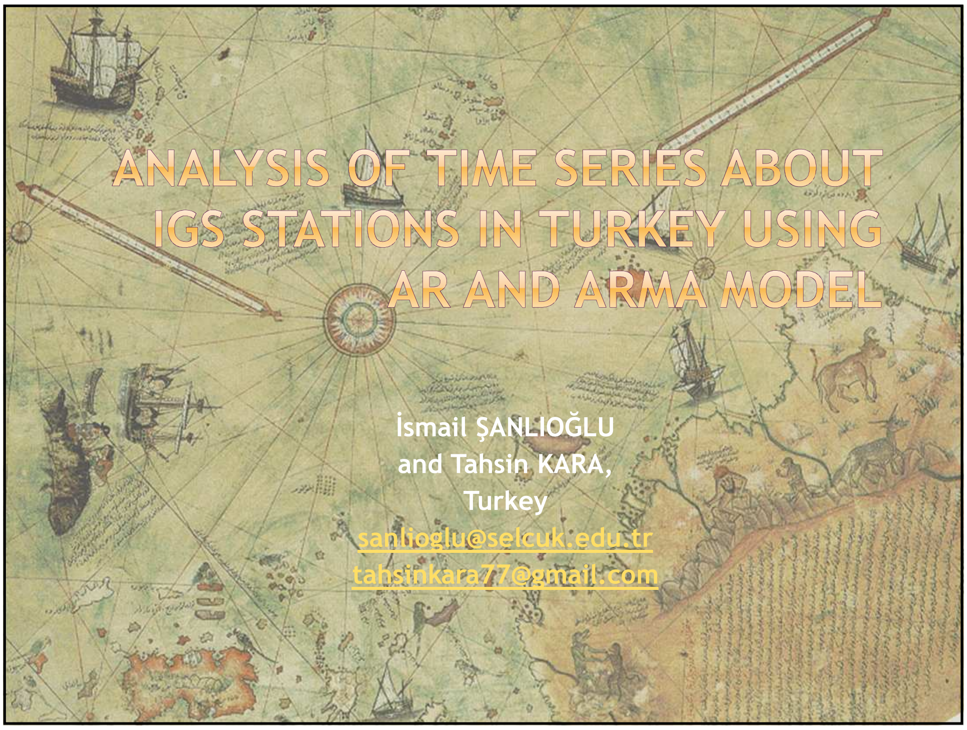


ANALYSIS OF TIME SERIES ABOUT IGS STATIONS IN TURKEY USING AR AND ARMA MODEL

İsmail ŞANLIOĞLU
and Tahsin KARA,
Turkey
sanlioglu@selcuk.edu.tr
tahsinkara77@gmail.com

CONTENTS OF PRESENTATION

INTRODUCTION

- Time series analysis is of great importance in the field of engineering. We consider that a very high cost and labor-intensive seen in their realization of the plan-project phases and steps of structures such as dams, bridges, towers and to ensure the continuation of engineering structures. These structures show a different behavior in life expectancy under different loads, such as deformation and displacement. The necessary measures will be provided in time with continuous monitoring of behavior and with this pre-determination of the possible accidents that may occur.

TIME SERIES

- Generally GNSS permanent station time series show various types of signals, some of which are real whilst the others may not have apparent causes: miss-modeled errors, effects of observational environments, random noise or any other effects produced by GNSS analysis software or operator choices of software parameters and settings of a prior stochastic models for different types of measurements

TIME SERIES

- In this study, data of coordinate component of permanent IGS stations in Turkey had been separately analyzed in terms of time series by using autoregressive (AR) and autoregressive moving average (ARMA) models which are the linear time series methods.

