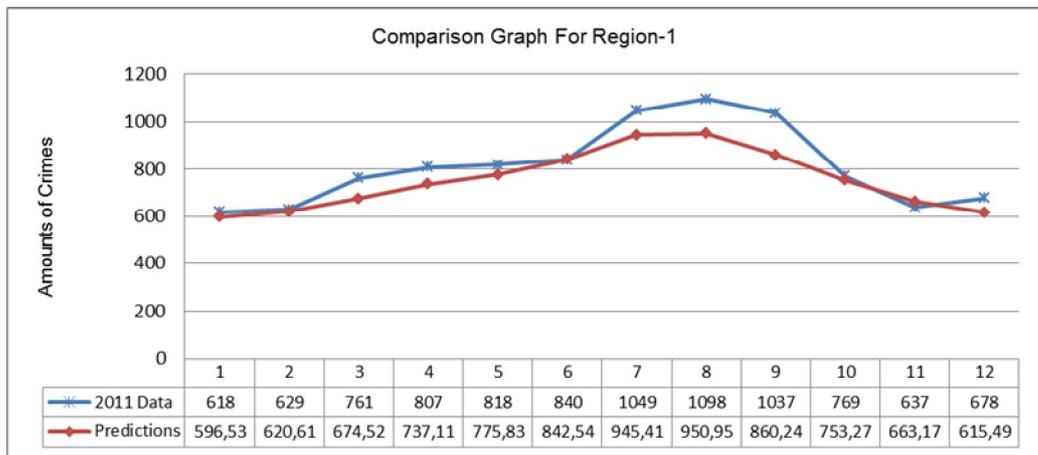


Figure-8: Results for Region-1 and comparison with the official data of 2011.

Table-1: Forecasted Number of Crimes for Region-1.

2011 Data	Forecasting	Std.Error	95% Conf.Interval	
618	604.0531	64.2709	478.0845	730.0217
629	621.1506	83.8317	456.8435	785.4577
761	661.9654	94.6171	476.5193	847.4116
807	715.2356	101.1391	517.0065	913.4646
818	753.2379	105.2426	546.9662	959.5096
840	835.6696	107.8785	624.2317	1047.1075
1049	929.5596	109.5919	714.7633	1144.3558
1098	903.9266	109.8084	688.7061	1119.1470
1037	810.9954	109.9473	595.5027	1026.4881
769	723.4266	110.0503	507.7320	939.1212
637	654.9520	110.1888	438.9859	870.9180
678	627.3651	114.9384	402.0901	852.6402

In the following part of our study; since the steps for obtaining appropriate time series models are similar; firstly, the models are given; secondly, forecast results and the comparison with the official data are given below for each region.

b. Time Series Model for Region-2

$$\hat{X}_{i,t} = -0.77873X_{i,t-1} + 0.07717X_{i,t-2} - 0.1317X_{i,t-3} + 0.25684X_{i,t-4} - 0.08707X_{i,t-7} - 0.00293X_{i,t-9} - 0.33291X_{i,t-11}$$

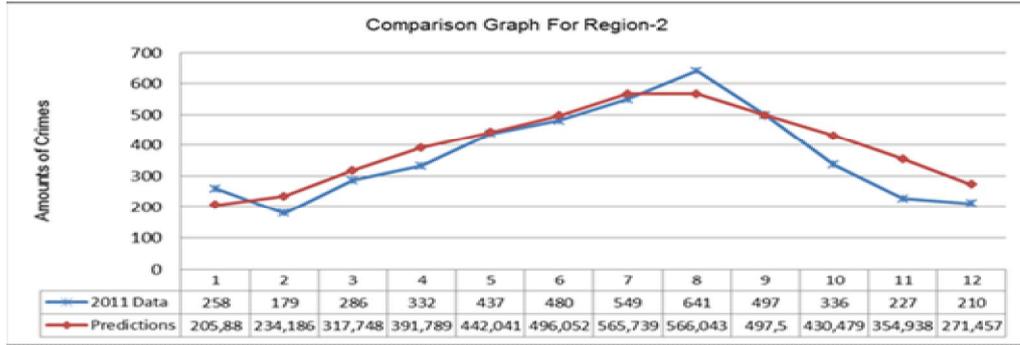


Figure-9: Results for Region-2 and comparison with the official data of 2011.

