Zoran Ivanovski, Ace Milenkovski, and Zoran Narasanov. 2018. Time Series Forecasting Using a Moving Average Model for Extrapolation of Number of Tourist. *UTMS Journal of Economics* 9 (2): 121–132.

Preliminary communication
(accepted May 5, 2018)

TIME SERIES FORECASTING USING A MOVING AVERAGE MODEL FOR EXTRAPOLATION OF NUMBER OF TOURIST

Zoran Ivanovski[1]
Ace Milenkovski
Zoran Narasanov

Abstract
Time series is a collection of observations made at regular time intervals and its analysis refers to problems in correlations among successive observations. Time series analysis is applied in all areas of statistics but some of the most important include macroeconomic and financial time series. In this paper we are testing forecasting capacity of the time series analysis to predict tourists' trends and indicators. We found evidence that the time series models provide accurate extrapolation of the number of guests, quarterly for one year in advance. This is important for appropriate planning for all stakeholders in the tourist sector. Research results confirm that moving average model for time series data provide accurate forecasting the number of tourist guests for the next year.

Keywords: seasonality, trend, regression, forecasting, centered moving average.

Jel Classification: C3; C32

INTRODUCTION

Contemporary business has to deal with uncertainty and operations in terms of complex and dynamic changes of business environment. Companies and public stakeholders use planning function in order to organize their work and provide necessary assets for the future periods. One of the most important parts of planning function is financial planning, as a continuous process of directing and allocating financial resources to meet companies' strategic goals and objectives. Financial planning usually starts with planning company's sales (revenues) that will be realized in next period.

[1] **Zoran Ivanovski**, PhD, Full Professor; **Ace Milenkovski**, PhD, Full Professor, University of Tourism and Management in Skopje; **Zoran Narasanov**, PhD, Assistant Professor, Winner Insurance, Vienna Insurance Group, Macedonia.

In tourist industry forecasting future revenues depends on capability to make prediction of number of tourists, having in mind that there are available statistics' data for average tourist spending per day. When all stakeholders in tourist industry have relevant and accurate forecasts of trends and number of tourists that will visit the country, they can provide reliable input for their financial planning. In order to do that it is important to use quantitative methods and statistics tools that can help to deal with past data where seasonality, irregularity and sometimes random variables appears.

There are many comprehensive theoretical studies in literature that suggest the necessity to use quantitative models and techniques for various types of data. In this paper we focus on a time series analysis. A time series is a collection of observations made sequentially in time. Examples are daily mortality counts, particulate air pollution measurements, and temperature data. Time series analysis refers to problems in which observations are collected at regular time intervals and there are correlations among successive observations as explained by Bartholomew (1971), Janacek (2009) and many other authors. Applications cover virtually all areas of statistics but some of the most important include macroeconomic time series as analyzed by Nelson and Plosser (1982), like monthly data for unemployment, hospital admissions, etc., in finance (eg. daily exchange rate, a share price, etc.), environmental (e.g., daily rainfall, air quality readings); medicine (e.g., ECG brain wave activity every 2^{-8} seconds).

Time series analysis include autoregressive and moving average processes and some discussion of the effect of time series correlations on the other kinds of statistical inference, such as the estimation of means and regression coefficients. The methods of time series analysis pre-date those for general stochastic processes and Markov Chains. The aims of time series analysis are to describe and summarize time series data, fit low-dimensional models, and make forecasts. One simple method of describing a series is a classical decomposition into four elements: Trend (Tt) — long term movements in the mean; Seasonal effects (It) — cyclical fluctuations related to the calendar; Cycles (Ct) — other cyclical fluctuations (such as a business cycles); Residuals (Et) — other random or systematic fluctuations. The idea is to create separate models for these four elements and then combine them, either additively:

$$X_t = T_t + I_t + C_t + E_t \tag{1}$$

or multiplicatively

$$X_t = T_t \cdot I_t \cdot C_t \cdot E_t \tag{2}$$

One of the most popular and frequently used stochastic series models is the Autoregressive Integrated Moving Average model (ARIMA) elaborated by many authors like Chatfield (1996), Box and Jenkins (1971), Cochrane (1997), Cottrell et al. (1995). One of the most used ARIMA model is Moving Average Model introduced by Armstrong (2006).

