PERBANDINGAN MODEL ARIMA DAN MODEL REGRESI DENGAN RESIDUAL ARIMA DALAM MENERANGKAN PERILAKU PELANGGAN LISTRIK DI KOTA PALOPO

Alia Lestari
Fakultas Teknik Universitas Andi Djemma

Abstrak

Persediaan sumber daya energi di bumi ini semakin hari semakin menipis. Upaya penghematan energi merupakan salah satu usaha yang perlu dilakukan, termasuk melakukan penghematan penggunaan listrik. Oleh karena itu diperlukan suatu metode yang tepat untuk mengkaji perilaku pelanggan listrik. Dalam makalah ini akan dibandingkan model *Autoregressive Integrated Moving Average* (ARIMA) dan model Regresi dengan residual ARIMA dalam menerangkan perilaku pelanggan listrik di Kota Palopo. Hasil analisis menunjukkan bahwa model ARIMA lebih baik dari model Regresi dengan residual ARIMA.

Kata kunci : ARIMA, Regresi, Pelanggan Listrik

PENDAHULUAN

Kampanye tentang hemat energi dewasa ini sedang digalakkan pemerintah. Hal ini mengingat ketersediaan sumber daya energi di bumi yang semakin hari semakin menipis. Berbagai penelitian dilakukan untuk mendukung program penghematan energi tersebut. Salah satu sumber energi yang sangat penting adalah listrik. Sebagaimana kita ketahui sebagian besar alat-alat yang digunakan oleh masyarakat menggunakan energi listrik. Oleh karena itu perlu dilakukan suatu penelitian tentang bagaimana perilaku masyarakat pelanggan listrik. Hasil dari penelitian ini diharapkan dapat memberi gambaran sehingga menjadi masukan bagi pemerintah, khususnya perusahaan listrik negara untuk dapat mengoptimalkan penggunaan listrik.

Penelitian terdahulu terhadap penggunaan energi listrik telah dilakukan oleh H.M. Al-Hamadi dan S.A. Soliman (2005) dengan pendekatan regresi. Pada makalah ini akan dibandingkan dua metode untuk menganalisis perilaku pelanggan listrik yaitu model ARIMA (*Autoregressive Integrated Moving Average*) dan model regresi dengan residual

ARIMA. Data yang digunakan dalam penelitian ini adalah jumlah pemakaian listrik selama bulan Juni tahun 2006 di Kota Palopo.

Pada bagian pertama akan digunakan model ARIMA, interpretasinya disajikan pada bagian kedua. Bagian ketiga dijabarkan penggunaan model Regresi Linier Sederhana, kemudian interpretasi model disajikan pada bagian keempat. Bagian berikutnya adalah membandingkan kedua metode yang telah digunakan diatas sehingga didapatkan model yang paling baik dalam menjelaskan data.

MODEL ARIMA

Dalam model deret berkala atau *time series*, pendugaan masa depan dilakukan berdasarkan data yang tersedia pada masa lalu. Tujuan metode peramalan deret berkala adalah menemukan pola dalam *time series* yang nantinya digunakan dalam penentuan data hasil ramalan di masa-masa mendatang (*forecast*).

Salah satu asumsi dalam tahap pembentukan model *time series* adalah stasioneritas. Data deret waktu yang stasioner dapat dijelaskan dimana relatif tidak terjadi kenaikan ataupun penurunan nilai secara tajam pada data deret waktu atau dengan kata lain, data berfluktuasi disekitar nilai rata-rata yang konstan. Kondisi stasioneritas dibedakan menjadi dua, yakni stasioner dalam mean dan stasioner dalam varians

