SEMINAR NASIONAL
MATEMATIKA DAN PENDIDIKAN MATEMATIKA
FMIPA UNIVERSITAS NEGERI YOGYAKARTA

PERBANDINGAN MODEL PADA DATA DERET WAKTU PEMAKAIAN LISTRIK JANGKA PENDEK YANG MENGANDUNG POLA MUSIMAN GANDA

Gumgum Darmawan [1], Suhartono[2]
[1] Staf Pengajar Jurusan Statistika FMIPA UNPAD
[2] Staf Pengajar Jurusan Statistika FMIPA ITS Surabaya

ABSTRAK

Pada makalah ini akan di bandingkan dua model dari data deret waktu yang mengandung pola musiman ganda. Kedua model yang dibandingkan adalah model ARIMA musiman dan model ARFIMA musiman. Data yang digunakan adalah data pemakaian listrik jangka pendek di Pulau Batam periode Januari – Desember 2006 yang dihitung dalam satuan jam. Berdasarkan pada gambar ACF dan PACF dari data pemakaian listrik di pulau Batam, data deret waktu memperlihatkan pola musiman ganda dengan periode harian $s_1 = 24$ dan mingguan $s_2 = 168$. Setelah dianalisis hasilnya menunjukkan model musiman ganda ARFIMA lebih baik dibandingkan dengan model musiman ganda ARIMA berdasarkan nilai-nilai AIC, MSE dan MAPE.

Kata Kunci : ARFIMA, Musiman Ganda, Periodogram.

1. PENDAHULUAN

Data deret waktu merupakan pengamatan terurut dari suatu pengamatan berdasarkan waktu. Observasi yang diamati merupakan barisan data yang bernilai kontinyu atau diskrit yang diperoleh pada interval waktu yang sama, misalnya harian, mingguan, bulanan dan sebagainya. Untuk mendapatkan model dari data yang diperoleh dari observasi tersebut diperlukan suatu pemodelan deret waktu seperti *autoregressive* (AR), *moving average* (MA), *autoregressive moving average* (ARMA) dan *autoregressive integrated moving average* (ARIMA). Model-model ini digunakan untuk memodelkan data deret waktu dengan ketergantungan jangka pendek (*short memory*), yaitu data dengan periode terpisah jauh diasumsikan tidak memiliki hubungan (tidak berkorelasi).

Data deret waktu akan mempunyai sifat ketergantungan jangka panjang (*long memory*) jika diantara pengamatan dengan periode yang terpisah jauh masih mempunyai korelasi yang tinggi. Sifat dari data deret waktu seperti ini mempunyai fungsi autokorelasi (ACF) ρ_k untuk lag ke-k, turun secara hiperbolik. Sedangkan data deret waktu yang stasioner akan mempunyai sifat ketergantungan jangka pendek (*short memory*) jika mempunyai fungsi autokorelasi ρ_k yang turun secara cepat atau turun secara eksponensial. Model data deret waktu ketergantungan jangka panjang sering disebut dengan Model ARFIMA (*Autoregressive Fractionally Integrated Moving Average*) dengan parameter pembeda berbentuk bilangan pecahan, ini berbeda dengan model ARIMA yang mempunyai parameter pembeda berupa bilangan bulat.

Penelitian mengenai model ARFIMA pertama kali dikembangkan oleh Granger dan Joyeux (1980) yang merupakan pengembangan dari model ARIMA. Hosking (1981) mengkaji sifat-sifat *long memory* dari model ARFIMA stasioner dan nonstasioner. Aplikasi dari model ARFIMA melalui regresi spektral diatas telah banyak dikaji, antara lain oleh Lopes, Olberman dan Reisen (2004) yang mengaplikasikannya untuk data tingkat suku bunga jangka pendek di Inggris. Lopes and Nunes (2006) mengaplikasikan model ARFIMA pada data DNA.

Dalam penelitian ini, penaksiran parameter ARFIMA akan dilakukan dengan menggunakan metode regresi spektral. Penaksiran parameter melalui metode ini pertama kali dikembangkan oleh Geweke dan Porter-Hudak (1983) yang disingkat dengan GPH, dengan menggunakan log Periodogram sebagai variabel tak bebasnya. Akurasi hasil peramalan dari model ARIMA musiman dan ARFIMA musiman akan dibandingkan berdasarkan nilai AIC, MSE dan MAPE.

