Journal of Theoretical and Applied Information Technology
15th October 2020. Vol.98. No 19

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

APPLICATION OF HIERARCHICAL TIME SERIES MODEL WITH TRANSFER FUNCTION

RAHMA FITRIANI [1], ACHMAD EFENDI[2], BAYU RAHAYUDI[3]

[1,2] Department of Statistics, Universitas Brawijaya Malang, Indonesia

[3] Department of Informatics Engineering, Universitas Brawijaya Malang, Indonesia

E-mail: [1]rahmafitriani@ub.ac.id, [2]a_efendi@ub.ac.id , [3]ubay1@ub.ac.id.

ABSTRACT

The prediction of the amount of water needed in each region is important since clean water is currently needed by every community worldwide. The monthly water consumption depends on how many users. The users vary from people, industry, agriculture, etc. Water providers, in this case local governments, need to estimate the needs in the short- and long-term periods. In the long run, water demand prediction is beneficial for policymakers as the population and water users are increasing such that efforts are needed to supply local water needs. In this study, we aim at predicting the need for monthly water consumption based on users or customers in the city of Malang, East Java, Indonesia and in each district within the city of Malang. There are five districts within this city. The water demand data are stratified from districts and city such that the modeling can accommodate hierarchy. Modeling and forecasting are done with ARIMA (Autoregressive and Moving Average) and transfer function models to accommodate one or several variables affecting water consumption. Accuracy results (Symmetric Mean Absolute Percentage Error) indicate an accuracy error of approximately 3 percent. This result is quite satisfactory and can be used to estimate the accuracy of water demand in Malang city and its districts.

Keywords: *Water, Time Series, Hierarchical, Prediction, Transfer Function*

1. INTRODUCTION

Forecasting is an activity needed to determine if an event occur, so that the right action can be taken. In making quantitative forecasting, we must adjust to the characteristics of existing data. If a time series data has a multilevel/hierarchical structure with a certain level, then it can be called a hierarchical time series data [1]. A forecasting hierarchy should be applied to data with these criteria. The hierarchical time series framework is a forecasting disaggregation that allows for forecasting together at all levels of the hierarchy [2,3]. Hierarchical forecasting requires that the data have a complete structure at each level [4]. Hierarchical forecasting is designed to produce coherent forecasting based on aggregation structures [5]. In other words, the amount of aggregated forecasting is equal to the number of forecasting of each aggregation component [1,6,7]. There are two commonly used hierarchical forecasting approaches, they are the bottom-up and top-down approaches.

Research on hierarchical time series data predicted the arrival of local tourism in Australia with an exponential smoothing forecasting model on hierarchical time series data to two levels [2]. The use of ratio trend variation (RTV) was able to grip the fuzziness in time series data and to reduce the estimation of forecasting errors [8]. Modeling and predicting the export value of Central Java, Indonesia, showed that for the Export Value, Bottom-up approach with Hybrid Autoregressive Integrated Moving Average - Radial Basis Function Neural Network (ARIMA-RBFNN) modeling can be used for long-term predictions [4]. All studies showed that hierarchical forecasting can be applied to various time series models.

There are two types of forecasting models, namely the scale series model and the causal or regression model [9,10]. One example of a causal model is the transfer function model. It assumes that the predicted factor shows a causal relationship to one or more independent variables. Transfer function is a model for solving problems if there are more than one time series data and one of the variables influences others [11]. The time series data mining technique is able to produce good model to forecast the distribution of rainfall pattern accurately [12] while research on the transfer function and time series modeling was carried out [13].

Daily data of urban water at East Doncaster zone in 1991-1999 were modeled using regression and time series methods [14]. Short-term water demand was able to be modelled using a fixed seasonal auto-regressive (FSAR) model and an adaptive seasonal auto-regressive (ASAR) [15]. Meanwhile, average demand patterns for different consumer profiles could be modelled using spatial and temporal forecasting methods [16]. We then come up with the hypothesis that there is a relationship between the consumption of PDAM (Perusahaan Daerah Air Minum; or district/city/local water company) water with the number of customers.

PDAM is one of the regional-owned business units that acts as a provider and distributor of clean water for the community. It continues to improve services and supply of clean water that is in line with water quality standards and community needs. PDAM is a business unit that is spread in almost all cities in Indonesia, one of which is in the city of Malang, East Java, Indonesia. At the end of 2017, the Malang Central Statistics Agency noted that the number of PDAM customers in Malang was approximately 156,216. PDAM customers are classified into several groups which include social groups, households, agencies, commerce, industry, and so forth. These customers are spread in various regions in Malang City, namely ABRI (Indonesian Army), regions of Blimbing, Karangploso, Kedungkandang, Sukun, Lowokwaru, Klojen, Pakis, and Pakisaji.

From this information and history, we are interested in conducting research on forecasting PDAM water consumption in Malang city based on the number of customers. It is also of interest to know the forecasted water demand in the second most populous region in East Java Province. This province is one of populous province, with about 40 million people, in Indonesia. This then motivates us to study the forecasting of water demand in Malang City. Forecasting in the next few periods is done through the hierarchy forecasting method. The approach used in hierarchical forecasting is bottom-up and top-down methods. These forecasting methods are carried out because PDAM water consumption data in Malang City has the structure of hierarchical time series. Level 0 is data on water consumption and the number of customers in Malang city. While level 1 is data on water consumption and the number of customers in the Malang sub-regions of Blimbing, Kedungkandang, Sukun, Lowokwaru, Klojen regions. The modeling is done using a single input transfer function model. The input series is the number of PDAM customers whereas the output series is the PDAM water consumption.

2. LITERATURE REVIEW

2.1 Identification and Estimation of the Transfer Function Model

Preparation for a series of inputs and outputs includes transformation and differentiating the series if there are nonstationarity. The nonstationary of variance is overcome by transformation while the mean nonstationarity is done through differentiation. The stationary series in mean and variance are called y_t and x_t [9]. The three key parameters in the transfer function model are b, r, s. Order r indicates the degree of function $\delta(B)$, s indicates the degree of function $\omega(B)$, and b indicates the delay in X_{t-b}. The meanings of these parameters are:

a. The value of b states that y_t is not affected by x_t until the period $t+b$ or $y_t = 0x_t + 0x_{t-1} + 0x_{t-1} + \ldots + \omega_0 x_{t-b}$

b. The value of s indicates the length of the output series y_t continuously affected by new values from the x_t input series, or y_t is affected by $x_{t-b}, x_{t-b-1}, \ldots, x_{t-b-g}$.

c. The value of r indicates that y_t is related to the previous values, or y_t is influenced by $y_{t-1}, y_{t-2}, \ldots, y_{t-r}$.

The transfer function model above is estimated as: $\delta = (\delta_1, \ldots, \delta_r)'$, $\omega = (\omega_0, \omega_1, \ldots, \omega_s)'$, $\phi = (\phi_1, \ldots, \phi_p)'$, $\theta = (\theta_1, \ldots, \theta_q)'$ and σ_α^2 hence:

$$\delta(B)\phi(B)y_t = \phi(B)\omega(B)x_{t-b} + \delta(B)\theta(B)a_t$$

$$c(B) = d(B) + e(B)a_t$$

$$c(B) = \delta(B)\phi(B) \quad = (1 - \delta_1 B - \ldots \delta_r B^r)$$

$$\begin{aligned} c(B) = \delta(B)\phi(B) \quad &= (1 - \delta_1 B - \ldots \delta_r B^r)(1 - \phi_1 B - \ldots - \phi_p B^p) \\ &= (1 - c_1 B - \ldots - c_{p+r} B^{p+r}) \end{aligned}$$

$$\begin{aligned} d(B) = \phi(B)\omega(B) &= (1 - \phi_1 B - \phi_1 B^r)(\omega_0 - \omega_1 B - \ldots - \omega_s B^s) \\ &= (d_0 - d_1 B - \ldots - d_{p+1} B^{p+s}) \end{aligned}$$

$$\begin{aligned} e(B) = \delta(B)\theta(B) &= (1 - \delta_1 B - \ldots - \delta_r B^r) = (1 - \theta_1 B - \ldots - \theta_q B^q) \\ &= (1 - e_1 B - \ldots - e_{r+q} B^{r+q}) \end{aligned}$$

with c_i, d_j, e_k is a transfer function from $\delta_i, \omega_j, \phi_k$ and θ_i. The used parameter estimation method is the Conditional Maximum Likelihood method. It is assumed that a_t is a white noise series that is normally distributed $N(0, \sigma_\alpha^2)$. We get the likelihood function as the following:

$$L(\delta,\omega,\phi,\theta,\sigma_\alpha^2 \mid b,x,y,x_0,y_0,a_0)$$
$$=(2\pi\sigma_\alpha^2)^{-\frac{n}{2}} e^{-\frac{1}{2\sigma_\alpha^2}\sum_{t=1}^{n} a_t^2}$$

with,

$$a_t = y_t - c_1 y_{t-1} - \ldots - c_{p+1} y_{t-p-1} - d_0 x_{t-b} + d_1 x_{t-b-1} +$$
$$\ldots + d_{p+s} x_{t-b-p-s} + e_1 a_{t-1} + \ldots + e_{r+q} a_{t-r+q}.$$