Table-2: Forecasted Number of Crimes for Region-2.

2011 Data	Forecasting	Std.Error	95% Conf.Interval	
258	205.8795	58.9518	90.3361	321.4230
179	234.1861	74.7182	87.7411	380.6311
286	317.7480	80.9708	159.0482	476.4478
332	391.7893	85.8461	223.5341	560.0445
437	442.0409	86.5110	272.4825	611.5993
480	496.0515	86.5248	326.4661	665.6369
549	565.7387	86.7529	395.7061	735.7713
641	566.0437	86.9308	395.6625	736.4250
497	497.4999	86.9759	327.0303	667.9694
336	430.4789	86.9787	260.0036	600.9541
227	354.9378	87.0834	184.2575	525.6180
210	271.4574	90.5944	93.8957	449.0192

In the first column of Table-2, crime numbers occurred in 2011 is given. For example, the number of crimes according to the official records in June is 258. The time series model for this region gives the forecasted number of crimes as 205. Standard error is 58,95 and confidence interval (95%) is between 90,33 and 312,42. For another example, the amount of crimes according to the official records in September is 497. The model gives the forecasted number of crimes as 497. Standard error is 86,97 and confidence interval (95%) is between 327,03 and 667,96. As seen from the Figure-9, the shape of the official crime data graph and forecasted amounts of crimes are approximately same.

Especially in the fifth, sixth, seventh and ninth months of 2011 the amount of crimes occurred and forecasted amount by the time series model are very close to each other.

c. Time Series Model for Region-3:

$$\hat{X}_{i,t} = -0.73514 X_{i,t-1} - 0.01151 X_{i,t-2} + 0.0186 X_{i,t-3} + 0.09953 X_{i,t-7} - 0.00827 X_{i,t-8} - 0.12157 X_{i,t-9} - 0.36734 X_{i,t-11} + 0.1257 X_{i,t-13}$$

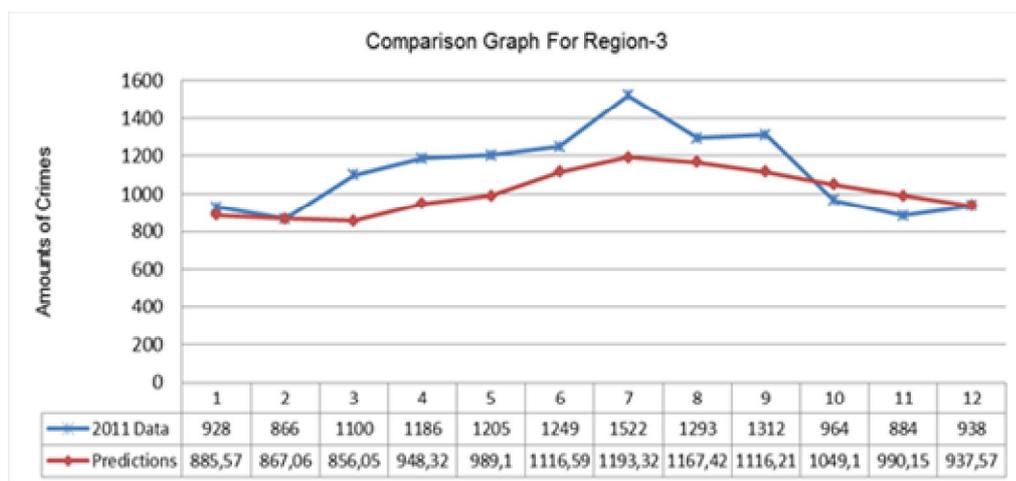


Figure-10: Results for Region-3 and comparison with the official data of 2011.