2. MODEL MATEMATIKA ARFIMA MUSIMAN GANDA

Model ARFIMA mampu memodelkan proses ketergantungan jangka pendek dan jangka panjang. Pengamatan-pengamatan yang dihasilkan oleh struktur ARMA menunjukkan ketergantungan jangka pendek, sedangkan parameter pembedaan pecahan d, yang menyebabkan nilai-nilai ACF turun secara hiperbolik menunjukkan ketergantungan jangka panjang. Model ARFIMA Musiman Ganda mempunyai persamaan sebagai berikut :

$$\phi_p(B)(1-B)^d\,\Phi_{P_1}(B^{s_1})(1-B^{s_1})^{D_1}\,\Phi_{P_2}(B^{s_2})(1-B^{s_2})^{D_2}\,\dot{z}_t = \theta_q(B)\times \Theta_{Q_1}(B^{s_1})\,\Theta_{Q_2}(B^{s_2})\,a_t \quad\text{.....................(1)}$$

$\dot{z}_t$ = Data deret waktu dengan rata-rata μ.

$\phi_p(B)$ = persamaan polinomial autoregresi nonmusiman

$\theta_q(B)$ = persamaan polinomial *moving average* nonmusiman

$\Phi_{P_1}\left(B^{s_1}\right)$ = persamaan polinomial autoregresi musiman pertama

$\Phi_{P_2}\left(B^{s_2}\right)$ = persamaan polinomial autoregresi musiman kedua

$\Theta_{Q_1}\left(B^{s_1}\right)$ = persamaan polinomial *moving average* musiman pertama

$\Theta_{Q_2}\left(B^{s_2}\right)$ = persamaan polinomial *moving average* musiman kedua

$(1-B)^d = \nabla^d = \sum_{k=0}^{\infty}\binom{d}{k}\left(\;{}^{k}\right) - 1$ operator pembeda pecahan

$\left(1-B^{s_1}\right)^{D_1}$ = operator pembeda musiman pertama dengan periode s_1

$\left(1-B^{s_2}\right)^{D_2}$ = operator pembeda musiman pertama dengan periode s_2

a_t = sisa dari data deret waktu yang diasumsikan *white noise*.

Dari persamaan (1) diatas, jika salah satu dari nilai *d*, D_1 dan D_2 bernilai pecahan maka dapat dinyatakan sebagai model *long memory* atau ARFIMA. Jika nilai *d*, D_1 dan D_2 semuanya bernilai integer (bilangan bulat) maka model (1) merupakan model *short memory* atau model ARIMA.

3. METODE PENELITIAN

Pemodelan ARFIMA yang dilakukan melalui metode Box-Jenkins dilakukan dalam beberapa tahap yaitu identifikasi, estimasi, verifikasi (*diagnostic check*) dan peramalan. Walaupun metode yang digunakan mempunyai tahapan yang sama dengan model ARIMA akan tetapi tiap tahapannya mempunyai perbedaan tersendiri.

3.1 Identifikasi Model ARFIMA

Langkah-langkah identifikasi dalam model ARFIMA sebagai berikut :

a) Membuat plot data deret waktu dan pilih transformasi yang sesuai untuk menstabilkan varians, jika data tidak stasioner dalam varian.

b) Mengidentifikasi pola *long memory* melalui Statistik Hurst (H). Jika nilai H-nya 0,5, maka data menunjukkan pola short memory, jika 0<H<0,5, maka data menunjukkan pola *inremediate memory* dan jika 0,5<H<1 maka data menunjukkan pola *long memory*.

c) Membuat plot ACF dan PACF untuk memperkirakan parameter *p* dan *q*.

Untuk memperkirakan nilai *p* dan *q* data terlebih dahulu dilakukan pembedaan setelah nilai pembedanya di ketahui. Data hasil pembedaan tersebut kemudian di plot

ACF dan PACF-nya. Parameter p ditentukan berdasarkan lag yang keluar/signifikan dari plot PACF sedangkan parameter q ditentukan dari lag yang keluar/signifikan dari plot ACF.