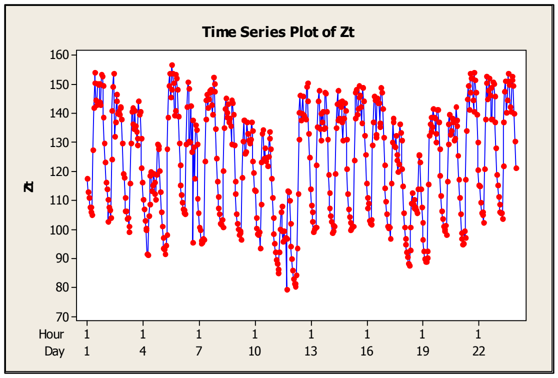
Model ARIMA merupakan model yang umum digunakan dalam analisis *time series*. Secara matematis model ARIMA dituliskan dalam bentuk:

$$\phi_p(B)(1-B)^d \dot{Z}_t = \theta_q(B)a_t \tag{1}$$

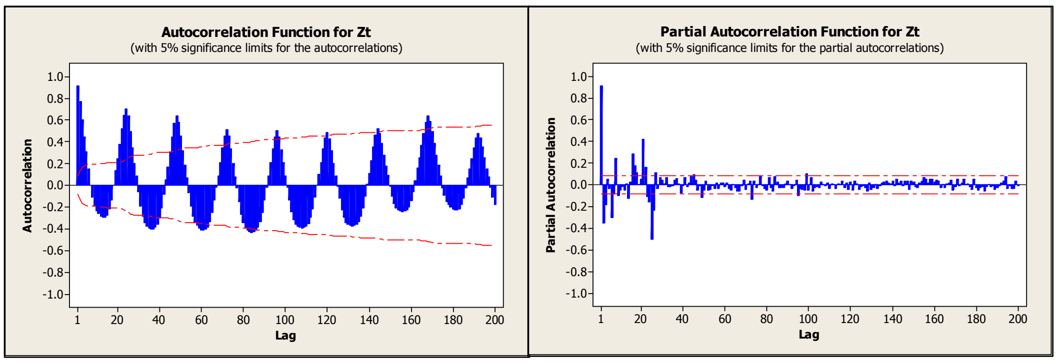
atau biasa dituliskan sebagai ARIMA (p,d,q).

Pemodelan ARIMA

Sebagai langkah awal dari metode ini, adalah identifikasi model melalui deskripsi data dengan menggunakan plot *time series*, ACF, dan PACF sebagai berikut :



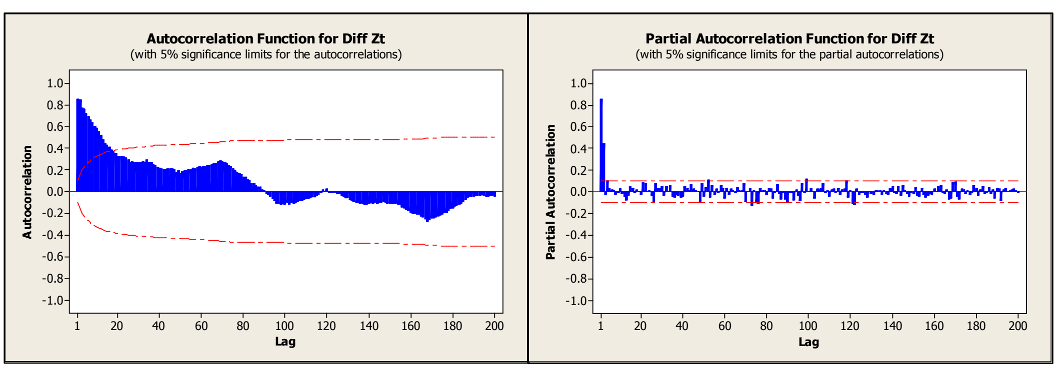
Gambar 1. Plot *Time series*



(a) (b)

Gambar 2. (a) Plot ACF dan (b) Plot PACF

Berdasarkan plot *time series* pada Gambar 1 terlihat bahwa data sudah stasioner dalam *varians* dan *mean.* Sedangkan plot ACF pada Gambar 2(a) memperlihatkan adanya indikasi pengaruh jam (*hour*) dan hari (*day*). Sehingga perlu dilakukan *differencing* untuk menghilangkan pengaruh jam dan hari. Pada permasalahan ini digunakan *differencing* 168 (24×7).