TIME SERIES

- **In a classical model, time serial has four component (Mann 1995).**

- Trend (T_t) - long term movements in the mean
- Cyclical (C_t) – cyclical fluctuations related to the calendar
- Seasonal (S_t) – other cyclical fluctuations (such as business cycles)
- Irregular (E_t) – other random or systematic fluctuations

These components may be combined in different ways.
It is usually assumed that they are multiplied or added

$$X(t) = T(t) \times C(t) \times S(t) \times E(t)$$

$$X(t) = T(t) + C(t) + S(t) + E(t)$$

TIME SERIES

- To correct for the trend in the first case one divides the first expression by the trend (T). In the second case it is subtracted. If we can estimate and extract the deterministic *T(t)* and *C(t)&S(t)*, we can investigate the residual component *E(t)*. After estimating a satisfactory probabilistic model for the process {*E(t)*}, we can predict the time series {*X(t)*} along with *T(t)* and *C(t)&S(t)*. Therefore, the GPS time series analysis actually refers to model the residual component *E(t)*

TIME SERIES MODELS AND ANALYSIS

- A time series is of a quantity of interest over time in an ordered set. The purpose of this analysis with time-series is face of the reality represented by a set of observation and over time in the future values of the variables to predict accurately
- Models for time series data can have many forms and represent different stochastic processes. When modeling variations in the level of a process, three broad classes of practical importance are the *autoregressive* (AR) models, the integrated (I) models, and the moving average (MA) models. These three classes depend linearly on previous data points. Combinations of these ideas produce autoregressive moving average (ARMA) and autoregressive integrated moving average (ARIMA) models

TIME SERIES MODELS AND ANALYSIS

- The different much examples are encountered when one examine types of time series such as;

  - Autocorrelation function
  - Partial autocorrelation function
  - The moving average (Moving Average, MA) series
  - Autoregressive moving average (ARMA) series

TIME SERIES MODELS AND ANALYSIS

- Autocorrelation function of time series is calculated with this equation

$$r_k = \frac{\sum_{i=1}^{N}(Z_i - \bar{Z})(Z_{i+k} - \bar{Z})}{\sum_{i=1}^{N}(Z_i - \bar{Z})^2}$$

- N is the number of observations,
- $\bar{Z}$ the number of observations in the middle of the corresponding data,
- k-delay
- r_k the autocorrelation coefficient of Z_i stochastic component

TIME SERIES MODELS AND ANALYSIS

- Partial Autocorrelation function of time series is calculated with this equation

$$\phi_{mm} = \frac{r_m - \sum_{j=1}^{m-1} \phi_{m-1,j} r_{m-1}}{1 - \sum_{j=1}^{m-1} \phi_{m-1,j} r_j}$$

- After adjusting for all other lagged observations, partial correlation examines the relationship between the X_t variable with the variable X_{t+k} obtained from X_t variable by any k-delay. Determining the coefficients of this relationship is called partial autocorrelation coefficient. ϕ_{mm} is indicated by the symbol

TIME SERIES MODELS AND ANALYSIS

- Partial Autocorrelation function:
  - In time series analysis, the partial autocorrelation function (PACF) plays an important role in data analyses aimed at identifying the extent of the lag in an autoregressive model. The use of this function was introduced as part of the Box-Jenkins approach to time series modeling, where by plotting the partial autocorrelative functions one could determine the appropriate lags **p** in an AR (**p**) model or in an extended ARIMA (**p**,**d**,**q**) model.

TIME SERIES MODELS AND ANALYSIS

- Partial autocorrelation plots are a commonly used tool for identifying the order of an autoregressive model. The partial autocorrelation of an AR(*p*) process is zero at lag *p* + 1 and greater. If the sample autocorrelation plot indicates that an AR model may be appropriate, then the sample partial autocorrelation plot is examined to help identify the order. One looks for the point on the plot where the partial autocorrelations for all higher lags are essentially zero. Placing on the plot an indication of the sampling uncertainty of the sample PACF is helpful for this purpose: this is usually constructed on the basis that the true value of the PACF, at any given positive lag, is zero

TIME SERIES MODELS AND ANALYSIS

- Autoregressive (AR) model, Moving-average model (MA), Autoregressive-moving-average model (ARMA)
- It is a special case of the more general ARMA model of time series.

$$Z_t = \mu + \phi_1 Z_{t-1} + \phi_2 Z_{t-2} + \dots + \phi_p Z_{t-p} + Z_t$$

- Where , former observation values, parameters of model; coefficients for observations, μ mean value is a constant , white noise; term of error. Some constraints are necessary on the values of the parameters of this model in order that the model remains stationary. For example, processes in the AR(1) model with ≥ 1 are not stationary.