In this paper we are testing the ability and accuracy of application of the time series analysis model for prediction of some important tourist trends and indicators. Many authors have analyzed accuracy of time series analysis like Fildes and Makridakis (1995), Chernick (1994), Montgomery, Jennings and Kulahci (2008). Gardner and Mckenzie (1985) in their paper argue that forecast accuracy can be improved by either damping or ignoring altogether trends which have a low probability of persistence. We are testing research hypothesis that moving average model for the time series data provide accurate forecasting of the number of tourist guests in next year.

The remainder of this paper is structured into three sections. In Section 1, we present a theoretical review and methodology. Section 2 presents research findings from a time series analysis using moving average model and last section summarizes the main conclusions.

1. LITERATURE REVIEW AND METHODOLOGY

There are many methods available for forecasting, so it is important to know their applicability and reliability to make appropriate selection before using in specific situation. Time series modeling is a very popular tool among researchers and practitioners for providing accurate forecast.

In fact, main task of time series modeling is to analyze past observations in order to develop an appropriate model for the specific data series and to use this as a model for forecasting future values for the series. A time series is a set of observations on the values that a variable takes at different times, collected at regular time intervals, such as monthly (eg. CPI), weekly (eg. money supply), quarterly (eg. GDP) or annually (eg. State budget). Time series models are used for the modelling and prediction of data collected sequentially in time in order to provide specific techniques for handling data (Brockwell and Davis, 2013) and Diggle (1990).

Depending on the frequency of the data (hourly, daily, weekly, monthly, quarterly, annually, etc) different patterns emerge in the data set which forms the component to be modeled. Sometimes the time series may just be increasing or decreasing over time with a constant slope or they may be patterns around the increasing slope (Bollerslev 1986). The pattern in a time series is sometimes classified into trend, seasonal, cyclical and random components. We can define trend as a long term relatively smooth pattern that usually persist for more than one year. Seasonal is a pattern that appears in a regular interval wherein the frequency of occurrence is within a year or even shorter, as it can be monthly arrival of tourist to a country. Cyclical means the repeated pattern that appears in a time-series but beyond a frequency of one year. It is a wavelike pattern about a long-term trend that is apparent over a number of years. Cycles are rarely regular and appear in combination with other components (eg. business cycles that record periods of economic recession and inflation, cycles in the monetary and financial sectors, etc.). Random is the component of a time-series that is obtained after these three patterns have been "extracted" out of the series. When we plot the residual series we can indicate a random pattern around a mean value.

Time series modeling is a dynamic research area which has become popular not only in academic circles but also widely used by business sector over last few decades. The main aim of time series modeling is to carefully collect and rigorously study the past observations of a time series to develop an appropriate model which describes the inherent structure of the series. This model then can be used to generate future values for the series, i.e. to make forecasts as Adhikari and Agrawal argue in their paper (2013). Idea for univariate time series modeling based on situations where an appropriate economic theory to the relationship between series may not be available and hence we consider only the statistical relationship of the given series with its past values.

Sometimes even when the set of explanatory variables may be known it may not be possible to obtain the entire set of such variables required to estimate a regression model and we can use only a single series of the dependent variable to forecast the future values.

There are different time series processes:

- White Noise: A series is called white noise if it is purely random in nature. Let $\{\varepsilon_t\}$ denote such a series then it has zero mean [E (ε_t)=0], has a constant variance [V $(\varepsilon_t)=\sigma^2$] and is an uncorrelated [E $(\varepsilon_t\ \varepsilon_s)$=0] random variable. The plot of such series across time will indicate no pattern and hence forecasting the future values of such series is not possible.
- Auto Regressive Model: An AR model is one in which Y_t depends only on its own past values Y_{t-1}, Y_{t-2}, Y_{t-3}, etc. Thus:

$$Y_t=f\ (Y_{t-1}, Y_{t-2}, Y_{t-3},\ldots\ \varepsilon_t). \tag{3}$$

A common representation of an autoregressive model where it depends on p of its past values called AR(p) model is:

$$Y_t=\beta_0 + \beta_1 Y_{t-1}+ \beta_2 Y_{t-2}+ \beta_3 Y_{t-3}+\ldots+ \beta_p Y_{t-p}+\varepsilon_t \tag{4}$$

- Moving Average Model – is one when Y_t depends only on the random error terms which follow a white noise process:

$$Y_t=f\{\varepsilon_{t,}\ \varepsilon_{t-1,}\ \varepsilon_{t-2,}\ \varepsilon_{t-3\ldots\ldots}\varepsilon_{t-n}\}. \tag{5}$$

A common representation of a moving average model where it depends on a *q* of its past values is called MA(q) model and represented below:

$$Y_t=\beta_0 + \varepsilon_t+\phi\ \varepsilon_{t-1}+ \phi\ \varepsilon_{t-2}+ \phi\ \varepsilon_{t-3}+\ldots+ \phi_q\ \varepsilon_{t-q} \tag{6}$$

The error terms ε_t are assumed to be white noise processes with mean zero and variance σ^2.