The estimated parameter of the transfer function model uses Conditional Maximum Likelihood which assumes $e_t = 0$ can be solved through nonlinear least square method by minimizing the following functions [17]:

$$S(\delta,\omega,\phi,\theta \mid b) = \sum_{t=1}^{n} a_t^2$$

with $t_0 = \max(p+r+1, b+p+s+1)$.

2.2 Diagnostic Checking of Transfer Function Model

The properness of the first model is checked through the residual autocorrelation a_t. The autocorrelation of the residual is used to determine the property of the white noise of the residual. The remaining autocorrelation calculation has the following hypothesis:

H_0: $\rho_1 = \rho_2 = \ldots = \rho_k = 0$

H_1: $\exists \rho_i \neq 0$, $i=1, 2,\ldots, k$

Test statistic for white noise assumption is

$$Q_1 = m(m+2)\sum_j^K (m-j)^{-1} \hat{\rho}_{\hat{\alpha}}^2(j).$$

Q follows the distribution of χ^2 with degree of freedom ($K-p-q$) which only depends on the parameters of the white noise model.

The second requirement to make the model feasible is that the cross-correlation between a_t and input series a_{t-1} is not significant. This calculation is used to test the independence of the noise series a_t and the input series a_{t-1}. The test hypothesis used is as the following:

H_0: no cross-correlation between a_t and a_{t-1},

H_1: there is cross-correlation between a_t and a_{t-1}.

Test statistics for the correlation between a_t and a_{t-1} is $Q_0 = m(m+2)\sum_j^K (m-j)^{-1} \hat{\rho}_{\hat{\alpha},\hat{\alpha}}^2(j)$,

with Q_0 follows χ^2 distribution with degree of freedom of $(K + 1) - M$; M is the number of parameter (ω_r and δ_r) and $m = n - t_0 + 1$.

2.3 Forecasting with the Transfer Function Model

It is important to note that the data are stationary for forecasting a time series model. Non-stationary data will affect the results of forecasting since it has been subjected to transformation or differencing. Suppose that Y_t and X_t are temporary non-stationary series, then $(1-B)^d Y_t$ and $(1-B)^d X_t$ is a stationary series. The series if stated in the form of transfer function model are as follows

$$(1-B)^d Y_t = \frac{\omega(B)}{\delta(B)}(1-B)^d X_t + \frac{\theta(B)}{\phi(B)}$$

and $\phi_x(B)(1-B)^d X_t = \theta_x(B) a_t$.

In order to have forecasting value we can modify the equation above and get the following equation

$$c(B)Y_t = d(B)X_{t-b} + e(B)a_t.$$

In this study, modeling of the transfer function is carried out on all variables, both at level 1 and 0 (Malang City/region and its 5 sub-region levels). Forecasting using the transfer function on all output variables is called the initial forecasting and is expressed in the form of: $\hat{Y}_{T,t}$: forecast value of total water consumption of Malang City PDAM, $\hat{Y}_{1,t}$: the forecasting value of PDAM water consumption in Blimbing, $\hat{Y}_{2,t}$: the forecasting value of PDAM water consumption in Kedungkandang, $\hat{Y}_{3,t}$: the forecasting value of PDAM water consumption in Klojen, $\hat{Y}_{4,t}$: PDAM water consumption forecasting in Lowokwaru, and $\hat{Y}_{5,t}$: the forecasting value of PDAM consumption in Sukun.

2.4. Hierarchical Time Forecasting Method

Hierarchical time series framework is a forecasting disaggregation that allows for forecasting to be done together at all levels of the hierarchies [2]. There are three types of approaches commonly used in forecasting hierarchies: top-down, bottom-up, and the optimal combination approach. The top-down approach can be divided into two, namely top-down with historical proportions and top-down with forecasting proportions [5].

In this study only two approaches were used, the top-down and bottom-up ones. We also use hierarchy up to level one. For example, the value $K = 2$ indicates the number of levels. The zero level is called the most aggregated level. The total series at zero level is disaggregated into two series at level one. Series A is disaggregated into three series at level two while Series B is disaggregated into two series at level two.

Observation at the t-time (t = 1,2, ..., n) of the series connected to nodes i (i = A, B, AA, AB, AC, BA, BB) on the hierarchical structure expressed in $Y_{t,\,t}$. Suppose the observation at the t-time at node A can be expressed as $Y_{A,\,t}$. While $Y_{T,\,t}$ represents the total aggregated value at time t. The number of time series at the jth level is expressed as m_j and the total time series is expressed by $m = m_0 + m_1 + ... + m_k$, $j = 0,1,2,...,K$

2.5. Selection of the Best Forecasting Method

In hierarchical forecasting, modeling is initially performed on all the rows in the hierarchy and then the best model is chosen precisely. The best model is used to forecast data and then the revised forecasting is done using a bottom-up and top-down approaches [4]. The selection of the best forecasting method is done using sMAPE (Symmetric Mean Absolute Percentage Error). The best forecasting result are those with the smallest criterion value of other results. The calculation of sMAPE is as follows [9]:

$$sMAPE = \frac{1}{H}\sum_{h=1}^{H}\frac{|Y_{i,h} - \hat{Y}_{i,h}|}{1/2(Y_{i,h} + \hat{Y}_{i,h})}x100\%$$

H : the number of data of out samples
$Y_{i,\,h}$: value of out sample in period h that corresponds to node i
$\hat{Y}_{i,h}$: predicted value for period h that corresponds to node i

3. RESEARCH METHOD

The data used in this study are secondary data obtained from the Regional Water Supply Company (PDAM) of Malang City, East Java, Indonesia. The water consumption data contain the number of monthly customers in the period of January 2011 until December 2018. It is divided into in-sample data, which are data in January 2011-June 2018 for modeling, and out-sample data, that contain data in July until December 2018 for forecasting.

The data consist of two variables, PDAM water consumption as an output variable and the number of PDAM customers as input variables. The output variable at level zero is the volume of drinking water consumption in Malang city while at level one is the volume of PDAM water consumption in sub-regions of Blimbing, Kedungkandang, Klojen, Lowokwaru, and Sukun areas. The input variable at level zero is the number of PDAM customers in Malang city while at level one is the number of PDAM customers in the sub-regions of Blimbing, Kedungkandang, Klojen, Lowokwaru, and Sukun areas.

The steps of ARIMA modeling used as part of the transfer function modeling include:

1. Preparing a series to be modeled by ARIMA.
2. Identifying time series data through plots. Then looking at the stationarity of the data on variety and averages. If it is not stationary to the average, then differencing is carried out, while the non-stationarity of the variance is overcome by Box-Cox transformation.
3. Identifying the ARIMA model through the ACF (Auto-correlation Function) and PACF (Partial Auto-correlation Function) plots and determining the tentative model.
4. Estimating the parameters p, d, q in ARIMA model.
5. Testing the parameter significance of the ARIMA model. If there is one non-significant parameter, then choosing another tentative model.
6. Testing the feasibility of the ARIMA model to check the assumptions of the remaining white noise.