(a) (b)

Gambar 3. (a) Plot ACF dan (b) Plot PACF setelah dilakukan *differencing*

Plot pada Gambar 3 (a) dan (b) terlihat indikasi bahwa sudah tidak ada pengaruh jam dan hari. Plot ACF turun *dies down,* pada plot PACF, data keluar pada lag 1, 2, dan 3. Hal ini menunjukkan bahwa kemungkinan model datanya adalah ARIMA (3,168,0). Setelah kemungkinan modelnya diperoleh maka langkah selanjutnya adalah melakukan Estimasi dan *Testing Hypothesis.* Dengan menggunakan *software* Minitab, diperoleh hasil sebagai berikut :

The ARIMA Procedure
Conditional Least Squares Estimation

Parameter	Estimate	Standard Error	t Value	Approx Pr > |t|	Lag
AR1,1	0.47630	0.05123	9.30	<.0001	1
AR1,2	0.46072	0.05162	8.93	<.0001	2
AR1,3	-0.01769	0.05130	-0.34	**0.7304**	3

Variance Estimate	15.30039
Std Error Estimate	3.911572
AIC	2140.238
SBC	2152.09
Number of Residuals	384

* AIC and SBC do not include log determinant.

Correlations of Parameter Estimates

Parameter	AR1,1	AR1,2	AR1,3
AR1,1	1.000	-0.465	-0.452
AR1,2	-0.465	1.000	-0.463
AR1,3	-0.452	-0.463	1.000

Gambar 4. Hasil analisis model ARIMA (3,168,0)

Dari output pada Gambar 4 terlihat bahwa lag ke-tiga tidak signifikan, ini berarti bahwa model yang diperoleh belum tepat untuk menjelaskan data. Oleh karena itu akan dicoba untuk menghilangkan lag ke-tiga pada model sebelumnya, sehingga diperoleh model ARIMA (2,168,0). Estimasi dan uji hipotesis terhadap model tersebut adalah sebagai berikut:

The ARIMA Procedure
Conditional Least Squares Estimation

Parameter	Estimate	Standard Error	t Value	Approx Pr > |t|	Lag
AR1,1	0.46831	0.04564	10.26	<.0001	1
AR1,2	0.45249	0.04570	9.90	<.0001	2

Variance Estimate	15.2651
Std Error Estimate	3.907058
AIC	2138.358
SBC	2146.259
Number of Residuals	384

* AIC and SBC do not include log determinant.

Correlations of Parameter Estimates

Parameter	AR1,1	AR1,2
AR1,1	1.000	-0.852
AR1,2	-0.852	1.000

Gambar 5. Hasil analisis model ARIMA (2,168,0)

Dari output pada Gambar 5 dapat diketahui bahwa model ARIMA (2,168,0) signifikan. Langkah selanjutnya setelah estimasi dan uji hipotesis adalah uji terhadap asumsi residual yaitu *white noise* dan uji normalitasnya.

Autocorrelation Check of Residuals

To Lag	Chi-Square	DF	Pr > ChiSq	--------------------Autocorrelations--------------------					
6	4.69	4	**0.3201**	0.008	-0.036	-0.095	0.022	-0.007	0.035
12	8.90	10	**0.5420**	0.032	-0.010	-0.015	0.043	0.054	0.066
18	11.01	16	**0.8091**	0.027	-0.059	-0.005	0.020	-0.010	0.022
24	18.77	22	**0.6596**	-0.021	-0.100	-0.007	0.036	0.041	0.074
30	25.13	28	**0.6209**	0.008	-0.115	-0.009	-0.019	0.032	0.024
36	33.07	34	**0.5130**	-0.030	-0.012	-0.013	0.126	0.033	0.027
42	34.35	40	**0.7221**	0.007	0.025	-0.005	-0.034	0.001	-0.034
48	42.72	46	**0.6102**	-0.055	0.004	0.037	0.094	0.036	-0.068

Gambar 6. Hasil analisis Residual ARIMA (2,168,0)

Asumsi *white noise* sudah dipenuhi, hal ini terlihat dari semua nilai Pr > Chi-Square yang lebih dari 0.05.