TIME SERIES MODELS AND ANALYSIS

- A moving-average model (MA) is conceptually a linear regression of the current value of the series against current and previous (unobserved) white noise error terms or random shocks. The random shocks at each point are assumed to be mutually independent and to come from the same distribution, typically a normal distribution, with location at zero and constant scale. Fitting the MA estimates is more complicated than with autoregressive models (AR models) because the lagged error terms are not observable. This means that iterative non-linear fitting procedures need to be used in place of linear least squares

TIME SERIES MODELS AND ANALYSIS

- The notation MA(*q*) refers to the moving average model of order *q*:

$$Z_t = \mu + a_1 - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \ldots - \theta_q a_{t-q}$$

- Where a_1, a_{t-1},… , a_{t-q}, terms of error, $\theta_1, \theta_2 \ldots \theta_q$ coefficients related to errors, μ is the expectation of Z_t (often assumed to equal 0).The value of *q* is called the order of the MA model.

TIME SERIES MODELS AND ANALYSIS

- The model is usually then referred to as the ARMA (*p*, *q*) model where *p* is the order of the autoregressive part and *q* is the order of the moving average part
-

$$Y_t = \delta + \phi_1 Y_{t-1} + \ldots + \phi_p Y_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \ldots. + \theta_q \varepsilon_{t-q}$$

- Where θ parameters of moving average part and ϕ parameters of autoregressive part and ε random variables of model, the term of cut δ and the error terms εt .

TIME SERIES MODELS AND ANALYSIS

- **Akaike Information Criterion**
  - The Akaike Information Criterion (AIC) is a measure of the relative quality of a statistical model, for a given set of data. As such, AIC provides a means for model selection. AIC deals with the trade-off between the goodness of fit of the model and the complexity of the model.

TIME SERIES MODELS AND ANALYSIS

- AIC is expressed by the following equation

$$AIC = N \cdot \ln\left(\sigma_e^2\right) + 2 \cdot (p+q)$$

- where N is the sample size used to estimate the model,

$$\sigma_e^2 = \frac{\text{Residual Sum of Squares}}{n}$$

- and , σ_e^2 variance, p degree of AR(p) model, q degree of MA(q) model. After selecting the model which gives the minimum value in AIC, statistic has been also calculated for this model. If this value is smaller than 95% confidence level, the model is considered to be suitable.

TIME SERIES MODELS AND ANALYSIS

- **V - Statistics**
- For the implementation of this test;

$$S = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2}$$

- S value must be calculated.

$$V = \frac{\left|X_E - \overline{X}\right|}{S}$$

TIME SERIES MODELS AND ANALYSIS

- **V - Statistics ;**
- If the hypothesis is smaller than V table values, the hypothesis will be considered valid according to the V table with n measurements and α significance level. The purpose of this test, one can understand whether the series are stationary or not when one investigate the long periodic time series. As a result, values not outlying must stay in the own time series.

NUMERICAL APPLICATION

- Time series raw data of IGNSS stations in Turkey
- ✓ ANKR (20805M002),
- ✓ TUBI (20806M001),
- ✓ ISTA (20807M001),
- ✓ TRAB (20808M001),

NUMERICAL APPLICATION

GNSS station used in time series analysis

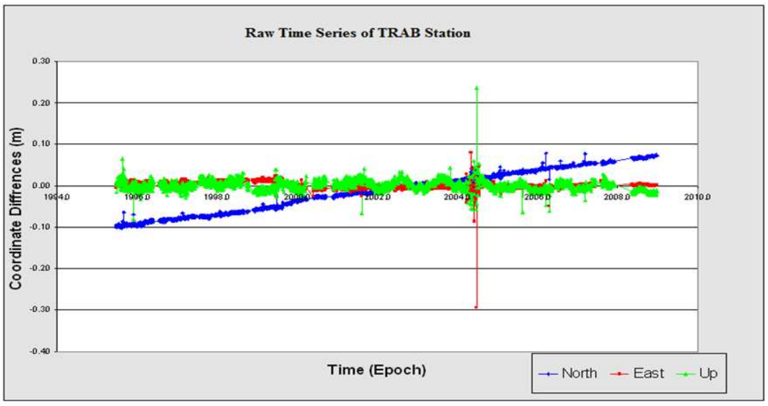
NUMERICAL APPLICATION

- The raw data of N (North), E (East) ve U (Up) local coordinates components of these stations from the Scripps Orbit and Permanent Array Center (SOPAC) GPS archive (web-1, 2009).