Once the best transfer function model is obtained then it is used to predict data using the hierarchical forecasting method. The forecasting results using the transfer function model called initial forecast. The next step is to revise forecasting using the hierarchy forecasting method through the following approach:

1. Bottom-Up Approach
   a) Calculating the forecasting hierarchical time series data on all level 1 variables with the transfer function model for the next 6 periods.
   b) Calculating the forecasting hierarchical time series on level 0 by adding up the forecasting results in step *a*.
2. Top-Down Approach with Historical Proportion (TDHP)
   a) Calculating the forecasting hierarchicall time series data at level 0 with the transfer function model for the next 6 periods.
   b) Calculating historical proportions for level 1 hierarchical time series with TDHP 1 and TDHP 2 proportions.
   c) Calculating the level 1 hierarchical forecasting by multiplying the historical proportion in step *b* with the forecasting result in step *a*.
3. Top-down Approach to Proportion Forecasting
   a) Calculating the forecasting hierarchical time series data at level 0 with the transfer function model for the next 6 periods.

Journal of Theoretical and Applied Information Technology
15th October 2020. Vol.98. No 19

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

b) Calculating the proportion of forecasting for *a* level 1 hierarchy time series.
c) Calculating the level 1 hierarchical forecasting by multiplying the proportions in step *b* with the forecasting results in step *a*.

4. Determining the best forecasting method.
This can be done by comparing the value of sMAPE and then choosing the approach with the smallest criteria. In the whole process of data analysis, software R 3.5.2 is used.

4. ANALYSIS AND RESULT

4.1. Data Description

A time series plot is able to explain the behavior of time series data in general. The following is a plot of the input and output time series used in modeling the transfer function in this study

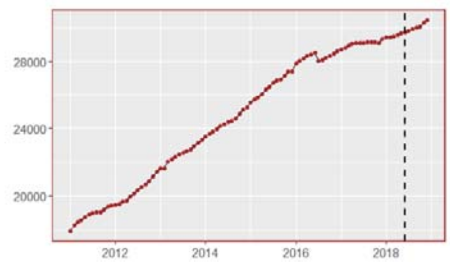

Figure 4.1. Time Series Plot of Number of PDAM Customers in Malang Based on Regional Aggregation for the Period 2011 – 2018 (x-axis: Time; y-axis: Customers)

It can be seen that the number of PDAM customers in Malang city tends to increase by month (Figure 4.1). Changes in the number of customers by time are not too volatile and are quite consistent. The nature of the aggregated series (customers in Malang city) is also the same as the results of the disaggregation series (PDAM customers in each area or district). Descriptively, it can be seen that the input series in Blimbing, Kedungkandang, Klojen, Lowokwaru, Sukun, and Malang City respectively have intercept values of 20000, 20000, 19000, 20000, 20000, and 100000. It is reflected that the number of PDAM customers in Malang City is the result of aggregation of the number of PDAM customers in its 5 districts.

There is something interesting about the input series plot in Klojen and Lowokwaru districts, from 2017 to 2018 it appears that the input series plot is starting to be flat. This indicates that the increase in the number of customers in the two districts is starting to shrink. As per 2019, 93% of the people of Malang City have been registered as PDAM customers. The majority of the people of Lowokwaru and Klojen districts have used PDAM services. While the Malang City PDAM is still working to develop services in the Kedungkandang district [18]. This is in line with the input series plots in Kedungkandang, which continued to rise until 2018.

As PDAM customers increase from time to time, PDAM water consumption also tends to increase per year. Graphically the behavior of the input and output series is appropriate and has an upward linear trend. It can be seen that the output series in districs of Blimbing, Kedungkandang, Klojen, Lowokwaru, and Sukun as well as Malang City have intercept values of 300000, 300000, 360000, 400000, 400000, and 1600000. It is descriptively reflected that PDAM water consumption in Malang City is the result of aggregation of PDAM water consumption in 5 districts. The value of PDAM water consumption is quite fluctuating by time. This occurs as changes in water consumption are not only influenced by the number of customers but can also be influenced by community activities, customer groups in the sub-district (shopping centers, households, places of worship, and government agencies), and the use of well water.

PDAM water consumption in Lowokwaru district has a unique characteristic, with data fluctuation that follows a seasonal pattern of around 6 months. This is not clearly seen descriptively and needs further verification statistically. Clean water supply for the Lowokwaru is largely supported by PDAM since the groundwater quality is quite low. Lowokwaru district is a center of education, especially major universities in Malang city, shopping centers and is populated by students living in boarding houses. Indirectly, community activities in Lowokwaru district follows the educational holiday season such that it also influences PDAM water consumption at that time. Similar to water consumption in Lowokwaru district, water consumption in Klojen district is quite fluctuating by time. This is caused by the characteristics of customers in Klojen which are mostly education centers or schools, shopping centers, and city centers. On the other hand, most people use a combination of PDAM water and well water [18].

Broadly speaking, it can be said that the aggregated series is similar in nature to the aggregated results in terms of data trends. Figure 4.1 also shows that the series is still coherent or the value of level 1 series is equal to the level 0 series. One way to maintain the coherent nature of forecasting is by the hierarchical forecasting method. Trend patterns in a time series indicate that

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

the data is not stationary to the average. This occurs in data on the number of customers and water consumption. Stationarity for averages and variances needs to be tested before data are modeled.

4.2. Stationarity of Time Series Data

The data stationarity is the first assumption that must be met when modeling time series data. First, the stationarity of a series for variance can be seen through the values λ in the Box-Cox transformation. Table 4.1 shows that the values of λ of all the variables in the hierarchical structure were not initially stationary to variance since the values λ were not equal to one. Then after the appropriate transformation, variables with a value of λ equals one is obtained. After the data met the assumption of variance stationarity, then it needs to be tested for stationarity of average. This test is performed using the Augmented Dickey Fuller test.

All input series $X^*_{i,t}$ are not stationary to the average because the p-value is greater than α (0.05). After the first differentiation is made on the input variable, only the variable of $X^*_{2,t}$, $X^*_{3,t}$ and $X^*_{5,t}$ are stationary, while the variable of $X^*_{T,t}$, $X^*_{1,t}$, $X^*_{4,t}$, requires two differentiation such that it is stationary to the average. Similar to the input variables, all output series of $Y^*_{i,t}$ are not stationary to the average as the p-value of the Augmented Dickey Fuller test is greater than α. After the data is differentiated once, the variables of $Y^*_{T,t}$, $Y^*_{1,t}$, $Y^*_{2,t}$, $Y^*_{3,t}$, $Y^*_{4,t}$ and $Y^*_{5,t}$ become stationary to the average.

From the six-output series, the series $Y^*_{4,t}$ are indicated as seasonal. Seasonal effects are seen in ACF series plots $\nabla^{-1}Y^*_{4,t}$, i.e. $Y^*_{4,t}$ series that are stationary to mean. Hence, the sequencing $Y^*_{4,t}$ needs to be done by differencing lag 12. The existence of this seasonal effect is supported by the characteristics of PDAM water use in the Lowokwaru sub-district as explained before.

Beside the Dickey Fuller Test, stationarity with the average can be seen from the ACF plot. Non-stationary time series are marked with ACF values that significantly exceed the first 3 consecutive lags and decrease exponentially. The results from the Dickey Fuller test are consistent with the ACF plot that was formed.

Tabel 4.1 Box-Cox Transformation Value λ of All Research Variables

Variable	λ	Transformed Variable	λ *
$X_{T,t}$	1.4856	$X^*_{T,t}$	1.0000
$Y_{T,t}$	1.3935	$Y^*_{T,t}$	0.9999
$X_{1,t}$	1.4318	$X^*_{1,t}$	0.9999
$Y_{1,t}$	1.6325	$Y^*_{1,t}$	1.0000
$X_{2,t}$	0.6248	$X^*_{2,t}$	0.9993
$Y_{2,t}$	0.7017	$Y^*_{2,t}$	0.9999
$X_{3,t}$	5.4741	$X^*_{3,t}$	1.0000
$Y_{3,t}$	1.1208	$Y^*_{3,t}$	0.9999
$X_{4,t}$	1.9361	$X^*_{4,t}$	1.0000
$Y_{4,t}$	1.2695	$Y^*_{4,t}$	0.9999
X_t	1.6517	$X^*_{5,t}$	1.0000
$Y_{5,t}$	1.3750	$Y^*_{5,t}$	0.9999

4.3. Transfer Function Modeling on All Output Variables

4.3.1. ARIMA Input Series Modeling

The input series in the transfer function modeling first must be modeled by ARIMA (*p, d, q*). ARIMA modeling of input series will be useful for whitening the input and output series. The ARIMA model identification of the input series is determined through the ACF and PACF plots of the input series which have been stationary. The identification of the ARIMA *p, q* model order is done through ACF and PACF plots. Order *p* which is an indication of the autoregressive model is based on the lag when partial autocorrelation is significant and can be seen in the PACF plot. While the order *q* which is an indication of the Moving Average model that is based on the lag when autocorrelation is significant and can be seen in the ACF plot.