MODEL REGRESI LINEAR SEDERHANA

Metode regresi merupakan metode yang memodelkan hubungan antara variabel respon (y) dan variabel bebas ($x_1, x_2, \ldots, x_p$). Model regresi sederhana dinyatakan dengan

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_p x_p + \varepsilon \tag{2}$$

Jika diambill sebanyak n pengamatan, maka model di atas dapat ditulis sebagai

$$y_i = \beta_0 + \sum_{k=1}^{p} \beta_k x_{ik} + \varepsilon_i \tag{3}$$

dengan $i = 1, 2, \ldots, n$; $k = 1, 2, \ldots, p$; $\beta_0, \beta_1, \ldots, \beta_p$ adalah parameter model dan $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n$ adalah *error* yang diasumsikan memiliki *mean* nol dan *varians* konstan σ^2. Pada model ini, hubungan antara variabel bebas dan variabel respon dianggap konstan. Estimator dari parameter model didapat dari persamaan

$$\hat{\boldsymbol{\beta}} = \left(\mathbf{X}^T \mathbf{X}\right)^{-1} \mathbf{X}^T \mathbf{y} \tag{4}$$

dengan

- $\boldsymbol{\beta}$: vektor dari parameter yang ditaksir berukuran $n \times (k+1)$
- $\mathbf{X}$: matrik data berukuran n × (k+1) dari variabel bebas yang elemen pada kolom pertama bernilai 1
- $\mathbf{Y}$: vektor observasi dari variabel respon berukuran ($n \times 1$)
- k : banyaknya variabel bebas

Dalam kasus ini, akan digunakan model yang telah dikembangkan oleh H.M. Al-Hamadi dan S.A. Soliman (2005), yang menggunakan *dummy variable* berdasarkan hari dan minggu sebagai variabel bebas untuk meramalkan penggunaan energi listrik beberapa tahun ke depan, dalam hal ini *dummy variable*-nya berdasarkan jam dan hari.

Model yang menggambarkan hubungan antar variabel tersebut dapat dinyatakan dalam persamaan regresi :

$$y = \beta_0 + \beta_1 H_1 + \beta_2 H_2 + \ldots + \beta_i H_j + \beta_{24} D_1 + \ldots + \beta_i D_k + \varepsilon \tag{5}$$

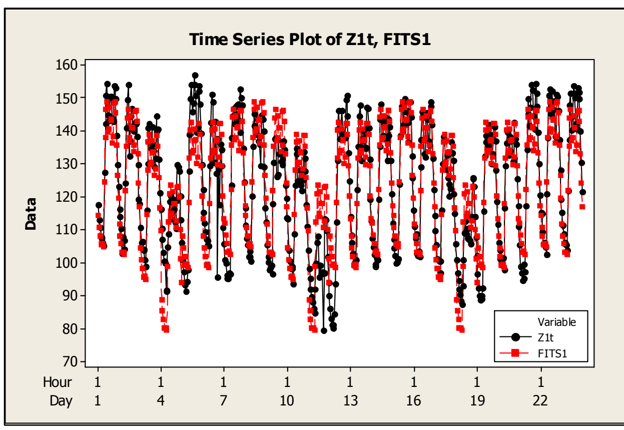
dengan β_i , i = 0, ..., 29 adalah parameter regresi pada jam ke-j, j = 1, ..., 24 dan pada hari ke-k, k= 1, ..., 7.

Pemodelan Regresi Linier Sederhana dengan Residual ARIMA

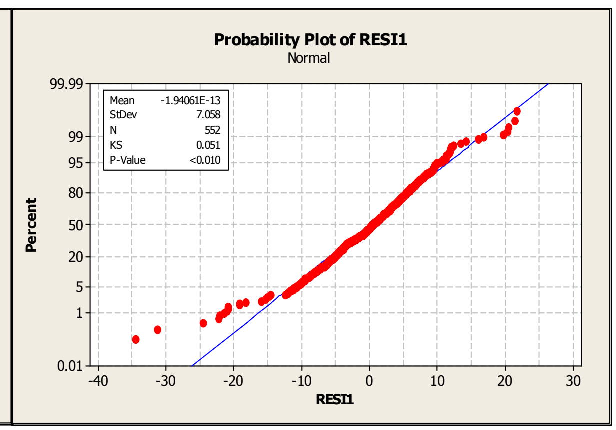
Proses pemodelan data dengan regresi linier sederhana menggunakan *software* Minitab 14, menghasilkan model :

