Model of Transfer Function In Series Time Analysis to Predict Rainfall

Ahmad Zaki

Department of Mathematics, Faculty of Mathematics and Science
State University of Makassar
Makassar, Indonesia

ABSTRACT: This study aimed to model and predict rainfall in the city of Makassar using transfer function model in time series analysis. Forecasting procedures performed by several stages, namely: identification, pre-whitening, the cross-correlation calculation, weighting impulse response, determining the value (b, r, s) and row noise, parameter estimation, diagnostic testing and forecasting. Data research using secondary data is data of rainfall and air temperature over 100 months Makassar city modeled and estimated using the basic theory of time series analysis and the theory of transfer function model. The results obtained by the data model of rainfall and air temperature using time series analysis model transfer function (b, r, s) as follows:

$$y_t = \frac{(\omega_0)}{(1-\delta_1 B-\delta_2 B^2)} x_t + (1-\Theta_1 B^{12}) a_t$$

forecasting results rainfall data showed that the rainy season ended in the fifth month of the year and the dry season only started in the sixth month of the year, while the next rainy season, or the end of the dry season will occur in the eleventh month of the year.

KEYWORDS: Time Series Analyis, Transfer Function Model, Rainfall.

1. INTRODUCTION

Transferfunction model is a model describing the future prediction value of a time series (***output series***or**Yt**) based on the past values from one or more inter-connected time series (***input series***or**Xt**) including their outputs.

For example, the model between sales total (Yt) and advertising expenditure (Xt) which is observed per month. (Makridakis, Wheelwright, and McGee, 1983), model between sales (Yt) and leading indicator (Xt) analyzed by Box and Jenkins (1976), model between water discharge in a barrage (Yt) and waterfall (Xt) which is observed in the same time interval (Gusman, 2000).

The general model of transfer function for single-input (***x***t) and single-output (***y***t) is as follows

$$y_t = \mu + v(B) x_t + \eta_t$$

Transfer function has a concept with more than one periodic series. To analyze the series from the data, it needs a series, i.e. input series, output series, and the obstacle. It is a main condition to use transfer function. Input series named as xt, output series named as ytand the obstacle. Next, an analysis is applied to a process which is called the formation stage of transfer function which implies the conclusion of a model which is called transfer function model. The analysis aims at apply the metho of transfer function in a periodic series analysis to obtain prediction model from an observation and to see the influence exsistence of both input and output series.

Prediction is an important element in decision making. Several unseen factors have influences to the effectivity of decision making. The prediction plays an important role in several fields, e.g. meteorlogy related to weather forecast.

Forecasting weather is the same as predicting what weather which will happen. Badan Meteorologi Klimatologi dan Geofisika (BMKG) applies two kinds of predicting, i.e. short-term prediction and long-term prediction. The short-term is a prediction which is done once a day to predict the weather in the next day. Meanwhile, the long-term prediction is a kind of prediction to predict when the rain season and dry season begin using criteria of the rain season begins when the rainfall is more than 50 mm for three consecutive decades and the dry season begins when the rainfall is less than 50 mm for three consecutive decades (Raodhatullah, 2006).

South Sulawesi province has variety of natural resources which can be developed, i.e. mining sector, tourism, industry, plantage, naval, fishery, and agriculture. Meanwhile, Makassar as the capital city of the province has variety of natural resources which can be developed, i.e. tourism, agriculture, and aviation sector. The development of those sectors is influenced by the climate and the weather in the city, e.g the waterfall. The data of the rainfall can be taken into account as the basis of the decision maker, planner, and the practitioner. By considering the rainfall condition, the prediction of the rainfall is definitely needed.

Time series analysis in the model of ARIMA is a single time (univariat). ARIMA (*Autoregressive Integrated Moving Avarage*) is a model has been developed and applied for a prediction. This method is developed by Box-Jenkins, frequently called Box-Jenkins method, referring to two statisticians who successfully develop and solidify the method. The method is the combination of smoothig method, regression method, and decomposition method. The method is used to predict the price of stock and other kinds of time series variables. This model relates to independent variable which is used to explain dependent variable. However, it needs high number of data and long time series. Whereas transfer function modelling is a multiple time series for multivariate model. Multivariate model is divided into two forms, i.e. bivariate model consisting of data which have two periodic series, meanwhile multivariate model consists of more than two time series. Since the transfer function model combines several characteristics of regression analysis and those of the time series (ARIMA), transfer function method is also called a method combining regression and time series approach. Regression method is a method using cause-effect approach and aims at predicting the future situation and measure several independent variables including their influences toward dependent variables.