Based on the analysis, the main tentative ARIMA model can be formed, ARIMA (3[9],2,1) for the series $X^*_{T,t}$, ARIMA (1 [6],2,3[6]) for the series $x^*_{1,t}$, ARIMA (1,1,1) for series $x^*_{2,t}$, ARIMA (1[13],1.2) for series $x^*_{3,t}$, SARIMA (Seasonal ARIMA) (3,2,1)(1,0,0) 12 for series $x^*_{4,t}$, and ARIMA (0,1,[3]) for $x^*_{5,t}$. From the main tentative model, estimation and significance parameter, testings are performed. Estimating the parameters to select the right model can be done by experimenting with a combination of lag values that are out of bounds [9].

The best ARIMA model of each input series that has overall significant parameters at a minimum level of 0.05, are the ARIMA model (0,2,1) for the series $x^*_{T,t}$, ARIMA (3[6],2,0) for the series $x^*_{1,t}$, SARIMA (0.2,1)(1,0,0)^{12} for the series $x^*_{4,t}$, and ARIMA (0,1,[3]) for the series $x^*_{3,t}$. While the ARIMA model (1,1,0) for a series $x^*_{2,t}$ of all parameters can be said to be significant at the 0.01. The ARIMA model (1,1,0) remained selected since among the three tentative models that were tried, only the parameters of those models approached significant criteria. Likewise, with the ARIMA model (0,1,2[5][12]) for the series $x^*_{5,t}$, all parameters can be said to be significant at the 0.10 level. The ARIMA model (0,1,2[5][12]) is still chosen because among the several models that have white noise residuals, the ARIMA model (0,1,2[5][12]) has the smallest AIC value and is able to make the input sequence to be white noise.

Based on the results of the Ljung Box test, the best model of each series of inputs is feasible. The model is said to be feasible as it gives a conclusion of accepting the null hypothesis in the Ljung Box test. Hence, the six models of the input series have been able to capture the factors that must be included in the model.

4.3.2. Whitening of Input and Output Series

Whitening of the input sequence requires that the input series be white noise. The following are the whitening inputs based on the formed ARIMA model. Each $x^*_{i,t}$ is a series that has been stationary of variance and average. As with whitening on the input sequence, to maintain the integrity of the functional relationship the output sequence must also be "whitened". Whitening in the output sequence is the same as changing α to β and *x* to *y*. Each $y^*_{i,t}$ is a series that has been stationary of variance and average.

4.3.3. Calculation of Cross-Correlation and Weight of Impulse Response Between $\alpha_{i,t}$ and $\beta_{i,t}$

Cross-correlation (CC) is done between the input and output series of the whitening result. In the CC function (CCF) plot, it is expected that the input sequence will affect the output one. This cross-correlation plot will assist the process of identifying the order of *b*, *s* and *r*. In order to find out the one-way causal relationship, the number of customers influencing PDAM water consumption, a positive lag is seen in the cross-correlation plot. On the positive lag, there is a significant CCF value on the variables by plots.

The lag which has almost significant CCF is still used as a reference for determining the transfer function order to maintain the transfer function modeling in this study. For example, in the plot of the cross-correlation between the series $\alpha_{2,t}$ and $\beta_{2,t}$ the CCF value that approaches the first significant significance is found in lag 9. While in the plot of the cross-correlation between the series $\alpha_{4,t}$ and $\beta_{4,t}$ the CCF value that approaches the first significance is in lag 3. If the cross-correlation value between $\alpha_{1,t}$ and $\beta_{i,t}$ has been obtained, the results of CCF can be used to calculate the weight value of the impulse response.

The order in the transfer function model can be determined by referring to the CCF pattern and the weight of the formed impulse response. Order *b* is marked by lag when CCF is firstly significant. Order *s* is determined based on the number of lags that have survived significantly for the second and third time since the first significant lag. While the order *r* is determined based on the pattern of the CCF plot. For instance, the CCF plot between the variable $\alpha_{T,t}$ and $\beta_{T,t}$ the first CCF value was almost significantly found at lag 0, but the CCF value that was ensured significant afterward was lag 1. So, the order value *b* was tried between 0 and 1. After lag 0, there were 2 lags that follow significantly. Then order *s* is equal to 2 or 1 if it assumes order *b* is equal to 1. The pattern of cross-correlation shows a sinus wave-like pattern, so that order *r* is equal to 2. However, order *r* can be tested between 0.1 and 2 [17].

4.3.4. Noise Series Calculation and Modeling

Noise series calculation will be useful in estimating the parameters of the transfer function model. In this model, the output series described and influenced by the input series and other inputs combined in a group called noise. The following are the results of the noise series calculation.

a) Noise series for the transfer function model of Malang City area

$$y^*_{T,t} = \upsilon_{T,t}(B)x^*_{T,t} + n_{T,t}$$

$$n_{T,t} = y^*_{T,t} - \upsilon_{T,0}x^*_{T,t} - \ldots - \upsilon_{T,n}x^*_{T,t-n}$$

$$n_{T,t} = y^*_{T,t} - 31.77x^*_{T,t} - 41.22x^*_{T,t-1} - 26.23x^*_{T,t-n} \quad (4.1)$$

Equation (4.1) can be calculated by firstly determining the values of $n_{T,1}$, $n_{T,2}$, $n_{T,15}$ equal to zero and then the values of $n_{T,166}$, $n_{T,17}$, ..., $n_{T,88}$ can be calculated.

b) Noise series for the Blimbing district transfer function model

$$y^*_{1,t} = \upsilon_{1,t}(B)x^*_{1,t} + n_{1,t}$$
$$n_{1,t} = y^*_{1,t} - \upsilon_{1,0}x^*_{1,t} - \ldots - \upsilon_{1,n}x^*_{1,t-n}$$
$$n_{1,t} = y^*_{1,t} - 967.92x^*_{1,t} - 2200.86x^*_{1,t-1} - \ldots + 668.32x^*_{1,t-15} \quad (4.2)$$

The noise series in line with equation (4.2) can be calculated by assuming the values of $n_{1,1}$, $n_{1,2}$, $n_{1,15}$ equal to zero and then the values of $n_{1,16}$, $n_{1,17}$, ..., $n_{1,88}$ can be calculated.

c) Noise series for the Kedungkandang district transfer function model

$$y^*_{2,t} = \upsilon_{2,t}(B)x^*_{2,t} + n_{2,t}$$
$$n_{2,t} = y^*_{2,t} - \upsilon_{2,0}x^*_{2,t} - \ldots - \upsilon_{2,n}x^*_{2,t-n}$$
$$n_{2,t} = y^*_{2,t} - 10.013x^*_{2,t} - 5.192x^*_{2,t-1} - \ldots + 3.09x^*_{2,t-15} \quad (4.3)$$

Noise series in equation (4.3) can be calculated by assuming the values of $n_{2,1}$, $n_{2,2}$, $n_{2,15}$ equal to zero and then the values of $n_{2,16}$, $n_{2,17}$, ..., $n_{2,88}$ can be calculated.

