PEMODELAN ARFIMA NONSTASIONER MELALUI METODE MODIFIKASI GPH (GEWEKE *AND* PORTER- HUDAK)

Gumgum Darmawan
Staf Pengajar Jurusan Statistika FMIPA UNPAD
e-mail : gumstat_1973@yahoo.com

ABSTRAK

Pada makalah ini akan di bandingkan dua metode pemodelan ARFIMA melalui Metode GPH (Geweke *and* Porter-Hudak) dan Metode MGPH (*Modification of* GPH). Kedua metode diatas mempunyai bentuk periodogram yang sama sebagai variabel tak bebas dari regresi spektral hanya berbeda pada variabel bebasnya. Berdasarkan hasil simulasi, akurasi penaksiran parameter Metode GPH mempunyai akurasi yang lebih baik untuk Model ARFIMA(1,*d*,0) dibandingkan dengan Metode MGPH, sedangkan untuk Model ARFIMA(0,*d*,1) Metode MGPH lebih baik dibandingkan dengan Metode GPH. Secara keseluruhan, Metode MGPH mempunyai standar deviasi penaksiran parameter Model ARFIMA yang lebih baik dibandingkan dengan Metode GPH.

Kata Kunci : ARFIMA, Periodogram, Regresi Spektral

1. Pendahuluan

Metode GPH (Geweke *and* Porter-Hudak) adalah metode penaksiran parameter dari Model ARFIMA (*Autoregressive Fractionally Integrated Moving Average*). Kelebihan Metode GPH dibandingkan dengan yang lainnya seperti Metode Maksimum Likelihood (Sowell,1992), Metode *Nonlinear Least Square* (Beran,1995) adalah fleksibilitas dalam penaksiran parameterya. Penaksiran parameter pembeda d pada Metode GPH dapat dilakukan secara langsung tanpa mengetahui nilai parameter p dan q terlebih dahulu.

Metode GPH pertama kali diusulkan oleh Geweke dan Porter-Hudak pada tahun 1983, dimana persamaan spektral ARFIMA dibentuk menjadi persamaan regresi spektral dengan log periodogram sebagai variabel tak bebasnya. Metode ini dikembangkan oleh Reisen (1994) yang dikenal dengan Metode SPR (*Smoothed Periodogram Regression*), Reisen melakukan pemulusan pada periodogram dengan menggunakan window Parzen. Pada tahun 1995, Robinson melalukan *Trimming* untuk mengubah nilai *bandwith* pada Metode GPH yang dikenal dengan GPHT*r* (GPH *trimming*). Hurvich dan Ray (1995) dan Velasco (1999a), melakukan *tapering* pada periodogram dengan menggunakan taper Cosine-Bell, metode ini dikenal dengan nama GPH*Ta* (GPH *tapering*). Kemudian Velasco (1999b), memodifikasi variabel bebas dari persamaan regresi spektral Metode GPH dari bentuk $2\sin\left(\omega_j/2\right)$ menjadi j, metode ini dikenal dengan nama MGPH (*Modification* of Geweke and Porter – Hudak).

Perbandingan akurasi penaksiran parameter Model ARFIMA dengan menggunakan metode-metode diatas telah banyak dikaji oleh para peneliti. Reisen dan Lopes (1999), membandingkan akurasi penaksiran Metode GPH dan SPR, dari hasil simulasi Metode SPR mempunyai akurasi yang lebih baik dibandingkan dengan metode GPH. Lopes, Olberman dan Reisen (2004) membandingkan akurasi penaksiran dari kelima metode diatas, yaitu GPH, SPR, GPHTr, GPHTa dan MGPH untuk Model ARFIMA Nonstasioner hasilnya Metode SPR menunjukkan akurasi terbaik, akan tetapi Metode MGPH mempunyai standar deviasi yang paling kecil.