There are some models in multivariate time series analysis of which one of them is called transfer function model, i.e. data consisting two or more time series. The concept of transfer function consists of input series denoted by X_t and output series denoted by Y_t, and all other variables called constraint denoted by $N_t$through several future periods. In understanding the method, it should be known that output time series called $Y_t$which is predicted being influenced by input periodic series and constraint N_t. Those systems are dynamic, in other words input series $X_t$gives influence to the *output*series $Y_t$through a method called transfer function which distributes the effect $X_t$through several future periods.

Bivariate time series analysis in the model of transfer function aims at using transfer function model to predict the value of $Y_t$in the future. The author would identify the form of transfer function model, parameter estimations and diagnostic test aiming at determining the future period value (Martika, 2007).

In the previous research from Raodhatullah with the title of the modelling of rainfall data in Makassar using ARIMA Box-Jenkins method, since it used one time series then it applied univariate model ARIMA Box-Jenkins time series analysis In addition, Rahmawati's research with the title of time series analysis using transfer functionmodel described the procedure of model transfer function formation. Moreover, M. Fathurahma's research with the title of while M. Fathurahma with title pemodelan multi input transfer function modelling used five variables based on the result instead of only three variables which have significant effect on the output variables. Furthermore, Suharnita's research on 2011 with the title of transfer function model to predict the rainfall in Makassar suggests the transfer function model and the result of the prediction.

BadanMeteorologiKlimatologidanGeofisika (BMKG) in forecasting weather has applied variety of statistical methods, i.e. regression analysis and time series analysis. Since the rainfall data are based on the time, then the author attempted to use time series analysis in rainfall modelling. One of the methods combining time series analysis and regression analysis (*causal*) is transfer function method. It has been used mostly on non-seasonal multivariate data. Therefore, this study would conduct literature and application review of the model for seasonal bivariate time series data. The literature review was focused on the statistics-mathematics theories appropriate to the configuration of transfer functionmodel in seasonal time series data. As a case study, ,the result of the review is applied in the modelling and the forecasting of rainfall data based on air temperature data.

2. LITERATURE REVIEW

2.1 TIME SERIES ANALYSIS

Time series analysisis an analysis of the observation, sensus, and composition of an event recorded over time. Commonly, the observation and sensus can be held in period of time, e.g. daily, weekly, monthly, three-monthly, semesterly, and yearly.

Several conditions that should be take into account regarding to the time series data are that measurement series is from the same source, an observation of a time and that of the other time are dependent each other statistically or correlatively, and some observations have certain pattern (Makridakis, wheelwright, & McGee, 1999).

The analysis is an analysis of data which are collected, observed, and written in a regular time interval. It is mostly used in economy as a forecasting tool based on past data.

Several concepts of time series analysis are described as follows:

Stochastic andStasioner

Compared to the future event which can be predicted based on the past experience, time series is deterministic and doesn't need further inquiry. Meanwhile, if the past experience just describe probabilistic structure of the future event, then the time series is stochastic.

The characteristics of model formation of time series analysis is assuming that the data is stationer. It is said stationer, if there is no tendency change in the term of mean and variance. In the other words, there is no significant value increase and decrease of data in stationary time series

Stationary conditions consists of two things, i.e. stationer in average and stationer in variance. The stationary condition test can apply time series diagram, i.e. scatter plot between the value of variable $Z_t$and that of t. If the diagram is fluctuative nearby the line parallel to time axis (t), then it can be said that stationary series is in average. If the stationary condition in average is not satisfied, then it needs *differencing*.

Average, Autocovariance, and autocorrelations

Stationer model {Z_t} has a constant average $E(Z_t) = \mu$, constant variance $Var(Z_t) = E(Z_t-\mu)^2 = \sigma^2$ and constant covariance $\mathrm{Cov}(Zt,Zs) = \gamma_{t,s}$ which is the function of time difference $|t-s|$.