NUMERICAL APPLICATION

- Dates of the data cover
  - the period from 26/06/1995 to 12/21/2008 in ANKR station,
  - the period from 21/12/1995 to 08/05/1998 in TUBI station,
  - the period from 26/12/1999 to 12/21/2008 in ISTA station,
  - the period from 26/12/1999 to 28/11/2007 in TRAB station.

NUMERICAL APPLICATION

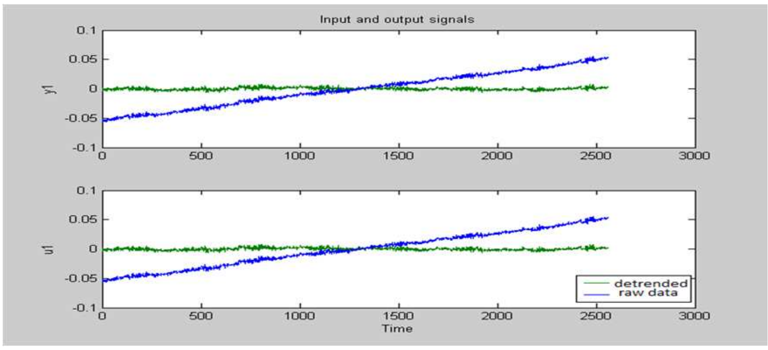


NUMERICAL APPLICATION

- Firstly, the data interruptions of stations and gaps in time series were examined. For these gaps and interruptions, time series were analyzed. The long-term discontinuities in the stations were analyzed separately for different ranges of epochs. N, E and U components of stations' coordinates divided periods of 6 months. V - Statistics were performed to determine outliers separately for every period of application because of big range of data. Thus the outliers were eliminated from time series.

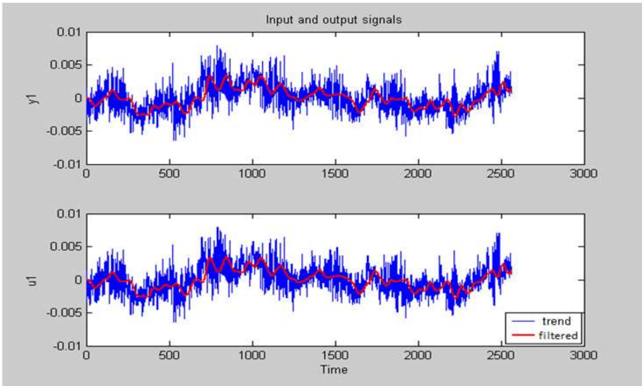
NUMERICAL APPLICATION

- Series models and analysis has been done by using MATLAB System Identification Toolbox module. By using Microsoft Excel the linear trend equation is determined and then the trend has been removed from data of all station time series. In figure 3 there is a detrended of series (raw data is blue color in series, detrended data is green color).



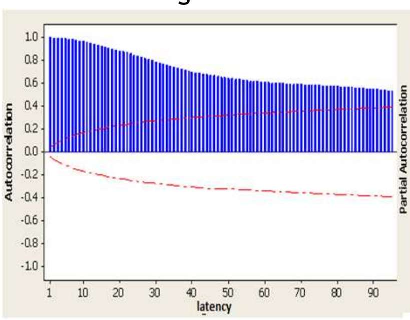
NUMERICAL APPLICATION

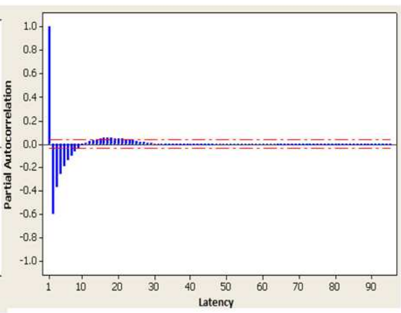
- The filtering of the series has been done in periodogram graph in MATLAB software