The simplest time series model is certainly the white noise. We use in our research MA(*q*) processes to linear processes and time series analysis of quarterly data for number of tourists in the Republic of Macedonia using actual past data for five years period (2012–2016) in order to provide extrapolation of number of guest that will visit the Republic of Macedonia in next year. In contrast to modeling in terms of mathematical equation, the moving average merely smooths the fluctuations in the data. A moving average model works well when the data have a fairly linear trend and a definite rhythmic pattern of fluctuations.

2. RESEARCH FINDINGS

We analyze quarterly data for total number of guests for the five years period from 2012–2016 that have visited the Republic of Macedonia. Official state statistics regularly provide monthly data for number of guests in the country. We summarize official data and calculate quarterly data for actual number of guests that have visited the Republic of Macedonia in five years period (2012–2016) as shown on next Table:

Table 1. Quarterly data for number of tourists in the Republic of Macedonia 2012–2016

Year	Quarter	Num. of guests
2012	Q_1	89.383
	Q_2	164.691
	Q_3	291.895
	Q_4	117.664
2013	Q_1	91.858
	Q_2	177.790
	Q_3	303.054
	Q_4	129.092
2014	Q_1	94.404
	Q_2	191.570
	Q_3	318.709
	Q_4	130.967
2015	Q_1	98.709
	Q_2	205.766
	Q_3	367.811
	Q_4	143.781
2016	Q_1	115.222
	Q_2	215.509
	Q_3	370.437
	Q_4	155.675

Source: State Statistical Office

The task of our research is to use a time series models for extrapolation number of guests, quarterly for year 2017. Forecasting number of guests enables appropriate planning in tourist sector (government, ministries, agencies, tour-operators, travel agencies, hotels and other relevant stakeholders).

In order to visualize data series, we plot them and create line chart:

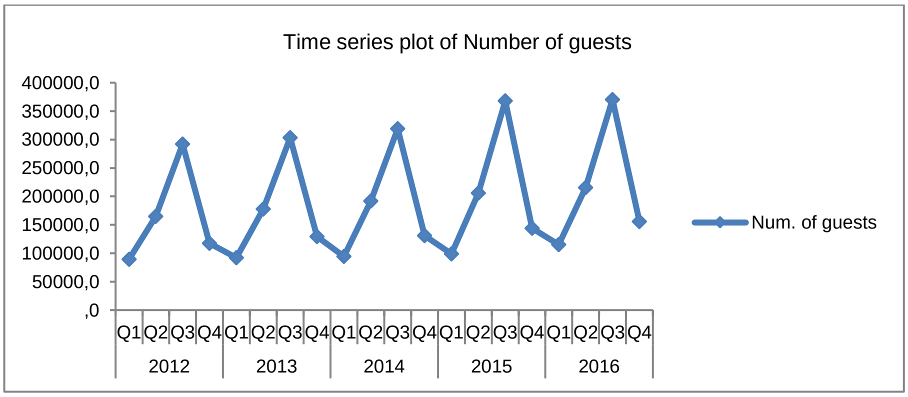

Figure 1. Time Series of number of tourist in the Republic of Macedonia 2012–2016

Analysis usually starts with charts where it is easier to visualize movement and behaviour of actual numbers. We can see from the chart that there is “pattern” that commonly repeats itself almost every year. There are “depths” and a “climaxes” in every period (years), in fact cycles repeats itself every year and this is clear example of seasonality. Beside this “up and down” movements, we can conclude that the overall direction of this plot is increasing, and that is trend component of actual data.