Pada makalah ini akan dikaji akurasi penaksiran parameter pembeda melalui Metode MGPH, walaupun tidak seakurat Metode SPR akan tetapi metode ini mempunyai nilai standar deviasi yang relatif kecil, sehingga perlu dipertimbangkan. Disamping itu, pada makalah ini akan dikaji akurasi peramalan dari kedua metode diatas untuk Model ARFIMA (1,*d*,0) dan Model ARFIMA (0,*d*,1)

2. Persamaan ARFIMA

Model ARFIMA (*p, d, q*) yang dikembangkan Granger dan Joyeux (1980) ditulis sebagai :

$$\phi(B)(1-B)^d\left(Z_t-\mu\right)=\theta(B)a_t, \qquad (1)$$

dengan :

t = indeks dari pengamatan,

d = parameter pembeda (bilangan pecahan),

μ = rata-rata dari pengamatan,

$a_t \sim$ IIDN(0, σ_a^2),

$\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \ldots - \phi_p B^p$ adalah polinomial AR(p),

$\theta(B) = 1 - \theta_1 B - \theta_2 B^2 - \ldots - \theta_q B^q$ adalah polinomial MA(q),

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

Untuk suatu d bernilai pecahan, operator *differencing* fraksional $(1-B)^d$ didefinisikan sebagai

$$(1-B)^d = 1 + \sum_{k=1}^{\infty} \frac{\Gamma(-d+k)}{\Gamma(-d)\,k!} B^k . \qquad (2)$$

Jika persamaan $\lambda_k(d) = \dfrac{\Gamma(-d+k)}{\Gamma(-d)\,k!}$ pada persamaan (2) dijabarkan untuk berbagai nilai k maka :

untuk k = 1, diperoleh $\dfrac{\Gamma(-d+1)}{\Gamma(-d)1!} = \dfrac{(-d)!}{(-d-1)!1!} = -d$,

untuk k = 2, diperoleh $\dfrac{\Gamma(-d+2)}{\Gamma(-d)2!} = \dfrac{(-d+1)!}{(-d-1)!2!} = \dfrac{-d(1-d)}{2}$,

untuk k = 3, diperoleh $\dfrac{\Gamma(-d+3)}{\Gamma(-d)3!} = \dfrac{(-d+2)!}{(-d-1)!3!} = \dfrac{-d(1-d)(2-d)}{6}$,

dan seterusnya. Persamaan (2) dapat ditulis kembali menjadi

$$(1-B)^d = 1 + \sum_{k=1}^{\infty} \lambda_k(d) B^k ,$$

dengan $\lambda_0(d) = 1$,

$\lambda_1(d) = -d$,

$\lambda_2(d) = -\frac{1}{2} d(1-d)$,

$\lambda_3(d) = -\frac{1}{6} d(1-d)(2-d)$ dan seterusnya.

Sehingga, persamaan (2) di atas dapat ditulis menjadi

$$(1-B)^d = 1 - dB - \frac{1}{2} d(1-d)B^2 - \frac{1}{6} d(1-d)(2-d)B^3 - \ldots \qquad (3)$$

Persamaan regresi spektral yang dibentuk oleh Geweke dan Porter-Hudak adalah,

$$\ln\left\{I_Z(\omega_j)\right\} = \ln\left\{f_W(0)\right\} + d \ln\left(2\sin(\omega_j/2)\right)^{-2} + \varepsilon , \qquad (4)$$

Nilai parameter d ditentukan dengan menggunakan metode OLS (*Ordinary Least Square*). Velasco (1999b) menyarankan bentuk periodogram yang sama dengan Metode GPH dengan menggantikan

$2\,sin\left(\omega_j\,/\,2\right)$ menjadi j sebagai variabel bebasnya. Penaksiran parameter ϕ dan θ dilakukan dengan menggunakan Metode Maksimum Likelihood.