3.2 Penaksiran Parameter Model ARFIMA

Penaksiran parameter pada model ARFIMA dilakukan dua tahap yaitu penaksiran parameter pembeda (d) dan penaksiran parameter ϕ serta θ. Penaksiran parameter ϕ dan θ dilakukan melalui metode Least Square. Sedangkan penaksiran parameter d dilakukan melalui metode regresi spektral.

Penaksiran parameter model ARFIMA melalui metode regresi spektral untuk parameter pembeda (d) diusulkan oleh Geweke dan Porter-Hudak tahun 1983, pertama membentuk fungsi densitas spektral menjadi persamaan regresi linier dan menaksir parameter d melalui metode *Ordinary Least Square* (*OLS*).

3.3 Uji Diagnostik dan Peramalan Pada Model ARFIMA musiman ganda

Uji diagnostik (*diagnostic checking*) dapat dibagi ke dalam dua bagian, yaitu uji signifikansi parameter dan uji kesesuaian model (terdiri dari uji asumsi *white noise* dan distribusi Normal dari sisa). Pada uji signifikansi parameter digunakan Distribusi t untuk melihat apakah parameter model ARFIMA sama dengan nol atau tidak. Uji asumsi *White Noise* menggunakan *Portmanteu* dan untuk menguji normalitas dari sisa menggunakan Kolmogorov Smirnov. Sedangkan peramalan model ARFIMA musiman ganda, karena cukup kompleknya persamaan menggunakan software SAS versi 9.

3.4 Kriteria Pemilihan Model Terbaik

Untuk memilih model terbaik digunakan beberapa kriteria pemilihan model, yaitu kriteria *in sample* (*training*) dan *out sample* (*testing*). Kriteria *in sample* dilakukan melalui persamaan *AIC* (Akaike's *Information Criterion*). Pada penentuan model terbaik melalui kriteria *out sample* dilakukan melalui persamaan *MSE* (*Mean Square Error*) dan *MAPE* (*Mean Absolute Percentage Error*).

4 HASIL PENELITIAN

Sesuai dengan permasalahan yang akan dibahas dalam penelitian ini, maka bahan yang digunakan adalah :

1. Data pendukung untuk menerapkan pendekatan ARFIMA pada penelitian ini. Data yang digunakan adalah data pemakaian listrik perhari di pulau Batam pada bulan Januari – Desember tahun 2006 yang diukur tiap jam dalam Mega Watt.

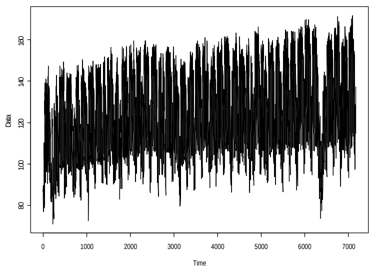
Data yang digunakan untuk pemodelan sebanyak 7172 pengamatan. Seratus (100) data terakhir digunakan untuk validasi peramalan, sehingga keseluruhan data sebanyak 7272 pengamatan.

2. Software yang digunakan untuk mengolah data adalah software R versi 2.7.2, SAS versi 9 dan Minitab versi 14.

4.1 Identifikasi Model

Langkah pertama dalam pembentukan model data deret waktu melalui metode Box Jenkins adalah identifikasi model. Pada tahapan identifikasi model, plot data untuk melihat stasioneritas, kemudian diperkirakan nilai *H* dan membuat plot ACF serta PACF.

a) Membuat plot data deret waktu.



Gambar 1 Plot Data Pemakaian Listrik di Pulau Batam

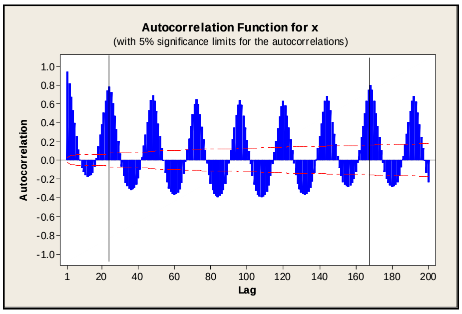
Plot data deret waktu dari data di atas memperlihatkan grafik yang datar, akan tetapi ada beberapa data yang menjurai kebawah. Secara keseluruhan dapat dikatakan data relatif stasioner secara rata-rata. Adanya beberapa data yang menjurai kebawah ini dapat mengakibatkan data tidak stasioner dalam varian untuk itu dilakukan transformasi *ln* pada data di atas. Transformasi ini didapat karena mempunyai nilai $\lambda = 0$.

b) Mengidentifikasi statistik Hurts. Setelah di identifikasi nilai dari statistik Hurts dengan bantuan Software R diperoleh H=0,88 ini menunjukkan data mengikuti pola *long memory*.