NUMERICAL APPLICATION

- After removed trend and noise-free, it is necessary to determine the type and *parameters of models. So autocorrelation coefficient and partial autocorrelation coefficient values have been computed by using MINITAB software. The red dashed line indicates the 95% confidence interval in the following figures.* Autocorrelation of time series is decreasing exponentially in these figures

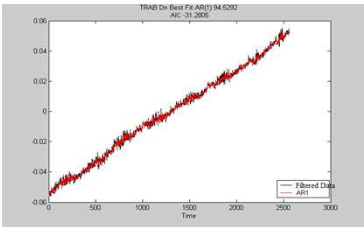




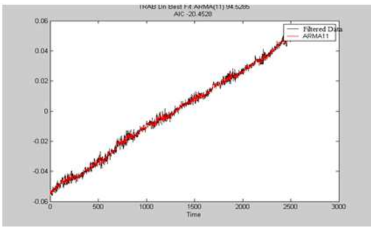
NUMERICAL APPLICATION

- Also the partial autocorrelation decreases in terms of absolute value. For this reason, we can say that the time series model is AR (p) model. Because partial autocorrelation value decreases after the first degree, we can say that the value of the degree of the model is one. These results are not enough. So, AIC for selecting the model has been also used. The harmonization of models has been given in percent and AIC has been seen on upper part of graph in the following graphics

NUMERICAL APPLICATION

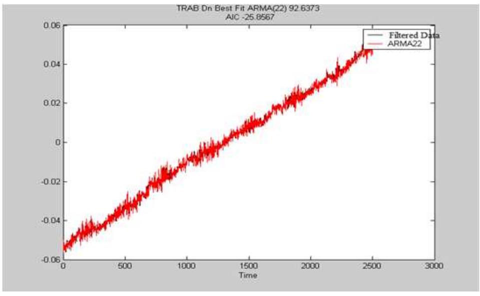


- AR(1) graph of TRAB station using north component of filtered time series



- ARMA (1 1) graph of TRAB station using north component of filtered time series.

NUMERICAL APPLICATION



- ARMA (2 2) graph of TRAB station using north component of filtered time series.

NUMERICAL APPLICATION

The time series models of ANKR station.

North Component			East Component			Up Component		
Between 26.06.1995 - 13.08.1999			Between 26.06.1995 - 13.08.1999			Between 26.06.1995 - 13.08.1999		
AR(1)	ARMA(1 1)	ARMA(2 2)	AR(1)	ARMA(1 1)	ARMA(2 2)	AR(1)	ARMA(1 1)	ARMA(2 2)
87.0597	86.955	82.4975	40.6973	40.7083	21.5074	31.7211	31.7116	10.4934
AIC	AIC	AIC	AIC	AIC	AIC	AIC	AIC	AIC
-26.31	-18.5321	-22.2725	-27.0197	-18.1905	-22.5565	-22.6288	-15.5597	-19.1472
Between 25.11.2000 - 17.03.2004			Between 25.11.2000 - 17.03.2004			Between 25.11.2000 - 17.03.2004		
AR(1)	ARMA(1 1)	ARMA(2 2)	AR(1)	ARMA(1 1)	ARMA(2 2)	AR(1)	ARMA(1 1)	ARMA(2 2)
87.463	87.4666	83.3067	3.9882	4.0272		45.3794	45.3744	45.3794
AIC	AIC	AIC	AIC	AIC	AIC	AIC	AIC	AIC
-27.7669	-19.6468	-23.7396	-28.6403	-19.5317		-26.5576	-16.5874	-21.8601
Between 20.07.2004 - 30.11.2007			Between 20.07.2004 - 30.11.2007			Between 20.07.2004 - 30.11.2007		
AR(1)	ARMA(1 1)	ARMA(2 2)	AR(1)	ARMA(1 1)	ARMA(2 2)	AR(1)	ARMA(1 1)	ARMA(2 2)
88.4305	88.5839	84.5457	19.0546	19.1023		40.2548	40.2818	21.2152
AIC	AIC	AIC	AIC	AIC	AIC	AIC	AIC	AIC
-26.6997	-18.9479	-22.7592	-26.6496	-18.9048		-25.4753	-16.2965	-21.0813