Setelah nilai koefisien pembeda dari data diketahui berdasarkan Metode GPH dan Metode MGPH dilakukan pembedaan sebesar *d* pada data tersebut, sehingga model dari data mengikuti Model ARMA(*p*,0,*q*). Peramalan dari data dilakukan melalui Model ARMA. Bentuk persamaan peramalan dari Model AR(1) adalah $\hat{Z}_t\left(h\right)=\mu+\phi^h\left(Z_t-\mu\right)$, dengan h adalah nilai periode yang akan diramalankan. Bentuk persamaan peramalan dari Model MA(1) adalah $\hat{Z}_t\left(h\right)=\mu-\theta a_t$ untuk $h = 1$ dan $\hat{Z}_t\left(h\right)=\mu$ untuk $h > 1$.

3. Kajian Simulasi

Simulasi menggunakkan Software R versi 2.7.2, dengan banyaknya data T = 300, 600, 1000 dengan perulangan 1000 kali. Untuk mengaktifkan fasilitas pembeda pecahan (*fractional Difference*) pada Sotware R, Sebelumnya di *install Package fracdiff*. Model ARFIMA yang dibangkitkan mengikuti Model ARFIMA(1,*d*,0) dan Model ARFIMA(0,*d*,1) dengan nilai *d* = 0,6 dan 0,8. Parameter ϕ dan θ masing-masing 0,5, a_t mengikuti Distribusi Normal dengan rata-rata nol dan varians 1. Akurasi penaksiran parameter *d* ditentukan dengan menghitung rata-rata dan standar deviasi dari 1000 nilai *d* untuk Model ARFIMA.

Pada simulasi untuk penentuan akurasi peramalan (Tabel 3), data yang dibangkitkan sebanyak T = 300, 600 dan 1000 masing-masing diambil 10 data terakhir sebagai data *testing*. Langkah-langkah dalam perbandingan peramalan, sebagai berikut :

i. Membangkitkan data Model ARFIMA sebesar T dengan parameter –parameter yang telah ditentukan.
ii. Menentukan nilai parameter pembeda (*d)* pada data Model ARFIMA melalui dua metode yaitu Metode GPH dan Metode MGPH.
iii. Melakukan pembedaan (*differencing*) berdasarkan nilai *d* yang telah ditentukan.
iv. Membagi data menjadi 2 bagian yaitu 10 data terakhir sebagai data *testing* dan T-10 data sebelumnya sebagai data *training*.
v. Menaksir parameter model untuk data *training* dan melakukan peramalan untuk 10 pengamatan kedepan untuk Metode GPH dan Metode MGPH.

vi. Menentukan nilai MSE dari kedua metode penaksiran.

Tabel 1. Akurasi Penaksiran Parameter Pembeda (*d*) dari Data Model ARFIMA Nonstasioner untuk Metode GPH dan Metode MGPH

T	Model ARFIMA	*d*=0,6		*d*=0,8	
		GPH	MGPH	GPH	MPGH
1000	ARFIMA(1,*d*,0)	0,622	0,643	0,823	0,843
	ARFIMA(0,*d*,1)	0,602	0,574	0,795	0,771
600	ARFIMA(1,*d*,0)	0,634	0,662	0,836	0,868
	ARFIMA(0,*d*,1)	0,596	0,555	0,790	0,746
300	ARFIMA(1,*d*,0)	0,651	0,697	0,830	0,891
	ARFIMA(0,*d*,1)	0,584	0,513	0,761	0,705

Tabel 2. Standar Deviasi Akurasi Penaksiran Parameter Pembeda (*d*) dari Data Model ARFIMA Nonstasioner untuk Metode GPH dan Metode MGPH

T	Model ARFIMA	*d*=0,6		*d*=0,8	
		GPH	MGPH	GPH	MPGH
1000	ARFIMA(1,*d*,0)	0,141	0,090	0,148	0,092
	ARFIMA(0,*d*,1)	0,142	0,095	0,138	0,090
600	ARFIMA(1,*d*,0)	0,162	0,110	0,166	0,109
	ARFIMA(0,*d*,1)	0,161	0,112	0,167	0,114
300	ARFIMA(1,*d*,0)	0,205	0,140	0,196	0,134
	ARFIMA(0,*d*,1)	0,209	0,139	0,208	0,141