Table-3: Forecasted Number of Crimes for Region-3.

2011 Data	Forecasting	Std.Error	95% Conf.Interval	
928	885.5703	93.5293	702.2563	1068.8844
866	867.0677	116.0831	639.5489	1094.5865
1100	856.0519	127.0441	607.0500	1105.0539
1186	948.3197	132.3228	688.9717	1207.6677
1205	989.1014	134.9535	724.5973	1253.6055
1249	1116.5926	136.2788	849.4911	1383.6940
1522	1193.3209	136.9512	924.9016	1461.7403
1293	1167.4242	136.9517	899.0039	1435.8446
1312	1116.2078	137.0825	847.5310	1384.8846
964	1049.1699	137.0994	780.4600	1317.8797
884	990.1536	137.2586	721.1317	1259.1755
938	937.5674	143.9248	655.4800	1219.6548

In the first column of Table-3, crime numbers occurred in 2011 is given. For example, the number of crimes according to the official records in June is 928. The time series model for this region gives the forecasted number of crimes as 885. Standard error is 93,53 and confidence interval (95%) is between 702,26 and 1068,88. For another example, the amount of crimes according to the official records in December is 938. The model gives the forecasted number of crimes as 937. Standard error is 143,93 and confidence interval (95%) is between 655,48 and 1296,65.

In the first, second and twelfth months of 2011 the amount of crimes occurred and forecasted amount by the time series model are very close to each other.

d. Time Series Model for Region-4:

$$\hat{X}_{i,t} = -0.61133X_{i,t-1} - 0.05883X_{i,t-2} + 0.03797X_{i,t-3} + 0.03677X_{i,t-4} - 0.16113X_{i,t-7} - 0.03301X_{i,t-8} - 0.08722X_{i,t-11} - 0.41816X_{i,t-12} + 0.25105X_{i,t-13} + 0.27284X_{i,t-18}$$

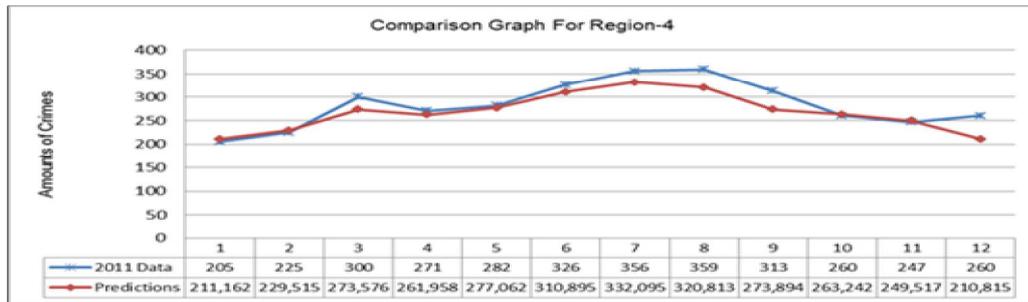


Figure-11: Results for Region-4 and Comparison With The Official Data of 2011.

Table-4: Forecasted Number of Crimes for Region-4.

2011 Data	Forecasting	Std.Error	95% Conf.Interval	
205	211.1616	34.7199	143.1119	279.2112
225	229.5153	40.6937	149.7571	309.2735
300	273.5764	43.3766	188.5600	358.5929
271	261.9577	44.3232	175.0859	348.8296
282	277.0624	44.5382	189.7691	364.3556
326	310.8953	44.5769	223.5262	398.2645
356	332.0951	44.5797	244.7206	419.4696
359	320.8126	44.9183	232.7744	408.8508
313	273.8941	45.5784	184.5621	363.2262
260	263.2416	46.2895	172.5158	353.9674
247	249.5170	46.8358	157.7205	341.3134
260	210.8146	47.5784	117.5628	304.0665

In the first column of Table-4, crime numbers occurred in 2011 is given. For example, the amount of crimes according to the official records in June is 205. The time series model for this region gives the forecasted number of crimes as 211. Standard error is 34,72 and confidence interval (95%) is between 143,12 and 279,21.