$$
\begin{aligned}
Z_t \;=\; & 94{,}1 - 5{,}20\,H_1 - 8{,}77\,H_2 - 11{,}4\,H_3 - 14{,}0\,H_4 - 13{,}9\,H_5 - 13{,}8H_6 - 14{,}6\,H_7 + 4{,}83\,H_8 \\
& + 18{,}3\,H_9 + 26{,}9\,H_{10} + 29{,}3\,H_{11} + 20{,}0\,H_{12} + 20{,}8\,H_{13} + 27{,}4\,H_{14} + 23{,}7\,H_{15} + \\
& 23{,}9\,H_{16} + 16{,}6\,H_{17} + 18{,}7\,H_{18} + 28{,}9\,H_{19} + 29{,}0\,H_{20} + 25{,}7\,H_{21} + 16{,}1\,H_{22} + \\
& 6{,}98\,H_{23} + 19{,}3\,D_1 + 19{,}2\,D_2 + 23{,}1\,D_3 + 25{,}5\,D_4 + 22{,}9\,D_5 + 15{,}6\,D_6
\end{aligned}
$$

Seluruh variabel dalam model memberikan pengaruh yang signifikan, nilai R^2 juga cukup besar, yaitu 86,6 persen, berarti model ini cukup bagus dalam menerangkan keragaman data. Hasil prediksi data dapat dilihat pada gambar berikut :



Gambar 7 (a). Fits data prediksi dan data asli



Gambar 7 (b). Plot Residual

Meskipun hasil fits data tampak sesuai dengan data asli, namun pada plot residual, terlihat bahwa asumsi kenormalan residual tidak terpenuhi, sehingga perlu dilakukan pemodelan terhadap residual. Dalam hal ini, residual akan dimodelkan dengan model ARIMA. Hasil pemodelan dengan menggunakan *software* SAS terlihat pada gambar berikut :

The ARIMA Procedure

Conditional Least Squares Estimation

Parameter	Estimate	Standard Error	t Value	Approx Pr > |t|	Lag	Variable	Shift
AR1,1	0.47855	0.04728	10.12	<.0001	1	y	0
AR1,2	0.47052	0.04732	9.94	<.0001	2	y	0
NUM1	0.07863	0.03121	2.52	0.0122	0	y	0
NUM2	-0.16437	1.13349	-0.15	0.8848	0	H1	0
NUM3	-0.02078	1.13305	-0.02	0.9854	0	H2	0
NUM4	-0.11270	1.35351	-0.08	0.9337	0	H3	0
NUM5	0.46357	1.43957	0.32	0.7476	0	H4	0
NUM6	0.07095	1.54301	0.05	0.9633	0	H5	0
NUM7	0.89181	1.60559	0.56	0.5789	0	H6	0
NUM8	0.37214	1.67064	0.22	0.8239	0	H7	0
NUM9	-3.06629	1.74633	-1.76	0.0800	0	H8	0
NUM10	-3.37021	1.92144	-1.75	0.0803	0	H9	0
NUM11	-4.18432	2.06733	-2.02	0.0437	0	H10	0
NUM12	-4.22638	2.11883	-1.99	0.0468	0	H11	0
NUM13	-2.08002	1.98733	-1.05	0.2960	0	H12	0
NUM14	-1.79024	1.98848	-0.90	0.3686	0	H13	0
NUM15	-2.93254	2.06926	-1.42	0.1573	0	H14	0
NUM16	-2.57409	1.98492	-1.30	0.1955	0	H15	0
NUM1	-2.42187	1.95577	-1.24	0.2164	0	H16	0
NUM18	-0.37963	1.82375	-0.21	0.8352	0	H17	0
NUM19	-2.33713	1.76730	-1.32	0.1869	0	H18	0
NUM20	-1.23175	1.88665	-0.65	0.5143	0	H19	0
NUM21	-2.46095	1.77997	-1.38	0.1677	0	H20	0
NUM22	-2.52612	1.64473	-1.54	0.1255	0	H21	0
NUM23	-1.68958	1.27039	-1.33	0.1844	0	H22	0
NUM24	-0.13601	1.18368	-0.11	0.9086	0	H23	0
NUM25	-2.47220	2.30535	-1.07	0.2843	0	D1	0
NUM26	-3.20818	2.91091	-1.10	0.2712	0	D2	0
NUM27	-6.26394	3.09870	-2.02	0.0440	0	D3	0
NUM28	-7.25180	2.95860	-2.45	0.0147	0	D4	0
NUM29	-4.92969	2.83572	-1.74	0.0830	0	D5	0
NUM30	-2.57412	2.28967	-1.12	0.2617	0	D6	0

Gambar 8. Hasil analisis Residual dengan ARIMA

Dari output pada Gambar 8 terlihat bahwa model ARIMA signifikan. Langkah selanjutnya setelah estimasi dan uji hipotesis adalah uji terhadap asumsi residual yaitu *white noise* dan uji normalitasnya.