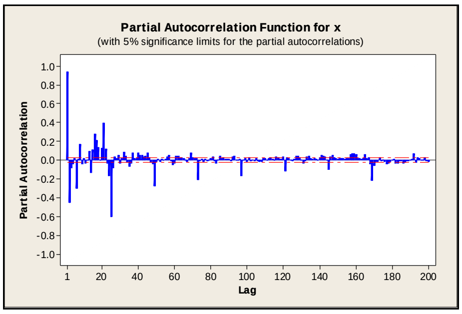
c) Membuat plot ACF dan PACF untuk memperkirakan parameter p dan q.

Sebelum memperkirakan parameter p dan q, terlebih dahulu di buat plot ACF dan PACF dari data deret waktunya. Data deret waktu yang dibuat ACF dan PACF-nya sudah ditransformasi sehingga stasioner dalam varians, hasilnya dapat dilihat pada Gambar 2.

Gambar 2, tampak bahwa data mempunyai pola musiman ganda, pada plot ACF pola musiman pertama mempunyai periode 24 artinya data mempunyai siklus 24 jam atau harian. Musiman ke dua mempunyai periode 168 ini terlihat dari plot ACF nilai autokorelasi kembali meningkat pada lag ke 168, artinya data mempunyai siklus mingguan. Dari plot PACF tampak banyaknya garis lag yang signifikan, ini menunjukkan bahwa data mempunyai koefisien dengan order tinggi. Berdasarkan plot ACF dan PACF, yang mengindikasikan pola musiman ganda, dengan nilai p dan q yang mungkin adalah 1, 24 dan 168.



(i) Plot ACF



(ii) Plot PACF

Gambar 2. Plot ACF dan PACF dari Data Pemakaian Listrik di Pulau Batam

4.2 Penaksiran Parameter

Setelah diidentifikasi nilai p dan q yang mungkin maka selanjutnya ditaksir nilai parameter *d* dengan menggunakan metode GPH. Berdasarkan software R didapat nilai *d* adalah 0,48 dengan standar deviasi 0,07. Sehingga nilai *d* yang memungkinkan berkisar antara 0,41 sampai 0,55. Setelah parameter *d* ditentukan, selanjutnya ditaksir juga parameter psi dan theta melalui Metode *Unconditional Least Square*.

Tabel 1 Penentuan Model ARFIMA Terbaik

NO	Parameter *p* dan *q*	Nilai *d*	*White Noise*	*MSE*	*AIC*
1	p = (1,2,3), q = (1,2,3,4)(24)(168)	0,447	Sampai lag 36	0,0258	-31105
2	p = (1,2,3), q = (1,2,3,4)(24,168)	0,447	Sampai lag 18	0,0258	-31101
3	p = (1,2,3)(24,168), q = (1,2,3,4)	0,447	Sampai lag 18	0,0290	-29674
4	p = (1,2,3)(24,168), q = (1,2,3,4)(24)	0,447	Sampai lag 18	0,0289	-29674
5	p = (1,2,3), q = (1,2,3,4)(24,48)(168)	0,447	Sampai lag 18	0,0259	-31038
6	p = (1,2,3), q = (1,2,3,4)(168)	0,447	Sampai lag18	0,0260	-31028
7	p = (1,2,3), q = (1,2,3,4)(24,168)	0,400	Sampai lag 18	0,0258	-3109,6
8	p = (1,2,3), q = (1,2,3,4)(168)	0,400	Sampai lag 18	0,0260	-31034
9	p = (1,2,3)(24,168), q = (1,2,3,4)(24)	0,400	Sampai lag 24	0,0289	-29664
10	p = (1,2,3)(24,48), q = (1,2,3,4)(24)(168)	0,535	Sampai lag 48	0,0257	-31147

Berdasarkan Tabel 1, model ARFIMA terbaik adalah model no 10 berdasarkan tiga kriteria di atas. Model ini *white* noise sampai lag 48 dan mempunyai nilai *MSE* dan *AIC* terkecil.