The amount of crimes according to the official records in November is 247. The model gives the forecasted number of crimes as 249. Standard error is 47,58 and confidence interval (95%) is between 117,56 and 304,07. In the first, second, fourth, fifth, tenth and eleventh months of 2011 the amount of crimes occurred and forecasted amount by the time series model are very close to each other. The results obtained from the model for Region-4 are very close to the official crime data as seen from Figure-11.

e. Time Series Model for Region-5:

$$\hat{X}_{i,t} = 0.6792 X_{i,t-1} + 0.02044 X_{i,t-2} + 0.07192 X_{i,t-3} + 0.0271 X_{i,t-4} - 0.16718 X_{i,t-5} + 0.11359 X_{i,t-9} + 0.3615 X_{i,t-11} - 0.11764 X_{i,t-13}$$

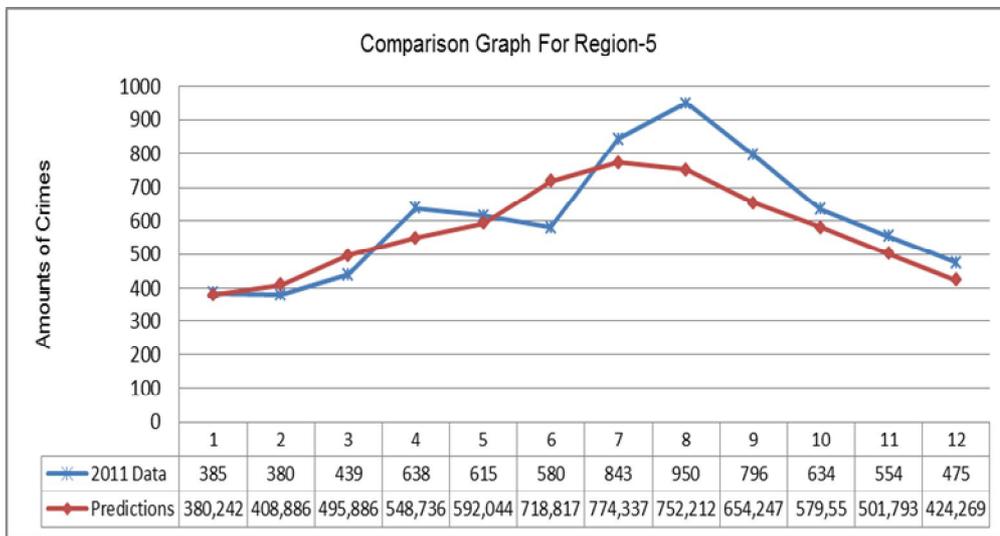


Figure-12: Results for Region-5 and comparison with the official data of 2011.

Table-5: Forecasted Number of Crimes for Region-5.

2011 Data	Forecasting	Std.Error	95% Conf.Interval	
385	380.2420	89.6986	215.2320	566.8439
380	408.8664	108.4321	205.2528	630.2989
439	495.8861	116.7256	267.1238	724.6799
638	548.7362	122.4635	330.5030	810.5513
615	592.0441	126.7949	374.9925	872.0194
580	718.8171	127.4431	508.5456	1008.1133
843	774.3373	127.4794	553.4904	1053.2003
950	752.2123	127.4885	531.9187	1031.6646
796	654.2473	127.6010	428.0426	928.2294
634	579.5501	127.6110	346.2118	846.4376
554	501.7930	127.7390	262.7041	763.4320
475	424.2694	133.9503	174.5935	699.6689

In the first column of Table-5, crime numbers occurred in 2011 is given. For example, the amount of crimes according to the official records in June is 385. The time series model for this region gives the forecasted number of crimes as 380. Standard error is 89,7 and confidence interval (95%) is between 215,23 and 566,84.

The amount of crimes according to the official records in March is 439. The model gives the forecasted number of crimes as 495. Standard error is 116,73 and confidence interval (95%) is between 267,12 and 724,68.