This mean that trend was clearly identified as well as the seasonal components in the actual numbers of the guests that have visited the Republic of Macedonia. Beside trend and seasonal component we need to check "irregularity" of actual data. The final component in the model that we are using here is going to be irregular component. That actually means that quarter to quarter variations of number of guests that exist does not follow any pattern. We can conclude that actual numbers have irregular aspect, also called "random aspect". That component is always present in data, no matter whether they are time series or not. So we have to deal with variability of time series data. This is usually a movement of variables that do not have predictable pattern in real world. In fact we need to deal with all three components: trend, seasonality and irregularity. In order to use time series analysis, next step is to "smooth out" our data and for that we will use moving averages above four periods MA(4). It is necessity to use MA(4) because we have identified cycles with four periods (quarters).

Moving averages calculation MA(4) are presented in final right column on the following Table 2:

Table 2. Moving averages calculation MA(4)

T	Year	Quarter	Num. of guests	MA (4)
1	2012	Q_1	89.383	
2		Q_2	164.691	
3		Q_3	291.895	165.908
4		Q_4	117.664	166.527
5	2013	Q_1	91.858	169.802
6		Q_2	177.790	172.592
7		Q_3	303.054	175.449
8		Q_4	129.092	176.085
9	2014	Q_1	94.404	179.530
10		Q_2	191.570	183.444
11		Q_3	318.709	183.913
12		Q_4	130.967	184.989
13	2015	Q_1	98.709	188.538
14		Q_2	205.766	200.813
15		Q_3	367.811	204.017
16		Q_4	143.781	208.145
17	2016	Q_1	115.222	210.581
18		Q_2	215.509	211.237
19		Q_3	370.437	214.211
20		Q_4	155.675	216.186

Next step of our time series data analysis is the centered moving averages calculation CMA(4). Results are presented and added to previous data in last right column, as shown on Table 3:

Table 3. Centered moving averages calculation CMA(4)

t	Year	Quarter	Num. of guests	MA (4)	CMA (4)
1	2012	Q_1	89.383		
2		Q_2	164.691		
3		Q_3	291.895	165.908	166.218
4		Q_4	117.664	166.527	168.164
5	2013	Q_1	91.858	169.802	171.197
6		Q_2	177.790	172.592	174.020
7		Q_3	303.054	175.449	175.767
8		Q_4	129.092	176.085	177.808

. . . .

Table 4. (*continued*)

t	Year	Quarter	Num. of guests	MA (4)	CMA (4)
9	2014	Q_1	94.404	179.530	181.487
10		Q_2	191.570	183.444	183.678
11		Q_3	318.709	183.913	184.451
12		Q_4	130.967	184.989	186.763
13	2015	Q_1	98.709	188.538	194.676
14		Q_2	205.766	200.813	202.415
15		Q_3	367.811	204.017	206.081
16		Q_4	143.781	208.145	209.363
17	2016	Q_1	115.222	210.581	210.909
18		Q_2	215.509	211.237	212.724
19		Q_3	370.437	214.211	215.198
20		Q_4	155.675	216.186	

This step was necessary to “smooth” the time series data and to “clean” the data from the seasonality and irregularity. Now we plot the CMA(4) data on the chart:

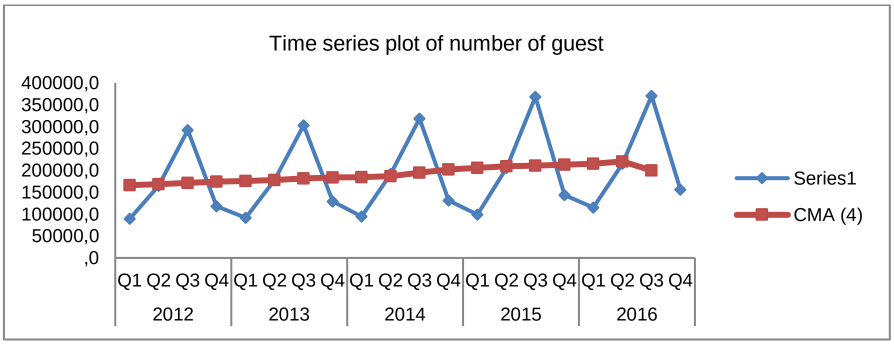

Figure 2. Time Series of number of tourist in the Republic of Macedonia 2012–2016