Sehingga berdasarkan hasil penelusuran didapat dua model terbaik yaitu Model ARIMA$(2,0,0)(1,1,0)_{24}(1,1,0)_{168}$ dan Model ARFIMA$(3,d,4)(0,0,2)_{24}(0,1,1)_{168}$ dengan d = 0,535.

Penaksiran Parameter

Berdasarkan hasil penentuan model terbaik, diperoleh persamaan matematik untuk kedua model sebagai berikut;

1) Model ARIMA$(2,0,0)(1,1,0)_{24}(1,1,0)_{168}$

$$\left(1-0{,}739B-1{,}87B^{2}\right)\left(1+0{,}452B^{24}\right)\left(1-B^{24}\right)\left(1-B^{168}\right)\left(1+0{,}534B^{168}\right)Z_t$$

2) Model ARFIMA$(3,d,4)(0,0,2)_{24}(0,1,1)_{168}$ dengan d = 0,535

$$\left(1-1{,}78B+1{,}72B^{2}-0{,}75B^{3}\right)\left(1-B\right)^{0{,}535}\left(1-B^{168}\right)Z_t=\left(1-1{,}51B+1{,}5B^{2}-0{,}56B^{3}+0{,}05B^{4}\right)$$
$$\times\left(1+0{,}12B^{24}+0{,}06B^{48}\right)\left(1-0{,}87B^{168}\right)a_t$$

4.3 Uji Diagnostik

Uji diagnostik dilakukan untuk menguji signifikansi parameter, independensi dan normalitas dari sisa. Untuk menguji signifikansi parameter digunakan distribusi t dengan α = 0,05, uji independensi atau *White Noise* melalui statistik Ljung Box dan uji normalitas melalui Kolmogorov-Smirnov.

a. **Uji Signifikansi Parameter**

Pengujian signifikansi parameter dilakukan untuk menguji apakah parameter dalam model sama dengan nol atau tidak. Jika dalam pengujian parameter tidak sama dengan nol ini menunjukkan bahwa parameter tersebut layak untuk diikutsertakan dalam model. Pada pengujian parameter model ARIMA maupun ARFIMA digunakan distribusi *student t* dengan menggunakan tingkat signifikansi 0.05. Berdasarkan hasil pengujian semua parameter yang terlibat pada kedua model tersebut bersifat signifikan pada α = 5%.

b. Uji *White Noise*

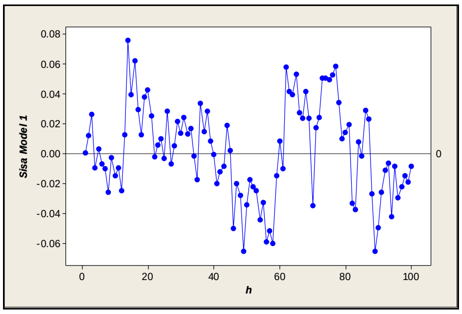
Uji *White Noise* dilakukan untuk melihat apakah sisa dari model sudah tidak berkorelasi satu dengan yang lainnya. Pengujian *White Noise* sisa dari kedua model diperiksa sampai lag ke 48. Berdasarkan pengujian *white noise*,model ARIMA *white noise* sampai lag 12, sedangkan Pengujian *white noise* dari model ARFIMA didapat bahwa model *white noise* sampai lag ke - 48.

c **Uji Normalitas Sisa**

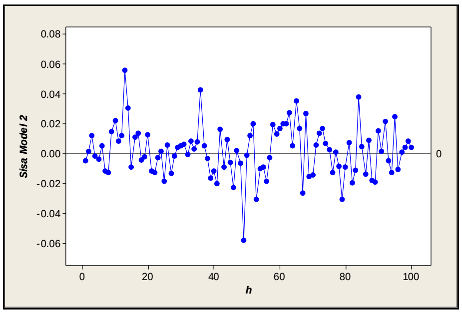
Dari hasil pengujian normalitas sisa dari keempat model ternyata tidak ada yang memenuhi syarat normalitas sehingga pada akhirnya akan dilakukan pemeriksaan terhadap nilai kurtosis dan *skewness* dari sisa. Melalui bantuan software SAS, diperoleh nilai kurtosis untuk Model ARIMA 9,699 dan untuk Model ARFIMA 20,08 dan keduadaunya mempunyai rata-rata di nol. Sehingga sisa dari kedua model membentuk kurva leptokurtik dengan pusat rata-rata di nol.