Covariance between Z_t and Z_{t+k} is $\gamma_k = \mathrm{cov}(Z_t, Z_{t+k}) = E(Z_t-\mu)(Z_{t+k}-\mu)$. Correlation between Z_t and Z_{t+k} is $\rho_k = \frac{\mathrm{cov}(Z_t,Z_{t+k})}{\sqrt{\mathrm{var}(Z_t)}\sqrt{\mathrm{var}(Z_{t+k})}} = \frac{\gamma_k}{\gamma_0}$, where $Var(Z_t) = Var(Z_{t+k}) = \gamma_0$ meanwhile γ_k is autocovariance function and ρ_k is autocorrelation function.

Autocorrelation function and Partial autocorrelation

A*utocorrelation function* (ACF) is a *function* which describes the correlation (linear relationship) between the observation in the time *t* (notated by Z_t) with the observation in the previous time (notated by $Z_{t-1}, Z_{t-2},\ldots\ldots Z_{t-k}$) for time series $Z_1, Z_2, \ldots.,Z_n$. ACF diagram can be used as a tool to identify the stationery of data. If it tends to slowly decrease and linearly decrease, it can be concluded that the data is not stationer in the term of average.

Partialautocorrelation function (PACF) is used to measure the association rate between Z_t and Z_{t-k}, is the effect of *timelag* 1, 2, 3, …, *k-1*is separated. PACF is also a function describing the partial correlation value between the observation in the time *t* (notated by Z_t) with the observation in the previous times (notated by $Z_{t-1}, Z_{t-2},\ldots\ldots Z_{t-k}$).

White noise process

A process {a_t} is named *white noise process* (free and identical) if the sequential variables is not correlative each other and follows certain distribution.The mean $E(a_t) = \mu_a$ of process is assumed zero and has constant variance, i.e. $Var(a_t) = \sigma_a^2$ and the value of covariance for this process, $\gamma_k = \mathrm{cov}(a_t, a_{t+k}) = 0$ for $k \neq 0$. A time series is *white noise process* if its mean and variance is constant and independent each other.

Differencing

Orde **d** in ARIMA (p, **d**, q) is used to model the event which is not stationer in the term of mean where **d** denotes differencing. Basically, *differencing* method forms a new data obtained by decreasing the value of observation in time *t* and the value of the observation in the previous time. Commonly, *differencing* operation resulting a new stationary event (process), e.g. W_t is:

$$W_t = Z_t - Z_{t-1}$$

Transformation

If the condition of stationer in variance is not satisfied, Box and Cox (1964) introduced *power transformation* $Z_t^{(\lambda)} = \frac{Z_t^{(\lambda)} - 1}{\lambda}$, where λ is called transformatio parameter. The transformation is also familiar as Box-Cox Transformation.

Operator Backshift (B)

Box-Jenkins method is used for time series analysis using *backshift* B operator defined as:

$$BZ_t = Z_{t-1}$$

Notation B put in Z_t has displacement effect of data in one past period.

Seasonal Model

The common form of seasonal ARIMA Box-Jenkins model of ARIMA $(p,d,q)(P,D,Q)^S$ is as follows:

$$\phi_p(B)\Phi_P(B^s)(1-B)^d(1-B^s)^D\dot{Z}_t = \theta_q(B)\Theta_Q(B^s)a_t$$

2.2 TRANSFER FUNCTION

Transfer function model is one of qualitative prediction model which can be used for mulivariate time series data prediction. The model can be used to describe one or more dynamic effect of input series on output series. It also can be used to obtain the prediction values simultaneously, i.e. input series and output series. The aim of the transfer function modellig is to determine simple model relating Y_t and X_t and N_t or to determine the role of certain indicator (*input* series) to determine (*output* series). The determination is closely related to ARIMA modelling (Makridakis, wheelwright & McGee, 1999).

The common pattern of bivariate transfer function model (*single-input*, x_t and *single-output*, y_t) is as follows:

$$y_t = v(B)x_t + n_t$$

Next, the transfer function $v(B)$ can be formulated as:

$$v(B) = \frac{\omega_s(B)B^b}{\delta_r(B)}$$

The Stage of Transfer function Formation

a. Model Pattern Identification

- Preparing the form of *input*and*output*series.
- *Prewhitening*of input series (x_t).
- *Prewhiteningoutput* series (y_t)
- The calculation of cross-correlation and autocorrelation for input and output serieswhich has been under *prewhitening*.
- Direct estimation of impulse response value.
- The determination of (b, r, s) for transfer functionmodel.
- Initial approximation of constraint series(n_t).
- The determination of (p, q) of ARIMA model (p, 0, q) of constraint seriesn_t.

b. Approximation of transfer function model parameters.

c. Diagnostic test for transfer function model.