In the first and fifth months of 2011 the amount of crimes occurred and forecasted amount by the time series model are very close to each other.

f. Time Series Model for Region-6

$$\hat{X}_{i,t} = -0.63943 X_{i,t-1} - 0.08464 X_{i,t-2} - 0.00427 X_{i,t-3} + 0.0122 X_{i,t-4} + 0.09045 X_{i,t-5} + 0.03248 X_{i,t-7} - 0.11765 X_{i,t-8} - 0.14541 X_{i,t-11} - 0.50249 X_{i,t-12} + 0.35876 X_{i,t-13}$$

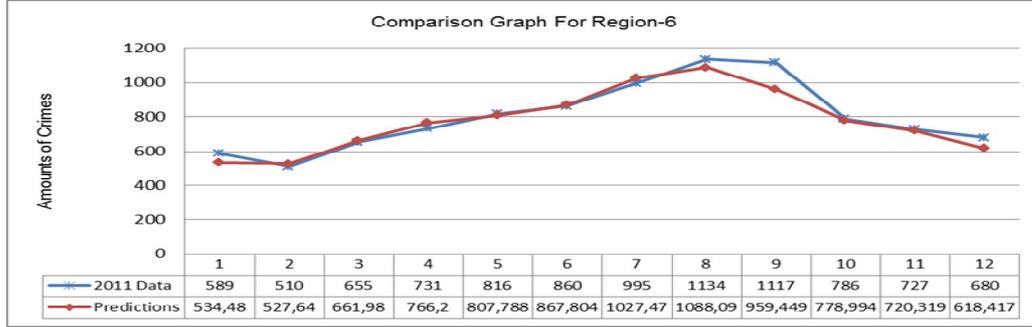


Figure-13: Results for Region-6 and Comparison With The Official Data of 2011.

Table-6: Forecasted Number of Crimes for Region-6.

2011 Data	Forecasting	Std.Error	95% Conf.Interval	
589	534.4804	94.1966	349.8584	719.1024
510	527.6482	111.8074	308.5096	746.7867
655	661.9804	121.0865	424.6552	899.3056
731	766.2020	126.1063	519.0383	1013.3657
816	807.7888	128.6718	555.5966	1059.9810
860	867.8041	129.0813	614.8095	1120.7987
995	1027.4249	129.1131	774.3678	1280.4819
1134	1088.0878	129.2054	834.8499	1341.3256
1117	959.4488	129.2356	706.1518	1212.7458
786	778.9936	129.3147	525.5414	1032.4458
727	720.3194	129.4765	466.5502	974.0886
680	618.4168	131.2238	361.2228	875.6107

In the first column of Table-6, crime numbers occurred in 2011 is given. For example, the amount of crimes according to the official records in February is 510. The time series model for this region gives the forecasted number of crimes as 527. Standard error is 111,81 and confidence interval (95%) is between 308,51 and 746,79. Crime amount according to the official records in June is 860. The model gives the forecasted number of crimes as 867.

In the second, third, fifth, sixth, tenth and eleventh months of 2011 the amount of crimes occurred and forecasted amount by the time series model are very close to each other. The results obtained from the model for Region-6 are very close to the official crime data as seen from Figure-13.

g. Time Series Model for Region-7

$$\hat{X}_{i,t} = 0.91328 X_{i,t-1} + 0.13268 X_{i,t-3} - 0.30161 X_{i,t-4} + 0.01122 X_{i,t-7} + 0.24756 X_{i,t-8} - 0.0097 X_{i,t-13}$$