4.4 Peramalan

Dengan menggunakan kedua model yaitu ARIMA, ARFIMA, didapat 100 peramalan ke depan. Gambar plot *fit* dan data aktual untuk 100 peramalan ke depan untuk masing masing model deret waktu dapat dilihat pada Gambar 4.



(i) ARIMA



(ii) ARFIMA

Gambar 4 Plot Sisa dari model ARIMA dan ARFIMA

Gambar 4 memperlihatkan plot sisa dari ke dua model. Tampak bahwa plot sisa dari Model ARFIMA relatif datar dengan garis tengahnya merupakan sumbu y = 0. Dari gambar tersebut menunjukkan bahwa Model ARFIMA memberikan taksiran yang baik dibandingkan dengan Model ARIMA.

4.5 Pemilihan Model Terbaik

Dari kedua model yang telah ditentukan perlu dilakukan pemilihan model terbaik yang paling layak untuk memodelkan data deret waktu. Pada penelitian ini ditentukan tiga kriteria yaitu *AIC*, MSE dan MAPE.

Tabel 2 Kriteria Pemilihan Model Terbaik

MODEL	AIC	MSE	MAPE
ARIMA$(2,0,0)(1,1,0)_{24}(1,1,0)_{168}$	-27.005	0,0346	1,8%
ARFIMA$(3,d,4)(0,0,1)_{24}(0,1,1)_{168}$, d = 0,535	-31.147	0,0245	0,9%

Dari Tabel 2 tampak bahwa Model ARFIMA memberikan hasil yang lebih baik dibandingkan Model ARIMA berdasarkan nilai AIC, MSE dan MAPE. Perbedaan yang cukup berarti terlihat dari nilai MAPE, dimana nilai MAPE dari Model ARIMA sebesar 1,8% sedangkan nilai MAPE dari Model ARFIMA sebesar 0,9%.

5 KESIMPULAN

Berdasarkan hasil analisis, model *Double Seasonal* ARFIMA lebih baik dibandingkan dengan model *Double Seasonal* ARIMA dalam memodelkan pemakaian listrik di Pulau Batam, dilihat dari nilai *AIC*, *MSE* dan *MAPE*. Model ARIMA$(2,0,0)(1,1,0)_{24}(1,1,0)_{168}$ mempunyai AIC -27.005, nilai *MSE* sebesar 0,0346 dan nilai *MAPE* sebesar 1,8%. Sedangkan, Model ARFIMA$(3,d,4)(0,0,2)_{24}(0,1,1)_{168}$ dengan d = 0,535 mempunyai nilai *AIC* sebesar -31.147, nilai *MSE* sebesar 0,0245 dan nilai *MAPE* sebesar 0,9%.

DAFTAR PUSTAKA

Geweke J dan Porter-Hudak,S. (1983), “The Estimation and Application of Long Memory Time Series Models”, *Journal of Time series Analysis,*Vol. 4, hal. 221-238.

Granger, C. W. J. dan Joyeux,R. (1980), “An Introduction to Long-Memory Time Series Models and Fractional Differencing”, *Journal of Time Series Analysis*, Vol. 1, hal. 15-29.

Hosking, J.R.M. (1981), “Fractional Differencing”, *Biometika*, Vol. 68, hal. 165-176.

Hurst, H.E. (1951), “Long-Term Storage of Reservoirs: An Experimental Study”, *Transactions of the American Society of Civil Engineers*, Vol. 116, hal. 770-799.

Lopes, S.R.C dan Nunes,M.A. (2006), ”Long Memory Analysis in DNA Sequences”, *Physica A*, Vol. 361, hal. 569-588.

Lopes, S.R.C.,Olberman,B.P dan Reisen,V.A. (2004), ”A Comparison of Estimation Methods in Non-Stationary ARFIMA Processes”, *Journal of Statistical Computation & Simulation*, Vol. 74, No. 5, hal. 339-347.