CMA(4) line is plotted on the chart with red color. This line in fact extract seasonal and irregular components and now we can clearly see the difference between original data and CMA(4) data. We proceed with our analysis and create new column S_t,I_t for seasonal and irregular components. This is necessary in order to “follow the model” requirements Table 4. As already explained in theoretical approach in first part of this paper, classical time series multiplicative model is:

$$Y_t = S_t \times I_t \times T_t \tag{7}$$

Table 5. S_t,I_t calculation

t	Year	Quarter	Num. of guests	MA (4)	CMA (4)	S_t,I_t
1	2012	Q_1	89.383			
2		Q_2	164.691			
3		Q_3	291.895	165.908	166.218	1,76
4		Q_4	117.664	166.527	168.164	0,70
5	2013	Q_1	91.858	169.802	171.197	0,54
6		Q_2	177.790	172.592	174.020	1,02
7		Q_3	303.054	175.449	175.767	1,72
8		Q_4	129.092	176.085	177.808	0,73

. . . .

Table 6. (*continued*)

t	Year	Quarter	Num. of guests	MA (4)	CMA (4)	S_t,I_t
9	2014	Q_1	94.404	179.530	181.487	0,52
10		Q_2	191.570	183.444	183.678	1,04
11		Q_3	318.709	183.913	184.451	1,73
12		Q_4	130.967	184.989	186.763	0,70
13	2015	Q_1	98.709	188.538	194.676	0,51
14		Q_2	205.766	200.813	202.415	1,02
15		Q_3	367.811	204.017	206.081	1,78
16		Q_4	143.781	208.145	209.363	0,69
17	2016	Q_1	115.222	210.581	210.909	0,55
18		Q_2	215.509	211.237	212.724	1,01
19		Q_3	370.437	214.211	215.198	1,72
20		Q_4	155.675	216.186	220.387	0,71

We calculate S_t,I_t , by dividing actual number of guest (Y_t) with CMA(4) data. In fact we have got coefficients that explains how much seasonal and irregular components are above or below "smooth" line of CMA(4). We continue our analysis with quantification of seasonal component S_t. In order to do that we calculate average of each seasonal irregular components (indexes) for each quarter in order to avoid irregularity. S_t calculated values are presented on the Table 5 shown below:

Table 7. Seasonal component S_t

Quarter	S_t
1	0,53
2	1,02
3	1,74
4	0,70

We add calculated data for S_t in Table 6 for entire length of our time series data:

Table 8. Time series analysis with S_t

t	Year	Quarter	Num. of guests	MA (4)	CMA (4)	S_t,I_t	S_t
1	2012	Q_1	89.383				0,53
2		Q_2	164.691				1,02
3		Q_3	291.895	165.908	166.218	1,76	1,74
4		Q_4	117.664	166.527	168.164	0,70	0,70
05	2013	Q_1	91.858	169.802	171.197	0,54	0,53
6		Q_2	177.790	172.592	174.020	1,02	1,02
7		Q_3	303.054	175.449	175.767	1,72	1,74
8		Q_4	129.092	176.085	177.808	0,73	0,70
9	2014	Q_1	94.404	179.530	181.487	0,52	0,53
10		Q_2	191.570	183.444	183.678	1,04	1,02
11		Q_3	318.709	183.913	184.451	1,73	1,74
12		Q_4	130.967	184.989	186.763	0,70	0,70
13	2015	Q_1	98.709	188.538	194.676	0,51	0,53
14		Q_2	205.766	200.813	202.415	1,02	1,02
15		Q_3	367.811	204.017	206.081	1,78	1,74
16		Q_4	143.781	208.145	209.363	0,69	0,70
17	2016	Q_1	115.222	210.581	210.909	0,55	0,53
18		Q_2	215.509	211.237	212.724	1,01	1,02
19		Q_3	370.437	214.211	215.198	1,72	1,74
20		Q_4	155.675	216.186	220.387	0,71	0,70

Our next step is to "clean" the data from seasonal component by dividing Y_t/S_t. Calculated values are presented in last column of the Table 7:

Table 9. Desseasonalized data

T	Year	Quarter	Num. of guests	MA (4)	CMA (4)	S_t,I_t	S_t	Dessesonalized
1	2012	Q_1	89.383				0,53	168.647
2		Q_2	164.691				1,02	161.462
3		Q_3	291.895	165.908	166.218	1,76	1,74	167.756
4		Q_4	117.664	166.527	168.164	0,70	0,7	168.091
5	2013	Q_1	91.858	169.802	171.197	0,54	0,53	173.317
6		Q_2	177.790	172.592	174.020	1,02	1,02	174.304
7		Q_3	303.054	175.449	175.767	1,72	1,74	174.169
8		Q_4	129.092	176.085	177.808	0,73	0,7	184.417
9	2014	Q_1	94.404	179.530	181.487	0,52	0,53	178.121
10		Q_2	191.570	183.444	183.678	1,04	1,02	187.814
11		Q_3	318.709	183.913	184.451	1,73	1,74	183.166
12		Q_4	130.967	184.989	186.763	0,70	0,7	187.096
13	2015	Q_1	98.709	188.538	194.676	0,51	0,53	186.243
14		Q_2	205.766	200.813	202.415	1,02	1,02	201.731
15		Q_3	367.811	204.017	206.081	1,78	1,74	211.386
16		Q_4	143.781	208.145	209.363	0,69	0,7	205.401
17	2016	Q_1	115.222	210.581	210.909	0,55	0,53	217.400
18		Q_2	215.509	211.237	212.724	1,01	1,02	211.283
19		Q_3	370.437	214.211	215.198	1,72	1,74	212.895
20		Q_4	155.675	216.186	220.387	0,71	0,7	222.393

In order to make prediction of number of guest that will visit the Republic of Macedonia in 2017 we create next column named T_t – "Trend component in time *t*" (*t* - reffered to the first column). We run simple linear regression using the deseassionalized variables as Y and *t* variable as X in our regression model where Anova software is used. Summary output table 8 of our regression analysis is presented bellow:

Table 10. Regression analysis

Summary Output

Regression Statistics	
Multiple R	0,96
R Square	0,92
Adjusted R Square	0,92
Standard Error	5420,9276
Observations	20

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	1	6,321E+09	6,32E+09	215,0984978	1,87725E-11
Residual	18	528956205	29386456		
Total	19	6,85E+09			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*
Intercept	156482,52	2518,1914	62,14084	1,85213E-22	151191,9967	161773
t	3083,0564	210,21455	14,66624	1,87725E-11	2641,412021	3524,701

We use regression results to create column T_t as trend component, as shown on next table 9:

Table 11. Trend Component calculation

t	Year	Quarter	Num. of guests	MA (4)	CMA (4)	S_t,I_t	S_t	Dessesonalized	T_t
1	2012	Q_1	89.383				0,53	168.647	159.566
2		Q_2	164.691				1,02	161.462	162.649
3		Q_3	291.895	165.908	166.218	1,76	1,74	167.756	165.732
4		Q_4	117.664	166.527	168.164	0,70	0,70	168.091	168.815
5	2013	Q_1	91.858	169.802	171.197	0,54	0,53	173.317	171.898
6		Q_2	177.790	172.592	174.020	1,02	1,02	174.304	174.981
7		Q_3	303.054	175.449	175.767	1,72	1,74	174.169	178.064
8		Q_4	129.092	176.085	177.808	0,73	0,70	184.417	181.147
9	2014	Q_1	94.404	179.530	181.487	0,52	0,53	178.121	184.230
10		Q_2	191.570	183.444	183.678	1,04	1,02	187.814	187.313
11		Q_3	318.709	183.913	184.451	1,73	1,74	183.166	190.396
12		Q_4	130.967	184.989	186.763	0,70	0,70	187.096	193.479
13	2015	Q_1	98.709	188.538	194.676	0,51	0,53	186.243	196.562
14		Q_2	205.766	200.813	202.415	1,02	1,02	201.731	199.645
15		Q_3	367.811	204.017	206.081	1,78	1,74	211.386	202.728
16		Q_4	143.781	208.145	209.363	0,69	0,70	205.401	205.811
17	2016	Q_1	115.222	210.581	210.909	0,55	0,53	217.400	208.894
18		Q_2	215.509	211.237	212.724	1,01	1,02	211.283	211.978
19		Q_3	370.437	214.211	215.198	1,72	1,74	212.895	215.061
20		Q_4	155.675	216.186	220.387	0,71	0,70	222.393	218.144

In order to make prediction we have to combine all components that we have calculated separately S_t and T_t. Results are in presented in the column "prediction", calculated as multiplied seasonal and trend components. We finally proceed with forecasting by adding new periods (four quarters for 2017 in our next Table 10 and calculate forecasted values:

Table 12. Forecasting

t	Year	Quarter	Num. of guests	MA (4)	CMA (4)	S_t,I_t	S_t	Desseso-nalized	T_t	Forecast
1	2012	Q_1	89.383				0,53	168.647	159.566	84.570
2		Q_2	164.691				1,02	161.462	162.649	165.902
3		Q_3	291.895	165.908	166.218	1,76	1,74	167.756	165.732	288.373
4		Q_4	117.664	166.527	168.164	0,70	0,70	168.091	168.815	118.170
5	2013	Q_1	91.858	169.802	171.197	0,54	0,53	173.317	171.898	91.106
6		Q_2	177.790	172.592	174.020	1,02	1,02	174.304	174.981	178.480
7		Q_3	303.054	175.449	175.767	1,72	1,74	174.169	178.064	309.831
8		Q_4	129.092	176.085	177.808	0,73	0,70	184.417	181.147	126.803
9	2014	Q_1	94.404	179.530	181.487	0,52	0,53	178.121	184.230	97.642
10		Q_2	191.570	183.444	183.678	1,04	1,02	187.814	187.313	191.059
11		Q_3	318.709	183.913	184.451	1,73	1,74	183.166	190.396	331.289
12		Q_4	130.967	184.989	186.763	0,70	0,70	187.096	193.479	135.435
13	2015	Q_1	98.709	188.538	194.676	0,51	0,53	186.243	196.562	104.178
14		Q_2	205.766	200.813	202.415	1,02	1,02	201.731	199.645	203.638
15		Q_3	367.811	204.017	206.081	1,78	1,74	211.386	202.728	352.747
16		Q_4	143.781	208.145	209.363	0,69	0,70	205.401	205.811	144.068
17	2016	Q_1	115.222	210.581	210.909	0,55	0,53	217.400	208.894	110.714
18		Q_2	215.509	211.237	212.724	1,01	1,02	211.283	211.978	216.217
19		Q_3	370.437	214.211	215.198	1,72	1,74	212.895	215.061	374.205
20		Q_4	155.675	216.186	220.387	0,71	0,70	222.393	218.144	152.701
21	2017	Q_1	123.122	224.589	200.281	0,61	0,53		221.227	117.250
22		Q_2	249.120	175.972			1,02		224.310	228.796
23		Q_3					1,74		227.393	395.664
24		Q_4					0,70		230.476	161.333

After finished calculations now we can plot forecasting series on the chart in order to visualize accuracy of the forecasting, as follows:

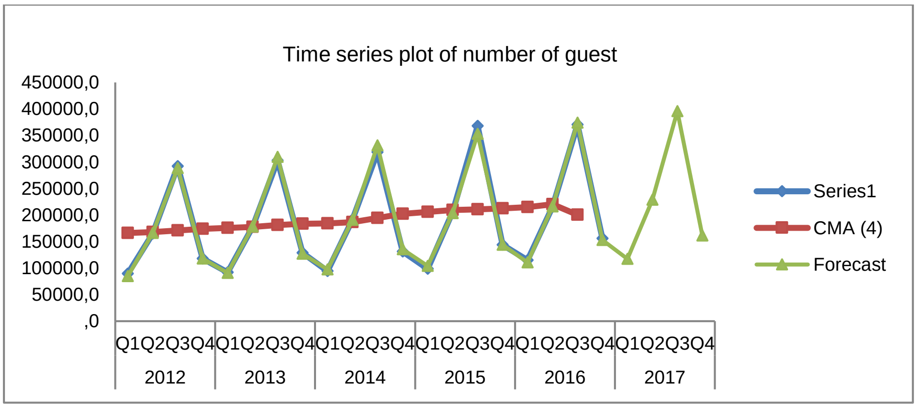

Figure 3. Time series of number of guest - forecasts

We can conclude from the chart that our forecasting line (green) – general flows is really well forecasted. This plot with 3 lines shows that actual data and predicted data for year 2017 are well approximated. Our forecasting values are appropriate with movement and behaviour of actual values. Finally we check our forecasting accuracy with actual data for 2017 and conclude that forecasting values are very close as the actual data for the first two quarters of 2017.

Table 13. Actual and forecasted data for number of tourist in the Republic of Macedonia 2017

Year	Quarter	Actual	Forecasted	Difference %
2017	Q_1	123.122	117.250	-4,8%
	Q_2	249.120	228.796	-8,1%

Source: State Statistical Office

Comparing actual and forecasted data we can conclude that there is relatively small difference between them and that time series forecasting moving average model provides accurate and reliable prediction of number of tourists with acceptable estimation. This finding confirms our research hypothesis that moving average model for the time series data provide accurate forecasting for the number of tourist guests in next year.