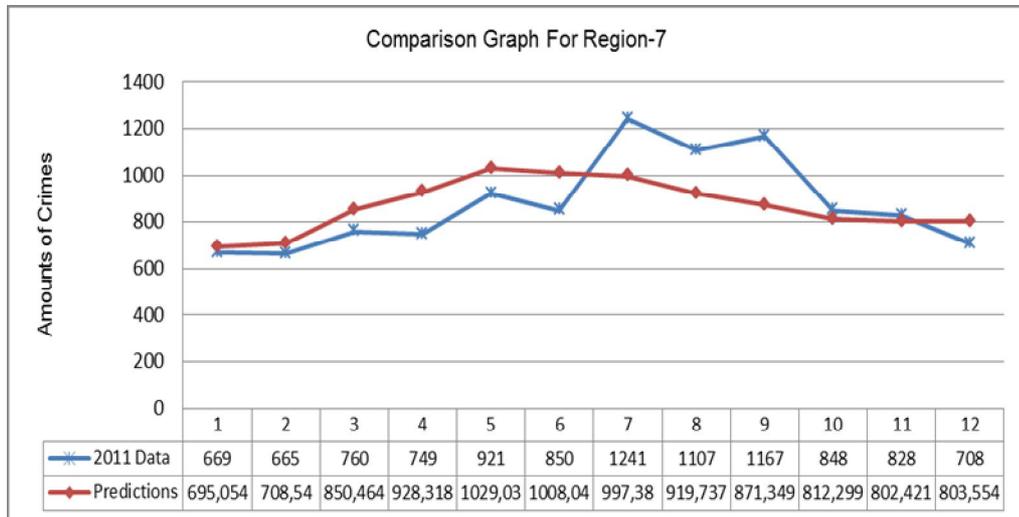


Figure-14: Results for Region-7 and Comparison With The Official Data of 2011.

Table-7: Forecasted Number of Crimes for Region-7.

2011 Data	Forecasting	Std.Error	95% Conf.Interval	
669	606.0876	131.3836	348.5805	863.5948
665	694.2221	177.9306	345.4846	1042.9597
760	848.2893	208.9691	438.7173	1257.8613
749	990.3986	239.7448	520.5074	1460.2899
921	1119.5676	253.9080	621.9171	1617.2180
850	1140.7464	259.7362	631.6728	1649.8200
1241	1116.1806	261.7627	603.1351	1629.2261
1107	1024.1222	261.8514	510.9029	1537.3415
1167	897.5779	262.7853	382.5282	1412.6277
848	796.6162	265.6579	275.9364	1317.2961
828	742.8432	271.3178	211.0702	1274.6163
708	741.8832	283.2253	186.7719	1296.9945

In the first column of Table-7, crime numbers occurred in 2011 is given. The amount of crimes according to the official records in August is 1107. The time series model for this region gives the forecasted number of crimes as 1024. Crime amount according to the official records in December is 708. The model gives the forecasted number of crimes as 741. In the first, second, tenth and eleventh months of 2011 the amount of crimes occurred and forecasted amount by the time series model are very close to each other.

7. CONCLUSION

This paper investigates and introduces a practical approach of using time series to forecast crime rates. The first responsibility of each state is to ensure safety and prevent crime. While obtaining crime prevention strategies, we need an accurate analysis of past data. Trend of the data, relationships between them, influence each other and by the help of these factors accurate forecasts can be done for the future.

For the security units, it is important to predict the number of crimes that might be happened in the near future. The unit whose real aim is to prevent crime can take the necessary measures through predictive capabilities. The study considering the rural areas of the seven geographical regions in Turkey particularly contributes to the literature because it takes into account the rural areas. There are studies usually taking into account the police's responsibility areas or urban centers. One of the reasons that the studies concentrate on the city centers is the ease of the data obtained. Another reason is the population density in the cities.

In this study, time series analysis has been applied to historical data of events against public order occurred in different regions of Turkey in period 2001-2011. Time series models suitable to each region have been obtained with the aid of these data. With the help of time series models, forecasts have been done. The forecasts have been compared with the official data. Determining the amount of crime in future periods of these regions is important for planning the security measures to be taken in this study area. The results of the time-series models obtained for the first, second, fourth and sixth (Figure-8,9,11 and 13) regions are very close to real crime data. It is seen from the graphs that especially the predictions for the fourth and sixth region obtained are very close to the actual crime data.