- Autocorrelation check for modelresidual.
- Calculation of cross-correlation between residual and constraint series which is*prewhitening.*

2.3 FORECASTING

Forecasting is a use of past data from a variable or group of variables to estimate their values in the future. In the other words, forecasting predicts the value of something in the future based on the past data which is analyzed naturally using statistics method.

3. RESEARCH METHOD

The present research is applied research to study the application of transfer function model in time series analysis to forecast the pattern of rainfall. The data in this research is secondary data, i.e. rainfall and temperature furing the last 100 month obtained from BadanMeteorologi, Klimatologi dan Geofisika (BMKG) in district IV Makassar. Analysis data technique which was used was statistical analysis using transfer function model in time series analysis. The stages of this research were (1) relevant literature and research finding review, (2) identification and sampling of data, (3) data tabulation and analysis, (4) data analysis interpretation, (5) discussion, and (6) conclusion and suggestion.

4. RESULT AND DISCUSSION

The research findings are described as follows:

Identification for *output*series and *input*series

Timeseriesplot of the rainfall data for the original data shows definitely fluctuative data meaning that the data are not statinonary in the term of mean which implies*differencing*.

The result of the plot of ACF and PACF after *defferencing*for one season (D=1).

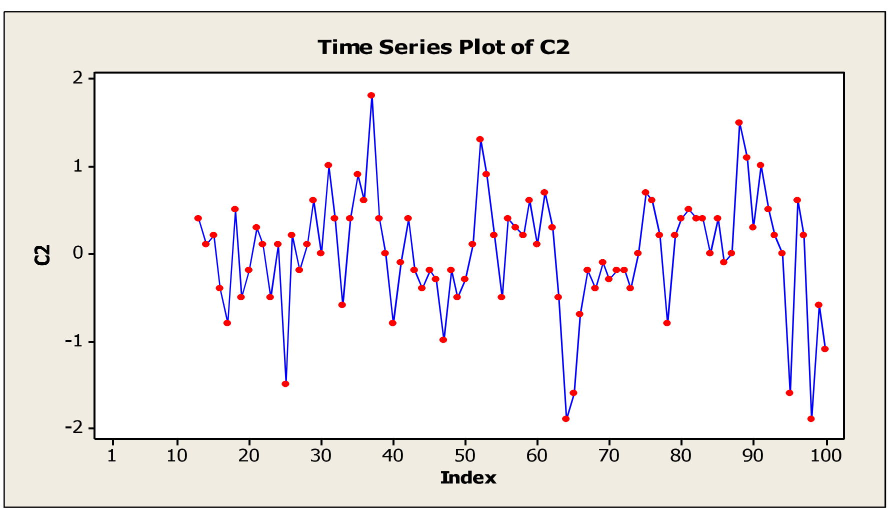

Fig. 1. Time SeriesPlot

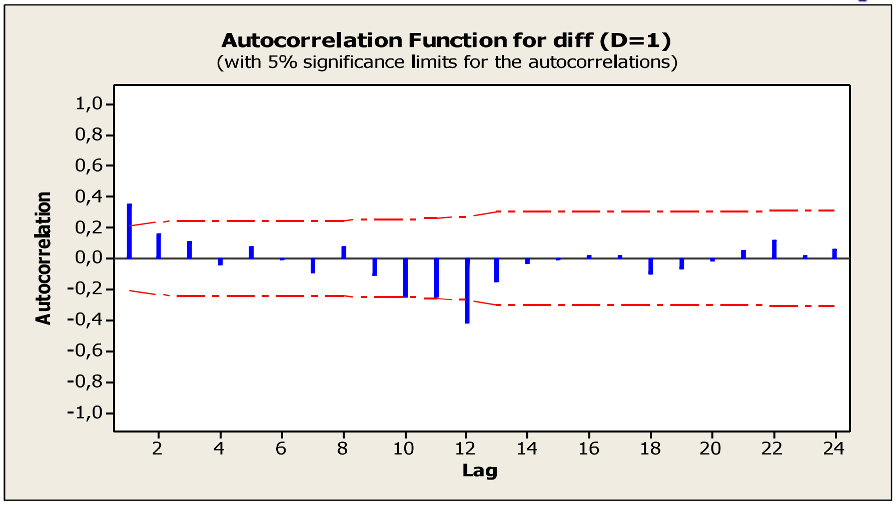

Fig. 2. Autocorrelation Function (ACF)Plot