d) Noise series for the Klojen district transfer function model

$$y^*_{3,t} = \upsilon_{3,t}(B)x^*_{3,t} + n_{3,t}$$
$$n_{3,t} = y^*_{3,t} - \upsilon_{3,0}x^*_{3,t} - \ldots - \upsilon_{3,n}x^*_{3,t-n}$$
$$n_{3,t} = y^*_{3,t} + 4.23x10^{-17}\,x^*_{3,t} - 1.69x10^{-17}\,x^*_{3,t-1} - \ldots + 4.23x10^{-17}\,x^*_{3,t-20} \quad (4.4)$$

Equation (4.4) can be calculated by first determining the values of $n_{3,1}$, $n_{3,2}$, $n_{2,20}$ equal to zero and then the values of $n_{3,21}$, $n_{3,22}$, $n_{2,89}$ can be calculated.

e) Noise series for Lowokwaru district transfer function model

$$y^*_{4,t} = \upsilon_{4,t}(B)x^*_{4,t} + n_{4,t}$$
$$n_{4,t} = y^*_{4,t} - \upsilon_{4,0}x^*_{4,t} - \ldots - \upsilon_{4,n}x^*_{4,t-n}$$
$$n_{4,t} = y^*_{4,t} + 0.041x^*_{4,t} + 0.047x^*_{4,t-1} + \ldots - 0.041x^*_{4,t-15} \quad (4.5)$$

The noise series based on equation (4.5) can be calculated by assuming the values of $n_{4,1}$, $n_{4,2}$, $n_{4,15}$ equal to zero and then the values of $n_{4,16}$, $n_{4,17}$, $n_{4,77}$ can be calculated.

f) Noise series for Sukun district transfer function model

$$y^*_{5,t} = \upsilon_{5,t}(B)x^*_{5,t} + n_{5,t}$$
$$n_{5,t} = y^*_{5,t} - \upsilon_{5,0}x^*_{5,t} - \ldots - \upsilon_{5,n}x^*_{5,t-n}$$
$$n_{5,t} = y^*_{5,t} - 3.33x^*_{5,t} - 0.54x^*_{5,t-1} - \ldots + 3.38x^*_{5,t-15} \quad (4.6)$$

Noise series in equation (4.6) can be calculated by assuming the values of $n_{5,1}$, ..., $n_{5,15}$ equal to zero and then the values of $n_{5,16}$, ..., $n_{5,89}$ can be calculated.

Noise series obtained from the calculation results do not need to be stationary, as these series are obtained from series that have been stationary. The autocorrelation value in the ACF plot of the noise series proves that the series is stationary to the mean. The order d of the noise series is equal to 0. The determination of order q is based on the lag that has a significant autocorrelation on the ACF plot. Whereas the order q is obtained based on lag when partial autocorrelation is significant. Estimating the parameters to select the right model can be done by experimenting with a combination of out of bounds out of bounded lag values [9].

The results of parameter estimation and significance testing show that the model with all the significant parameters at the 5% level is ARIMA (3,0,0) for the $n_{T,t}$, ARIMA (0,0,1) series for $n_{1,t}$, ARIMA series ([2],0,0) for series $n_{3,t}$, SARIMA (3[7],0,0) $(0,0,0)^{12}$ for series $n_{4,t}$ and ARIMA (0,0, [2]) for series $n_{5,t}$. While the ARIMA model (3[6],0.0) for series $n_{2,t}$ has significant parameters at the 10% level. The ARIMA model (3[6],0.0) was chosen because of the seven tentative ARIMA models for series $n_{2,t}$, the model (3[6],0,0) fulfills the white noise assumption and has the smallest AIC among the models whose parameters are significant.

4.3.5. Estimating the Transfer Function Model Parameters

In estimating the transfer function parameters, it is necessary to first determine the candidate values (b, s, r) and the model of the noise series. This order (b, s, r) value has previously been estimated in Section 4.3.3 while the best noise series has been explained in Section 4.3.4. Estimation of the transfer function model parameters is carried out on a number of possible tentative models. The following, Table 4.2, are the results of parameter estimation and testing of significance of the best transfer function for each output series.

The transfer function model (b,s,r) (p_n,q_n) which has all the significant parameters is the transfer function model with constant of (1,1,0)

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

(3,0) for the series $y^*_{T,t}$, the transfer function model with constant of (13,0,0)(0,1) for series $y^*_{1,t}$, transfer functions (2,0,1)([2],0) for variables $y^*_{3,t}$, and transfer functions (3,0,1)(0,[2]) for $y^*_{5,t}$. While the transfer function model (9,0,0)(3[6],0) to $y^*_{3,t}$ have a significant parameter $\hat{\phi}_6$ at the level of 10%. This is due to modeling the series $n_{2,t}$, requires parameters $\hat{\phi}_6$ to be entered into the model as described in Section 4.3.4. Seasonal transfer function models (*b*,*s*,*r*) (S,R)(*p*,*q*)(P,Q) (6,0,1)(0.0) (3 [7],0)(0.0)^{12} for a series $y^*_{4,t}$ with parameters $\hat{\omega}_0$ are not significant. The seasonal transfer function model (2,0,1)(0,0)(3[7],0)(0,0)$_{12}$ is still used because of the three tentative models that were tried, the model has estimates which are close to significant. In addition, order *b* cannot be significant because in the CCF plot, there is no truly significant lag. This model is still used for hierarchical forecasting in this study as the method requires all variables to be modeled with a consistent model. In determining whether the six models in Table 4.2 can be used for forecasting, the model must go through a stage of diagnostic testing.

Tabel 4.2 Estimation and significance test of the parameters

Var.; Transfer Function Model **(b, s, r)** **(p_n,q_n)**	Para-meter	p-*value*	AIC
$y^*_{1,t}$; (1,1,0) (3,0) *with const.*	$\hat{\mu}$ = 4717224	0.0045	3221.70
	$\hat{\phi}_1$ = -0.3921	0.0004	
	$\hat{\phi}_2$ = -0.4010	0.0002	
	$\hat{\phi}_3$ = -0.2908	0.0065	
	$\hat{\omega}_0$ = 46.4846	0.0348	
	$\hat{\omega}_1$ = -84.2200	0.0001	
$y^*_{1,t}$; (13,0,0) (0,1) *with constant*	$\hat{\mu}$ = 7675311	<0.0001	2901.52
	$\hat{\theta}_1$ = 0.7575	<0.0001	
	$\hat{\omega}_0$ = 2959.7	0.0141	
$y^*_{2,t}$; (9,0,0) (3[6],0)	$\hat{\phi}_1$ = -0.5322	<0.0001	1147.76
	$\hat{\phi}_2$ = -0.3263	0.0057	
	$\hat{\phi}_3$ = -0.4309	<0.0001	
	$\hat{\phi}_6$ = -0.1891	0.0681	
	$\hat{\omega}_0$ = 9.91387	0.0023	
$y^*_{3,t}$; (2,0,1) ([2],0)	$\hat{\phi}_2$ = -0.2461	0.0211	2183.83
	$\hat{\omega}_0$ = 2.4E-17	0.0181	
	$\hat{\delta}_1$ = -0.8049	<0.0001	
$y^*_{4,t}$; (6,0,0) (0,0) (3[7],0) (0,0)$_{12}$	$\hat{\phi}_1$ = -0.6598	<0.0001	2303.80
	$\hat{\phi}_2$ = -0.486	<0.0001	
	$\hat{\phi}_3$ = -0.3472	0.0006	
	$\hat{\phi}_7$ = 0.1985	0.0201	
	$\hat{\omega}_0$ = 0.0758	0.319	
$y^*_{5,t}$; (3,0,1) (0,[2]) *with constant*	$\hat{\mu}$ = 572882.7	0.0117	2678.43
	$\hat{\theta}_2$ = 0.3463	<0.0001	
	$\hat{\omega}_0$ = -5.9948	<0.0001	
	$\hat{\delta}_1$ = -0.7739	<0.0001	

4.3.6. Diagnostic Test of Transfer Function Model

The diagnostic test is done after testing the significance of the parameters of the transfer function model before finally selecting the best transfer function model. This diagnostic test in Table 4.3 is only presented on the six best transfer function models for Malang city and its 5 districts to ensure that it meets the assumptions of the transfer function modeling. The first diagnostic test of the model is checking the residual autocorrelation of the model, i.e. $a_{i,t}$. The statistical test to ensure that residuals are white noise is the Ljung Box test. Based on results in Table 4.2, it can be concluded that the best transfer function model for each output level one variable is already white noise because the Ljung Box test p-value gives a decision to accept the null hypothesis. This means the model has been able to capture the factors that must be included in the model. The second test is to ensure that the value of the cross-correlation (CCF) between residuals and series $\alpha_{i,t}$ is insignificant in all lags. The significance test shows that the CCF between residuals and series in the five transfer function models is not significant until the 24th lag. This shows that there is no correlation between residuals and input lines. These have been whitened based on time lag and the model can be said to be feasible.