The security policies for governments are very crucial and to provide against crimes lots of expedients are taken into account. Since time series analysis can be used to obtain forecasts for the future with the help of

historical data; we aimed to give a different point of view for security policy makers by using time series approach for forecast. Examining the amount of the events against public order occurred in our country's rural areas, revealing the time series model based on the amount of crime it will help the effectiveness and efficiency of the security units whose main purpose is to prevent crime.

Analysis to be performed with the help of preventive measures against crime, especially by type of crime, the predictions can be made in this direction and this will provide development of law enforcement units themselves and provide the community with a safer environment. Using of the result of the analyses and results will contribute to prevention of crimes and controlling the crimes, to finding new strategies for the development of the security units, to determine the specialties that security units must have in the future and to planning the personnel and sources appropriately.

The time series models and results of predictions obtained from the study will show new ways for planning new strategies in crime preventions and planning that will be made by law-security units. Using scientific analyses means growing the efficiency and productivity of the security units as for each foundation. As a result, moral and material damages caused by crime in the society are going to be minimized and safety expectations of the public will be met effectively.

Finally, this paper advised that security policy makers to be aware of this kind of analysis in the crime preventing strategies.

ACKNOWLEDGMENTS

All the views expressed in this paper belong to the authors and do not represent the views of Turkish Gendarmerie, Gendarmerie Schools Command or its staff.

We would like to thank for the editor and two anonymous referees for giving us an opportunity to revise our manuscript based on their suggestions.

REFERENCES

- Akdi, Y.(2010). **Zaman Serileri Analizi: Birim Kökler ve Kointegrasyon**, Gazi Kitapevi, Ankara.
- Akkuş, Y. (2003). **Suçun Ekonomik Analizi: Türkiye Örneği**, Atatürk Üniversitesi Sosyal Bilimler Enstitüsü Yüksek Lisans Tezi, Erzurum.
- Andrienko, Y. (2001). **Explaining Crime Growth in Russia During Transition: Economic and Criminometric Approach**, Centre for Economic and Financial Research, Moscow, February.
- Aslan, A., Öcal, O. (2012). **Türkiye’de Suç Oranlarının Devamlılığı**, Niğde Üniversitesi, İktisadi ve İdari Bilimler Dergisi, **5(1)** 85-92.
- Bal, A.M. (2003). **Modern Devlet ve Güvenlik**, IQ Kültür Sanat Yayıncılık, İstanbul.
- Beki, C., K. Zeelenberg and N. Fielding. (1999). **An Analysis of the Crime Rate in the Netherlands, 1950-1993**, British Journal of Criminology, Oxford, **39(3)** 401-415.
- Cook, J., Cook, S. (2011). **Are US Crime Rates Really Unit Root Processes?**, Journal of Quantitative Criminology, **27** 299-314.
- Dickey, D. A. and Fuller, W. A. (1979). **Distribution of the Estimators for Autoregressive Time Series with a Unit Root**, Journal of the American Statistical Association **74**.

Deadman, D. (2000). **Forecasting Residential Burglary**, Public Sector Economics Research Centre, Department of Economics, University of Leicester, February.

Göktaş, Ö. (2005). **Teorik ve Uygulamalı Zaman Serileri Analizi**, Beşir Kitapevi, İstanbul.

Gujarati, D.N. (2003). **Basic Econometrics**, McGraw-Hill, New-York.

İçli, T. (2007). **Kriminoloji**, Seçkin Yayınevi, 7. Baskı, Ankara.

Kaygısız, M., Sever, H. (2006). **Suç Analizi**, Adalet Yayınevi, Ankara.

Kirchgassner, G. and Wolters, J. (2007). **Introduction to Modern Time Series Analysis**, New York, Springer.

Sanlı, S. F. (1998). Türkiye'de **Suçların Yapısını Etkileyen Sosyo-Ekonomik Faktörler** (Basılmamış yüksek lisans tezi), Ankara, Devlet İstatistik Enstitüsü.