NUMERICAL APPLICATION

The time series models of ISTA, TUBI, TRAB stations

ISTA station			TUBI station			TRAB station		
North between 21.08.1999 - 13.12.2008			North between 26.12.1999 - 06.11.2006			North between 26.12.1999 - 06.11.2006		
AR(1)	ARMA(1 1)	ARMA(2 2)	AR(1)	ARMA(1 1)	ARMA(2 2)	AR(1)	ARMA(1 1)	ARMA(2 2)
94.2231	94.2165	92.2221	92.4869	92.4791	89.9978	94.5292	94.5285	92.6373
AIC	AIC	AIC	AIC	AIC	AIC	AIC	AIC	AIC
-28.8191	-18.8281	-23.8019	-30.4257	-20.0129	-25.3495	-31.2805	-20.4528	-25.8567
East between 21.08.1999 - 13.12.2008			East between 26.12.1999 - 06.11.2006			East between 26.12.1999 - 06.11.2006		
AR(1)	ARMA(1 1)	ARMA(2 2)	AR(1)	ARMA(1 1)	ARMA(2 2)	AR(1)	ARMA(1 1)	ARMA(2 2)
95.7083	95.6755	94.3099	95.0616	95.0644	93.4528	94.8976	94.8914	93.2043
AIC	AIC	AIC	AIC	AIC	AIC	AIC	AIC	AIC
-27.7043	-17.9143	-22.7883	-28.2692	-19.0887	-23.5571	-32.274	-20.2712	-26.2516
Up between 21.08.1999 - 13.12.2008			Up between 26.12.1999 - 06.11.2006			Up between 26.12.1999 - 06.11.2006		
AR(1)	ARMA(1 1)	ARMA(2 2)	AR(1)	ARMA(1 1)	ARMA(2 2)	AR(1)	ARMA(1 1)	ARMA(2 2)
38.8958	38.9003	17.889	48.2008	48.19	31.3275	38.1557	38.1396	17.232
AIC	AIC	AIC	AIC	AIC	AIC	AIC	AIC	AIC
-30.5854	-17.7834	-24.674	-29.6703	-17.1804	-24.0413	-32.6195	-17.7205	-25.933

NUMERICAL APPLICATION

These processes have been applied to north, east and up coordinate component of ANKR, ISTA, TRAB and TUBI station filtered time series. According to results of model graphs all models have been determined as AR (1) model. Also the best fit model AR (1) model determined in AIC analysis. Additionally, some ARMA (p, q) series that conform to the model have been tested. The harmonization of models has been given in percent and AIC has been seen at the table 1 and table 2. At the table 1, time series of ANKR station have been divided into three parts because of Gölcük/İzmit earhtquake and receiver&antenna changing, and the other stations' series are full.

CONCLUSIONS

- As shown in the graph of time series, there are data discontinuities in time series for the hardware reasons, i.e. ANKR station. Thus we can say that these interruptions could adversely affect the results of time series analysis and analysis will result in incorrect results. So for the modeling of time series, the available data must be compatible with each other as possible and loss of data by the long-and short-term in series should not be. Also, data have to be obtained from a long-term series. If these rules are performed, better results can be achieved.

CONCLUSIONS

- AR and ARMA models has been used time series analysis of the stations and autocorrelation and partial autocorrelation coefficients of the AR (1) and ARMA (1 1) models the values are close to each other out. Degrees of northing and easting components are one degree in auto-regression and partial- auto-regression graphics and their autocorrelations are in downward tendency in the positive direction; and decreasing suddenly in the positive direction after the first degree, their partial-autocorrelations make small changes in both positive and negative directions.

CONCLUSIONS

- It can be said that these tendencies are consistent with the auto-regression model (AR), among the models of time series. Akaike Information Criterion has tested to model results and as a result the station data that they have made in the past period the movements, in the future to allow for the best interpretation of the model AR (1) that has been seen.

- Thank you for listening