Table 4.3 The Ljung Box Test on the Residual Transfer Function Model

Vari-able	Transfer Function Model $(b,s,r)(p_n,q_n)$	*To Lag*	*p-value*	Note
$y^*_{T,t}$	(1,1,0)(3,0) *with constant*	6	0.2666	Fit
		12	0.207	Fit
		18	0.3817	Fit
		24	0.1911	Fit
$y^*_{1,t}$	(13,0,0) (0,1) *with constant*	6	0.6688	Fit
		12	0.2189	Fit
		18	0.507	Fit
		24	0.6579	Fit
$y^*_{2,t}$	(9,0,0) (3[6],0)	6	0.1792	Fit
		12	0.2587	Fit
		18	0.2388	Fit
		24	0.166	Fit
$y^*_{3,t}$	(2,0,1) ([2],0)	6	0.0808	Fit
		12	0.3313	Fit
		18	0.2142	Fit
		24	0.2241	Fit
$y^*_{4,t}$	(6,0,0)(0,0) $(3[7],0)(0,0)_{12}$	6	0.7614	Fit
		12	0.2266	Fit
		18	0.2565	Fit
		24	0.2571	Fit
$y^*_{5,t}$	(3,0,1) (0,[2]) *with constant*	6	0.7834	Fit
		12	0.7269	Fit
		18	0.6313	Fit
		24	0.7007	Fit

4.3.7. Interpretation of the Best Transfer Function Model

The results of the modeling process of transfer function on the PDAM water consumption in Malang city and its districts of Blimbing, Kedungkandang, Klojen, Lowokwaru, and Sukun areas are shown as the following:

a) Transfer Function Model (1,0,0)(3,0) with Constant for Malang City

$$y^*_{T,t} = \frac{\omega(B)}{\delta(B)} x^*_{T,t-b} + \frac{\theta(B)}{\phi(B)} a_{T,t-b}$$

$$\hat{y}^*_{T,t} = \hat{\mu} + (\hat{\omega}_0 - \hat{\omega}_1) x^*_{T,t-b} + \frac{1}{(1-\hat{\phi}_1 B^1 - \hat{\phi}_2 B^2)^{a_{T,t}}}$$

$$= 4717224 + 46.48 x^*_{T,t-1} - 8.75 x^*_{T,t-2} - 121.96 x^*_{T,t-3} + 84.22 x^*_{T,t-4} + 0.61 y^*_{T,t-1} - 0.009 y^*_{T,t-2} + 0.11 y^*_{T,t-3} + 0.009 y^*_{T,t-4} + a_{T,t} \tag{4.7}$$

Based on the model (4.7), PDAM water consumption in Malang City this month was affected by water consumption in the previous 1 to 4 months, the number of customers in the previous 1 to 4 months and factors that are not explained in the model.

b) Transfer Function Model (13,0,0)(0,1) with Constant for the Blimbing district

$$y^*_{1,t} = \frac{\omega(B)}{\delta(B)} x^*_{1,t-b} + \frac{\theta(B)}{\phi(B)} a_{1,t-b}$$

$$\hat{y}^*_{1,t} = \hat{\mu} + \hat{\omega}_0 x^*_{1,t-b} + (1-\hat{\phi}_1 B) a_{1t}$$

$$y^*_{1,t} = 7675311 + 2959.7 x^*_{1,t-13} - 5919.4 x^*_{1,t-14} + 2959.7 x^*_{1,t-15} - y^*_{1,t-1} + a_{1,t} - 0.75 a_{1,t-1} \tag{4.8}$$

Model (4.8) can be interpreted as that PDAM water consumption in Blimbing district this month is influenced by the number of customers in the previous 13, 14, and 15 months, water consumption in the previous month, and other factors not included in the model.

c) Transfer Function Model (9,0,0)([1,2,3,6], 0) for the Kedungkandang district

$$y^*_{2,t} = \frac{\omega(B)}{\delta(B)} x^*_{2,t-b} + \frac{\theta(B)}{\phi(B)} a_{2,t-b}$$

$$\hat{y}^*_{2,t} = \hat{\omega}_0 x^*_{2,t-b} + \frac{1}{(1-\hat{\phi}_1 B^1 - \hat{\phi}_2 B^2 - \hat{\phi}_3 B^3 - \hat{\phi}_6 B^6)} a_{2,t}$$

$$\hat{y}^*_{2,t} = 0.468 y^*_{2,t-1} + 0.206 y^*_{2,t-2} - 0.105 y^*_{2,t-3} + 0.431 y^*_{2,t-4} - 0.189 y^*_{2,t-6} + 0.189 y^*_{2,t-7} + 9.914 x^*_{2,t-9} - 9.914 x^*_{2,t-10} + a_{2,t} \tag{4.9}$$

This can be interpreted that PDAM water consumption in Kedungkandang district this month was affected by water consumption in 1,2,3,6 and 7 months before, the number of customers in the previous 9 and 10 months, and unexplained factors in the model.

d) Transfer Function Model (2,0,1)([2],0) for the Klojen district

$$y^*_{3,t} = \frac{\omega(B)}{\delta(B)} x^*_{3,t-b} + \frac{\theta(B)}{\phi(B)} a_{3,t-b}$$

$$\hat{y}^*_{3,t} = \frac{\hat{\omega}_0}{(1-\hat{\delta}_1 B)} x^*_{3,t-2} + \frac{\hat{\omega}_0}{(1-\hat{\phi}_2 B^2)}$$

$$\hat{y}^*_{3,t} = 0.195 y^*_{3,t-1} + 0.559 y^*_{3,t-2} + 1.051 y^*_{3,t-3} + 0.198 y^*_{3,t-4} + 2.35x10^{-17} x^*_{3,t-2} - 2.35x10^{-17} x^*_{3,t-1} + a_{3,t} \tag{4.10}$$

Transfer function model (4.10) can be interpreted that PDAM water consumption in Klojen district this month is affected by water consumption in the previous 1 to 4

Journal of Theoretical and Applied Information Technology
15th October 2020. Vol.98. No 19

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

months, the number of customers in the previous 2 and 3 months and factors not explained in the model.

e) Seasonal Transfer Function Model (6,0,0)(0,0)([1,2,3,7],0)(0,0)^{12} for the Lowokwaru district

$$y_{4,t}^{*}=\frac{\omega_i(B)\omega_{s,i}(B^s)}{\delta_i(B)\delta_{s,i}(B^s)}x_{4,t-b}^{*}+\frac{\theta_q(B)\Theta_{s,p}(B^s)}{\phi_p(B)\Theta_{s,q}(B^s)}a_{4,t}$$

$$\hat{y}_{4,t}^{*}=\hat{\omega}_0x_{4,t-6}^{*}+\frac{1}{(1-\hat{\phi}_1B^1-\hat{\phi}_2B^2-\hat{\phi}_3B^3-\hat{\phi}_7B^7)}a_{4,t}$$

$$\begin{aligned}\hat{y}_{4,t}^{*}&=0.34y_{4,t-1}^{*}+0.174y_{4,t-2}^{*}+0.139y_{3,t-3}^{*}+\\&\quad 0.35y_{4,t-4}^{*}+0.198y_{4,t-7}^{*}+0.198y_{4,t-8}^{*}+y_{4,t-12}^{*}\\&\quad -1.7y_{4,t-13}^{*}+0.49y_{4,t-14}^{*}-0.14y_{4,t-15}^{*}-0.35y_{4,t-16}^{*}\\&\quad -0.198y_{4,t-19}^{*}+0.198y_{4,t-20}^{*}+0.076x_{4,t-6}^{*}\\&\quad -0.152x_{4,t-7}^{*}-0.076x_{4,t-8}^{*}+a_{4,t}\qquad(4.11)\end{aligned}$$