Autocorrelation Check of Residuals

Variance Estimate	15.48069
Std Error Estimate	3.934551
AIC	2172.336
SBC	2298.757
Number of Residuals	384

* AIC and SBC do not include log determinant.

To Lag	Chi-Square	DF	Pr > ChiSq	Autocorrelations					
6	5.48	4	**0.2416**	0.015	-0.031	-0.110	0.017	-0.023	0.002
12	8.93	10	**0.5391**	0.004	0.011	0.007	0.043	0.040	0.071
18	11.09	16	**0.8041**	-0.026	-0.050	0.017	0.028	-0.033	0.001
24	17.54	22	**0.7328**	-0.018	-0.098	-0.006	0.039	0.047	0.047
30	23.73	28	**0.6957**	0.025	-0.111	-0.021	-0.025	0.029	0.004
36	32.71	34	**0.5306**	-0.037	0.015	0.020	0.132	0.019	0.037
42	34.44	40	**0.7181**	0.007	0.031	0.014	-0.029	-0.012	-0.042
48	46.06	46	**0.4698**	-0.061	0.002	0.043	0.091	0.031	-0.108

Gambar 9. Hasil analisis Residual ARIMA

Hasil analisis dengan SAS pada Gambar 9 menunjukkan residualnya juga sudah memenuhi asumsi *white noise*.

PERBANDINGAN MODEL ARIMA DAN MODEL REGRESI DENGAN RESIDUAL ARIMA

Berdasarkan analisis pada bagian-bagian sebelumnya maka dapat dibandingkan kedua metode tersebut dengan melihat SSE (*Standard Error Estimate*) dan AIC (*Akaike's Information Criterion*) sebagai berikut:

Tabel 1. Perbandingan nilai SSE dan AIC kedua metode

Model ARIMA		Model Regresi dengan Residual ARIMA	
Variance Estimate	15.2651	Variance Estimate	15.48069
Std Error Estimate	3.907058	Std Error Estimate	3.934551
AIC	2138.358	AIC	2172.336
SBC	2146.259	SBC	2298.757
Number of Residuals	384	Number of Residuals	384

Dari hasil Tabel 1 dapat dilihat bahwa nilai SSE dan AIC model ARIMA lebih kecil yang menunjukkan bahwa kesalahan menggunakan model ARIMA lebih kecil dibanding model Regresi dengan residual ARIMA. Oleh karena itu model ARIMA lebih sesuai digunakan untuk menerangkan perilaku pelanggan listrik di Kota Palopo.

SIMPULAN

Nilai SSE dan AIC model ARIMA yang lebih kecil dari model Regresi dengan Residual ARIMA dapat dijadikan acuan untuk mengambil kesimpulan bahwa model ARIMA lebih sesuai untuk menerangkan perilaku pelanggan listrik di Kota Palopo daripada model Regresi dengan Residual ARIMA.

DAFTAR PUSTAKA

Cryer, J. D. 1986. *Time series Analysis*. Boston: PWS-KENT Publishing Company.

Gujarati, D.N. 2003. *Basic Econometrics*. Singapore: McGraw-Hill Education (Asia).

H.M. AL-Hamadi, S.A.Soliman. 200). Long-term/mid-term Electric Load Forecasting Based On Short-term Correlation and Annual Growth, *Electric Power System Research* Vol. 74: 353-361.

Draper, N.R. & Smith, H. 1981. *Applied Regression Analysis, 2nd edition*. New York: John Wiley & Sons.

Wei, W.W.S. 1990. *Time series Analysis, Univariate and Multivariate Methods*. Otawa: Addison Wesley Publishing Company.