CONCLUSION

Time series modeling and forecasting has fundamental importance to various practical domains having in mind that time series forecasting enables predicting the future by understanding the past. In tourist industry forecasting future revenues depends on capability to make prediction of number of tourists, having in mind that there are data for average tourist spending per day. If all stakeholders in tourist industry have relevant and accurate forecast of number of tourists that will visit the country they can provide reliable input for their financial planning. In order to do that it is important to use quantitative methods and statistics tools that can help to deal with past data where seasonality, irregularity and sometimes random variables appears.

There are many comprehensive theoretical studies in literature that suggest the necessity to use quantitative models and techniques for various types of data. In this paper we focus on a time series analysis. A time series is a collection of observations made sequentially in time. Time series analysis refers to problems in which observations are collected at regular time intervals and there are correlations among successive observations.

Analysis of the data for the number of guest shows strong seasonal component, obvious pattern in data variations and trend. Having in mind that the moving average merely smooths the fluctuations in the data we chose this time series forecasting model for our research. A moving average model works well when the data have a fairly linear trend and a definite rhythmic pattern of fluctuations.

Research finding confirms our research hypothesis that moving average model for the time series data provide accurate forecasting for the number of tourist guests in next year.

This study outlines directions for future researches that could be investigated to improve the forecasting of other tourist and economic indicators for the Macedonian economy. Due to the fact that we use limited data and time series of number of guest (2012–2016) and compare with actual number of tourist for first two quarters of 2017, longer time series would allow estimation with greater precision.

REFERENCES

Armstrong, Scott, J. 2006. Findings from evidence-based forecasting: Methods for reducing forecast error. *International Journal of Forecasting* 22: 583–589.

Adikhari, Ratnadip, and R. K. Agrawal. 2013. *An Introductory Study on Time Series Modeling and Forecasting.* Berlin: LAP lambert Academic Publishing, Germany.

Bartholomew, David J. 1971. Time series analysis forecasting and control. *Journal of the Operational Research Society* 22 (2): 199–201

Brockwell, Peter J., and Richard A. Davis. 2013. *Time Series: Theory and Mathods.* New York: Springer Science & Business Media.

Bollerslev, Tim. 1986. Generalized autoregressive conditional heteroscedasticity. *Journal of Econometrics* 31 (3): 307–327.

Box, George E. P., and Gwilym M. Jenkins. 1971. Time Series Analysis: Forecasting and Control. *Operational Research Quarterly* 22 (2): 199–201.

Chatfield, Chris. 1996. Model uncertainty and forecast accuracy. *Journal of Forecasting* 15 (7): 495–508.

Chernick, Michael R., 1994. Time Series: Forecasting, Simulation, Applications by Gareth, Janacek and Louise Swift, Reviewed Work. *The American Statistician* 48 (1): 58–59.

Cochrane, John H. 1997. *Time Series for Macroeconomic and Finance.* Chicago: Graduate School of Business. University of Chicago.

Cottrell, Marie, Bernanrd Girard, Yvone Girad, Morgan Mangeas, and Corinne Muller. 1995. Neural modeling for time series: A statistical stepwise method for weight elimination. *IEEE Trans. Neural Networks* 6 (6): 1355–1364.

Diggle, Peter J. 1990. *Time Series: A Biostatistical Introduction.* London: Oxford University Press.

Fildes, Robert, and Spyros Makridakis. 1993. *The Impact of Empirical Accuracy Studies on Time Series Analysis and Forecasting*. France, Fontainebleau: INSEAD.

Gardner, Everette S. Jr., and Ed Mckenzie. 1985. Forecasting Trends in Time Series. *Management Science* 31 (10): 1237–1246.

Janacek, Gareth. 2009. Time series analysis forecasting and control. *Journal of Time Series Analysis* 31 (4): 303.

Montgomery, Douglas C., Cheryl L. Jennings, and Murat Kulachi. 2008. *Introduction to Time Series Analysis and Forecasting.* London: Wiley.

Nelson, Charles R., and Charles R. Plosser. 1982. Trends and random walks in macroeconomic time series: Some evidence and implications. *Journal of Monetary Economics* 10 (2): 139–162.