Model (4.11) can be interpreted that the consumption of PDAM water in Lowokwaru district this month is influenced by the number of customers in the previous 6, 7 and 8 months, water consumption in 1, 2, 3, 4, 7, 8, 12, 13, 14 , 15, 16,19, and 20 months earlier, as well as other factors not included in the model.

f) Transfer Function Model (3,0,1)(0,[2]) with Constant for Sukun district

$$y_{5,t}^{*}=\frac{\omega(B)}{\delta(B)}x_{5,t-b}^{*}+\frac{\theta(B)}{\phi(B)}a_{5,t}$$

$$y_{5,t}^{*}=\mu+\frac{\hat{\omega}_0}{(1-\hat{\delta}_1B)}x_{5,t-2}^{*}+(1-\theta_2B^2)a_{5,t}$$

$$\begin{aligned}y_{5,t}^{*}&=572882.7+0.226y_{5,t-1}^{*}+0.774y_{5,t-2}^{*}\\&\quad -5.99x_{5,t-8}^{*}+5.99x_{5,t-4}^{*}+a_{5,t}-0.346a_{5,t-2}\end{aligned}\qquad(4.12)$$

Equation (4.12) can be interpreted that PDAM water consumption in Sukun district this month is influenced by water consumption in the previous 1 to 2 months, the number of customers in the previous 2 to 4 months, as well as factors not explained in the model.

4.4. Prediction Results of All Output Variables Based on the Transfer Function Model

After obtaining the best function model in Malang city and its district in the hierarchical structure, then the prediction is possible using models in Section 4.3.7. The model can be used to predict the value of output variables based on in-sample or out-sample periods. The plot of the results of predictions based on the model is useful for seeing the goodness of the model descriptively. The following plots display the results of prediction of output variables in all regions based on the in-sample period. It can be seen that the predicted results of PDAM water consumption based on the transfer function model are quite close to the value of the data in the sample. This is reflected by the shape of the plot and their similar trends, particularly in the prediction results in Lowokwaru district. While prediction plots in Malang city, Kedungkandang, Klojen and Sukun districts have mostly followed fluctuations of in-sample data and data patterns. Unlike the other areas, the prediction plot in Blimbing district does not follow the fluctuations of in-sample data but still follows the data trends by time.

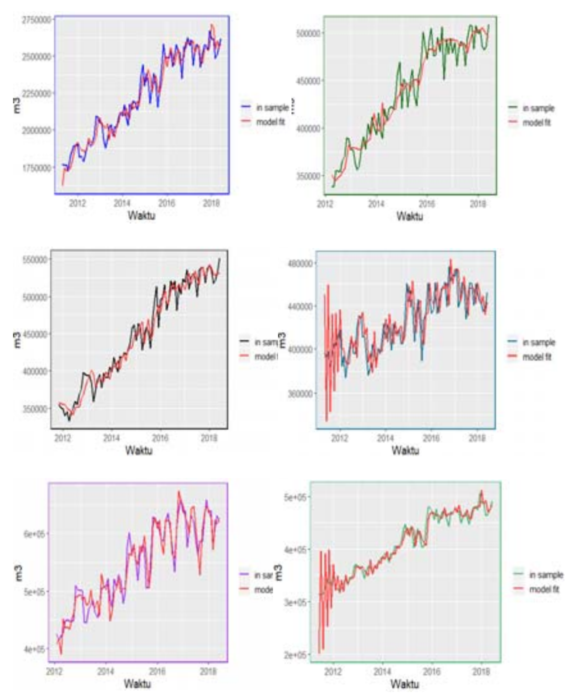

Figure 4.2 Plot of Results of Prediction Output Variables in Sample

Figure 4.2 is a prediction plot based on the best transfer function model of out-sample period. The value of the out-sample data in Malang City decreased in July and August then increased continuously until December. This is consistent with most districts in Malang City where the first 2 months of water consumption has fallen then 4 months later consumption had risen quite sharply. The results of the initial forecast output variable show that forecasting for the next six months mostly tends to increase by month. However, the results of this forecasting did not follow the fluctuations in the out-sample data, especially in

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

November and December. Only the results of the initial forecast output variable in the Lowokwaru district followed the fluctuation of the sample data well. Initial forecasting results do not produce a coherent forecasting because the number of disaggregated series results is not the same as the aggregated series value.

After the prediction value is obtained based on the transfer function model, then the prediction results in-sample and out-sample can be used to calculate the accuracy value of the model, sMAPE. The sMAPE value is calculated for the in-sample and out-sample periods to see the consistency of the model's ability to predict output variables. sMAPE on in-sample and out-sample periods are represented by the its average values. It is the sMAPE calculated from the average between sMAPE of in-sample and out-sample with weights of 93.75% and 6.25%, respectively. These weights are in accordance with the partition of periods in-sample and out-sample.

Table 4.3 sMAPE Value Prediction Results of the Best Transfer Function Models

Region	Transfer function model (b,s,r) (p_n,q_n)	sMAPE(%)	
		In sam-ple	*Out sam-ple*
Malang city	(1,1,0)(3,0) *with Constant*	2.541	3.1820
Blim-bing	(13,0,0)(0,1) *with constant*	2.896	3.5820
Kedung-kandang	(9,0,0)(3[6],0)	2.534	3.7470
Klojen	(2,0,1)([2],0)	3.799	2.7570
Lowok-waru	(6,0,0)(0,0) $(3[7],0)(0,0)_{12}$	3.037	0.7080
Sukun	(3,0,1)(0,[2]) *with constant*	4.383	3.6500

4.5. Method of Hierarchical Time Series Forecasting

The results of forecasting revisions with the time series hierarchical method using three approaches are explained below. They are bottom-up, top-down based on historical proportions, and top-down based on the proportion of forecasting with the transfer function model [4].

4.5.1 Bottom-Up Method

In this approach, the forecasting results obtained at level 0, forecasting PDAM water consumption in Malang city, by adding up the forecasting results at level 1. The mathematical form of bottom-up forecasting for level 0 is as follows

$$\tilde{Y}_{T,h} = \sum_{i=1}^{5} \hat{Y}_{i,h}$$

$$\tilde{Y}_{T,1} = \hat{Y}_{1,1} + \hat{Y}_{2,1} + \ldots + \hat{Y}_{5,1}$$

$$\tilde{Y}_{T,2} = 487075 + 535321.7 + \ldots + 490374)$$

$$\vdots$$

$$\tilde{Y}_{T,6} = 507458.3 + 543909 + \ldots + 500293 . \qquad (4.13)$$

$\tilde{Y}_{T,h}$ is forecasting PDAM water consumption at level 0 (Malang City) and i is notation for each district in Malang City. In the bottom-up method, forecasting results at level 1 are equivalent to the value of initial forecasting based on the formed transfer function model.

4.5.2 Top-Down Method with Historical Proportion

This second method is useful for breaking down initial forecasting at level 0 to get forecasting at level 1. The assumption in this calculation is the proportion in each month for each level 1 variable is the same. There are two kinds of equations for finding historical data proportions. The first equation is Top-Down Historical Proportions 1 (TDHP 1). For instance, the following values for the proportion of series one for Blimbing district will be calculated:

$$p_{1i} = \frac{\sum_{t=1}^{n} \frac{y_{i,t}}{y_{T,t}}}{n}$$

$$p_{11} = \frac{\sum_{t=1}^{90} \frac{y_{i,t}}{y_{T,t}}}{90}$$

$$= \frac{1}{90}\left[\frac{y_{1,1}}{y_{T,1}} + \frac{y_{1,2}}{y_{T,2}} + \ldots + \frac{y_{1,90}}{y_{T,90}}\right]$$

$$= \frac{1}{90}\left[\frac{336338}{1793694} + \frac{301203}{1643813} + \ldots + \frac{508825}{2621152}\right]$$

$$= 0.190616 \qquad (4.14)$$

each p_{1i} is a historical proportion average of series at level 1 ($Y_{i,t}$) at time t=1,2,…,n that useful in breaking aggregated total forecasting ($Y_{T,t}$).