Sookram, S., Basdeo, M., Sumesar-Rai, K., Saridakis, G. (2010). **Serious Crime in Trinidad and Tobago: An Empirical Analysis Using Time-Series Data between 1970-2007**, Journal of Eastern Caribbean Studies, **35**(1), March.

Sutherland, E.H., Cressey, D.R. (1966). **Principles of Criminology**, J.B. Lippincott Company, New York.

Williams, F.P., McShane, M.D. (1999). **Criminological Theory**, USA: Prentice Hall.

APPENDIX

Table A1: Seasonal Unit Root Test Results³.

Regions ⁴	Lag	ADF (None)	Critical Values (%1, %5, %10)			ADF (Intercept)	Critical Values (%1, %5, %10)			ADF (Intercept and Trend)	Critical Values (%1, %5, %10)		
			-2.586753	-1.943853	-1.614749		-3.492523	-2.888669	-2.581313		-3.492523	-2.888669	-2.581313
Region-1	1	-3.491809	-2.586753	-1.943853	-1.614749	-3.674406	-3.492523	-2.888669	-2.581313	-3.700911	-4.046072	-3.452358	-3.151673
			-2.586350	-1.943796	-1.614784	-9.530497	-3.491345	-2.888157	-2.581041	-9.479919	-4.044415	-3.451568	-3.151211
			-2.586753	-1.943853	-1.614749	-4.688885	-3.492523	-2.888669	-2.581313	-4.742223	-4.046072	-3.452358	-3.151673
Region-2	1	-4.296026	-2.586753	-1.943853	-1.614749	-4.599893	-3.492523	-2.888669	-2.581313	-4.665432	-4.046072	-3.452358	-3.151673
			-2.586753	-1.943853	-1.614749	-4.599893	-3.492523	-2.888669	-2.581313	-4.665432	-4.046072	-3.452358	-3.151673
			-2.586753	-1.943853	-1.614749	-4.599893	-3.492523	-2.888669	-2.581313	-4.665432	-4.046072	-3.452358	-3.151673
Region-3	1	-4.571116	-2.586753	-1.943853	-1.614749	-4.645565	-3.492523	-2.888669	-2.581313	-4.680710	-4.046072	-3.452358	-3.151673
			-2.586753	-1.943853	-1.614749	-4.645565	-3.492523	-2.888669	-2.581313	-4.680710	-4.046072	-3.452358	-3.151673
			-2.586753	-1.943853	-1.614749	-4.645565	-3.492523	-2.888669	-2.581313	-4.680710	-4.046072	-3.452358	-3.151673
Region-4	1	-4.637666	-2.586960	-1.943882	-1.614731	-3.326160	-3.493129	-2.888932	-2.581453	-3.295566	-4.046925	-3.452764	-3.151911
			-2.586960	-1.943882	-1.614731	-3.326160	-3.493129	-2.888932	-2.581453	-3.295566	-4.046925	-3.452764	-3.151911
			-2.586960	-1.943882	-1.614731	-3.326160	-3.493129	-2.888932	-2.581453	-3.295566	-4.046925	-3.452764	-3.151911
Region-5	1	-3.337469	-2.585773	-1.943714	-1.614834	-9.018054	-3.489659	-2.887425	-2.580651	-9.169162	-4.042042	-3.450436	-3.150549
			-2.585773	-1.943714	-1.614834	-9.018054	-3.489659	-2.887425	-2.580651	-9.169162	-4.042042	-3.450436	-3.150549
			-2.585773	-1.943714	-1.614834	-9.018054	-3.489659	-2.887425	-2.580651	-9.169162	-4.042042	-3.450436	-3.150549

³ In each table, ADF values and their critical values of three models (*model without intercept, with intercept, with intercept and trend*) are given. When ADF value of the model is bigger than the critical value first time, it is written in bold.

⁴ Region-1: *Mediterranean*; Region-2: *Eastern Anatolian*; Region-3: *Aegean*; Region-4: *Southeastern Anatolia*; Region-5: *Central Anatolia*; Region-6: *Blacksea* and Region-7: *Marmara*.