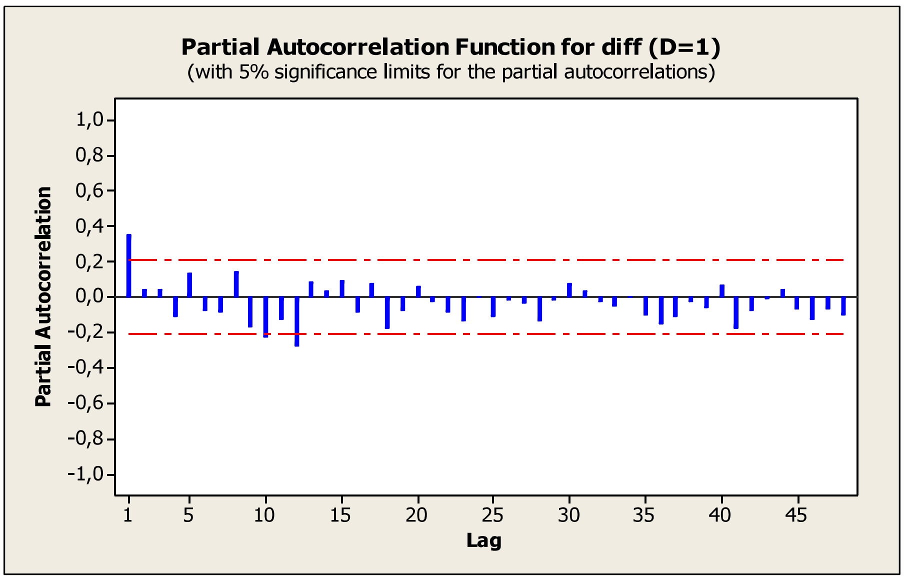

Fig. 3. Partial Autocorrelation Function (PACF) Plot

Based on the *time series*, FAK and FAKP plots, *differencing* result suggests that the data is stationer with the estimation of appropriate model, i.e. $(0{,}0{,}0)(2{,}1{,}0)^{12}$.

The analysis suggests that all parameters in the model are significant. The remains have satisfied the *white noise*condition and normally distributed, then the appropriate model for *input*series (air temperature) is ARIMA model $(0{,}0{,}0)(2{,}1{,}0)^{12}$. Mathematically it can be written as:

$$\Phi_2(B^{12})(1-B^{12})X_t = a_t$$
$$(1-\Phi_1 \mathrm{B}^{12}-\Phi_2 \mathrm{B}^{24})(1-B^{12})X_t = \alpha_t$$
$$(1-\Phi_1 \mathrm{B}^{12}-\Phi_2 \mathrm{B}^{24})x_t = \alpha_t$$

Where$x_t = (1-B^{12})X_t = X_t - X_{t-12}$

Next, the stage of *prewhitening*is conducted for inputseries, output series, cross-correlation for both input and output which is under *prewhitening*, the direct estimation of impulse response value, the identification of (*r, s, b*) for transfer function model, the initial estimation of constraint series($\boldsymbol{n_t}$)and autocorrelation calculation, and partial autocorrelation. Then the value of parameter estimation of transfer function for the noise series is ARIMA $(0{,}0{,}1)^{12}$i.e.:

$$Z_t = a_t - \Theta_1 a_{t-12}$$
$$Z_t = (1-\Theta_1 B^{12})a_t$$
$$n_t = (1-\Theta_1 B^{12})a_t$$

Then, the best model for noise series is $n_t = (1 - \Theta_1 B^{12}) a_t$

The best transfer function model for rainfall and temperature data is as follows:

$$y_t = \frac{(\omega_0)}{(1-\delta_1 B-\delta_2 B^2)} x_t + (1 - \Theta_1 B^{12}) a_t$$