Whereas the second equation for finding historical proportions is TDHP 2. For instance, we are interested in calculating the proportions for a variable $Y_{i,t}$, then mathematically it can be written as:

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

$$p_{2i} = \sum_{t=1}^{n} \frac{y_{i,t}}{n} / \sum_{t=1}^{n} \frac{y_{T,t}}{n}$$

$$p_{21} = \sum_{t=1}^{90} \frac{y_{i,t}}{90} / \sum_{t=1}^{90} \frac{y_{T,t}}{90}$$

$$= 0.190907 \qquad (4.15)$$

each $p_{2,i}$ is a historical proportion average of series at level 1 ($Y_{i,t}$) at time t=1,2,…,n that useful in breaking aggregated total forecasting ($Y_{T,t}$).

After obtaining the proportion value for each series, the forecast revision results can be calculated using the historical proportion top-down method. For instance, forecasting calculation for variable $Y_{i,t}$ one step ahead based on the proportion of TDHP 1 as follows:

$$\tilde{y}_{1,h} = p_{1,i} \ x \ \tilde{y}_{T,h}$$

$$\tilde{y}_{1,1} = p_{1,1} \ x \ \hat{y}_{T,1}$$

$$= 0.19062 \text{ x } 2608479$$

$$= 497216.7 \qquad (4.16)$$

where $\tilde{y}_{T,h}$ is forecasting h step forward for the variable of water consumption in Blimbing district used bottom-up method whereas $\hat{y}_{T,h}$ is initial forecast h step forward for the variable of water consumption in Malang. City.

4.5.3 Top-Down Method with Forecasting Proportion

Disaggregation using forecasting proportion includes the assumption that the proportion in each month for each level 1 variable is not similar. This equation can be used to calculate the value of the proportion of forecasting variable $Y_{i,t}$ where one step ahead can be written as follows:

$$p_{3_{i,h}} = \prod_{i=0}^{K-1} \frac{\hat{Y}_{i,h}^{l}}{\hat{S}_{i,h}^{l+1}}$$

$$p_{3_{1,1}} = \prod_{i=0}^{0} \frac{\hat{Y}_{1,h}^{0}}{\hat{S}_{1,h}^{1}}$$

$$= \frac{497075.2}{2568748}$$

$$= 0.19351 \qquad (4.17)$$

with $p_{3_{i,h}}$ is forecasting proportion to calculate variable forecasted value $Y_{1,t}, \hat{Y}_{1,h}^{0}$ is the result of initial forecasting for $Y_{1,t}$ at h next month and $\hat{S}_{1,h}^{1}$ is sum of initial forecasting from each i at h the next month.

The obtained proportion is used to calculate the variable forecasting $Y_{1,t}$ one month ahead with calculation as follows:

$$\tilde{y}_{1,h} = p_{3_{i,h}} \ x \ \hat{y}_{T,h}$$

$$\tilde{y}_{1,1} = p_{3_{1,1}} \ x \ \hat{y}_{T,h}$$

$$= 0.19351 \text{ x } 2608479$$

$$= 497216.7 \qquad (4.18)$$

with $\tilde{y}_{1,h}$ is forecasting h the next month for water consumption variable in Blimbing district based on top-down method, whereas $\hat{y}_{T,h}$ is initial forecasting h next step for water consumption variable at Malang city.

From our results above, we realize that this result can still be broadened. Some factors influencing the water consumption should be accommodated [16]. These factors include for instance accommodating other factors influencing the water consumption, demographic profiles, industrial activities in the city, and so forth. We would consider the factors in the future research, however for the moment the factor we include has been enough as we stressed the consumption prediction on leveling (two levels in our cases). The leveling is of important in predicting the consumption based on specific region and volumes.

5. CONCLUDING REMARKS

Modeling PDAM water consumption in Malang city using the hierarchical forecasting method with the transfer function model produces six best transfer function models from each region (Malang city and its 5 districts). The best transfer function model for Malang City is the transfer function model (b,s,r)(p_n,q_n)(1,1,0)(3,0), with transfer function of (13,0,0)(0,1) for Blimbing district, transfer function of (9,0,0)([1,2,3,6],0) for Kedungkandang district, transfer function of (2,0,1)([2],0) for Klojen district, seasonal transfer function of (6,0,0)(0.0)([1,2,3,7],0)(0,0) 12 for Lowokwaru district, and the transfer function model of (2,0,1)(0,[2]) for Sukun district. Based on the revised forecasting results using the hierarchical forecasting method, the best approach to the hierarchical forecasting method for the time-series data of the PDAM water consumption hierarchy in Malang city is the bottom-up method.

REFERENCES:

[1] Weiss, C. 2018. *Essays in Hierarchical Time Series Forecasting and Forecast Combination*. Unpublished Desertation. Cambridge University.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

[2] Anthanasopoulus, G., Ahmed, R.A., and R.J. Hyndman. 2007. Hierarchical forecasts for Australian Domestic Tourism. *International Journal of Forecasting*, **25**(1), 146–166.

[3] Hyndman, R.J., Ahmed, A.G., Athanasopoulos, R., and Shang, H. L. 2011. Optimal Combination Forecasts for Hierarchical Time Series. *Computational Statistics and Data Analysis,* **55**:2579-2589.

[4] Makhya, D.A., Ulama, B.S., and Suhartono. 2016. Hierarchical time series bottom-up approach for forecast the export value in Central Java. *Journal of Physics: Conference Series*, **893**: 012033.

[5] Hyndman, R.J and George, A. 2018. *Forecasting: Priciples and Practice, 2nd Edition*. Australia: Otexts.

[6] Brockwell, P.J. and Davis, R.A. 1996. *Introduction to Time Series and Forecasting, 2nd Edition*. New York: Springer.

[7] Hanke, J.E. and Reitsch, A.G. 1992. *Business Forecasting, 4th Edition*. Allyn and Bacon. USA.

[8] Ali, N.R. and Ku-Mahamud, K.R. 2018. High-order RTV-Fuzzy Time Series Forecasting Model Based on Trend Variation. *Journal of Theoretical and Applied Information Technology*, **96** (21), 7151-7163.

[9] Makridakis, S., Wheelwright, S.C., and McGee, V.E. 1999. *Forecasting: Methods and Aplications 2nd Edition*. New York: John Wiley and Sons, Inc.

[10] Cryer, J.D. and Chan, K. 2008. *Time Series Analysis with Applications in R*. New York: Springer.

[11] Bowerman, B.L. and O'Cornell, R.T. 1993. *Forecasting and Time Series: An Applied Approach, 3rd Edition*. California: Duxbury Press.

[12] Othman, Z.A., Ismail, N., Hamdan, A.R., and Sammour, M.A. 2016. Klang Vally Rainfall Forecasting Model Using Time Series Data Mining Technique. *Journal of Theoretical and Applied Information Technology*, **92** (2), 372-379.

[13] Abraham, B. 1985. Seasonal Time Series and Transfer Function Modeling. *Journal of Business & Economic Statistics*, **3**(4), 356-361, DOI: 10.1080/07350015.1985.10509469

[14] Gato, S., Jayasuriya, N., Roberts, P., and Hadgraft, R. 2005. A Simple Time Series Approach to Modeling Urban Water Demand. *Australian Journal of Water Resources,* DOI: 10.1080/13241583.2005.11465252.

[15] Chen, J. and Bocelli, D.I. 2014. Demand forecasting for water distribution systems. *Procedia Engineering* 70: 339-342. 12th International Conference on Computing and Control for the Water Industry, CCWI2013.

[16] Mamade, A., Loureiro, D., Covas, D., Coelho, S.T., and Amado, C. 2014. Spatial and temporal forecasting of water consumption at the DMA level using extensive measurements. *Procedia Engineering* 70: 1063-1073. 12th International Conference on Computing and Control for the Water Industry, CCWI2013.

[17] Wei, W. 2006. *Time Series Analysis: Univariate and Multivariate Methods (2nd ed.)*. USA: Pearson Education, Inc.

[18] Indrawan, M.F. 2019. Characteristic of Water Consumption of PDAM Malang City. *Interview Document*, June 2019 at PDAM Malang Office.