Berdasarkan hasil simulasi tabel 1 dan tabel 2, akurasi penaksiran parameter Metode GPH mempunyai akurasi yang lebih baik untuk Model ARFIMA(1,*d*,0) dibandingkan dengan Metode MGPH, sedangkan untuk Model ARFIMA(0,*d*,1) Metode MGPH lebih baik dibandingkan dengan Metode GPH. Secara keseluruhan, Metode MGPH mempunyai standar deviasi penaksiran parameter Model ARFIMA yang lebih baik dibandingkan dengan Metode GPH.

Tabel 3. MSE hasil peramalan dari Metode GPH dan MGPH untuk Model ARFIMA Nonstasioner

T	**Model ARFIMA**	***d*=0,6**		***d*=0,8**	
		GPH	**MGPH**	**GPH**	**MPGH**
1000	ARFIMA(1,*d*,0)	1,298	1,268	1,383	1,1340
	ARFIMA(0,*d*,1)	1,288	1,254	1,280	1,251
600	ARFIMA(1,*d*,0)	1,351	1,303	1,347	1,297
	ARFIMA(0,*d*,1)	1,272	1,233	1,266	1,233
300	ARFIMA(1,*d*,0)	1,228	1,162	1,267	1,201
	ARFIMA(0,*d*,1)	1,229	1,176	1,272	1,203

Berdasarkan tabel 3, MSE (*Mean Square Error*) dari Metode MGPH lebih kecil dari Metode GPH baik untuk d = 0,6 maupun untuk d = 0,8. Hasil simulasi ini menunjukkan bahwa peramalan menggunakan Metode MGPH lebih akurat dibandingkan dengan menggunakan Metode MGPH untuk Model ARFIMA(1,*d*,0) dan Model ARFIMA(0,*d*,1).

4. Kesimpulan

Metode GPH mempunyai akurasi penaksiran parameter pembeda yang lebih baik dari Metode GPH untuk Model ARFIMA(1,*d*,0) dan Model ARFIMA(0,*d*,1), tetapi standar deviasinya lebih besar jika dibandingkan dengan Metode MGPH. Akurasi peramalan dari Metode MGPH lebih baik dibandingkan Metode GPH berdasarkan nilai MSE-nya.

DAFTAR PUSTAKA

Beran, J. (1994), "Maximum Likelihood Estimation of the Differencing Parameter for Invertible Short and Long Memory Autoregressive Integrated Moving Average Models", *Journal of the Royal Statistical Society*, Vol. 57, hal. 659-672.

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.

Hurvich, C.M. dan Ray, B.K. (1995), "Estimation of the Memory Parameter for Non stationary or Noninvertible Fractionally Integrated Processes", *Journal of Time series Analysis,* Vol. 16, hal.17-42.

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.

Reisen,V.A dan Lopes,S.R.C. (1999), "Some Simulations and Applications of Forecasting Long-Memory Tome Series Models" , Journal of Statistical and Planning Inference, Vol.80, hal 269-287.

Robinson, P.M. (1995), "Log-Periodogram Regression of Time Series with Long Range Dependence", *Annals of Statistics,* Vol. 23, hal. 1048-1072.

Sowell, F. (1992), "Maximum Likelihood Estimation of Stationary Univariate Fractionally Integrated Time Series Models", *Journal of econometrics,* Vol.53, hal.165 – 188.

Velasco, C. (1999a), "Non-Stationary Log-Periodogram Regression", *Journal of Econometric*, Vol. 91, hal. 325-371.

Velasco, C. (1999b), "Gaussian Semiparametric Estimation of Non-Stationary Time Series", *Journal of Time Series Analysis*, Vol. 20, No.1, hal. 87-127.