Transfer function model completed with the new parameter estimation is as follows:

$$y_t = \frac{(-3.0144)(X_t - X_{t-12})}{(1+0.4485B+0.7610B^2)} + (1 - 0.6891 B^{12}) a_t \text{atau}$$

$$Y_t = (-0.4485)Y_{t-1} - 0.7610Y_{t-2} + Y_{t-12} + 0.4485Y_{t-13} + 0.7610Y_{t-14} - 3.0144X_t + 3.0144X_{t-12} + a_t + 0.4485a_{t-1} + 0.7610a_{t-2} - 0,6891a_{t-12} - 0,3091a_{t-13} - 0,5244a_{t-14}$$

after the bivariate model from the rainfall is obtained, the prediction of the square root value for twelve next periods is conducted using the criteria that the rain season begins when the rainfall is more than 50 mm for three consecutive decades and the dry season begins when the rainfall is less than 50 mm for three consecutive decades.

5. CONCLUSION

Transferfunctionmodel suggest that the value of quadratic square of raiinfall in the time *t*is influenced by the quadratic square of the rainfall in the last month(Y_{t-1}), two months ago months(Y_{t-2}), twelve monthsago (Y_{t-12}), thirteen months ago(Y_{t-13}), and fourteen months ago $(Y_{t-14})$and is influenced by the temperature in the same months respectively, (the t-th monthhor(X_t))and twelve months ago (X_{t-12}). Besides that it is influenced by the residual value in the same month, one month ago(a_{t-1}), two months ago(a_{t-2}), twelve months ago(a_{t-12}), thirteen months ago(a_{t-13})andfourteen months ago(a_{t-1}).

ACKNOWLEDGMENT

Further thanks to all those who helped research and writing articles. hereinafter also conveyed his gratitude to those who have provided research funding of PNBP.

REFERENCES

[1] Aswi&Sukarna. *AnalisisDeretWaktu*. Makassar: Andira Publisher. 2006.

[2] Fathurahman, M. *Pemodelan Transfer function Multi Input*. Jurnal. Samarinda: jurusanstatistika FMIPA universitasmulawarman. 2009.

[3] Latikan B. *Dasar-DasarKlimatologi*. Jakarta: PT. Raja GrafindoPersada. 1994.

[4] Makridakis, S., wheelwright, S.C., & McGee, V.E., *MetodedanAplikasiPeramalan*. Jilid 1. EdisiKedua. Jakarta: Erlangga. 1998.

[5] Martika. PeramalanJumlahPengunjungdenganMetodeDekomposisisertaKontribusiJumlahPengunjungTerhadapPendapatanobjek WisataPantaiPurwahamba Indah Tegal. Skripsi. Semarang: JurusanMatematika FMIPA UniversitasNegeri Semarang. 2007.

[6] Rahmawati. *AnalisisDeretBerkalaBivariatdenganMenggunakan Model Transfer function*. Skripsi. Makassar: JurusanMatematika FMIPA UniversitasNegeri Makassar. 2007.

[7] Raodhatullah. *Pemodelan Data Curah Hujan Di Kota Makassar DenganMenggunakanMetode ARIMA Box-Jenkins*. Skripsi. Makassar: JurusanMatematika FMIPA UniversitasNegeri Makassar. 2006.

[8] Suharnita, *StudiTentang Model Transfer function untukMeramalkan Curah Hujan Di Kota Makassar*. Skripsi. Makassar: JurusanMatematika FMIPA UniversitasNegeri Makassar. 2011.

[9] Supranto, J. *StatistikaTeoridanAplikasi*. Jakarta: Erlangga. 2000.

[10] Subagyo, P. *Forecasting KonsepdanAplikasi*. Yogyakarta :BPFE Yogyakarta. 1986.

[11] Sudjana. *MetodeStatistika*. Bandung: Tarsito. 1989.

[12] Wardhani, S.L &Algifari. *TeknikProyeksiuntukBisnisdanEkonomi*. EdisiPertama. Yokyakarta: BPFE. 2007.

[13] Wei, W. W. S. *Time Series Analysis: Univariate and Multivariate Methods*. California: Addison-Wesley Publishing Company. 1994.