 ISBN : 978-979-16353-8-7

T-7

PERAMALAN CURAH HUJAN DENGAN WAVELET

Garini Widosari[1]

[1] Politeknik Negeri Samarinda
[1]garini_72@yahoo.com

Abstrak

Peramalan adalah salah satu unsur yang sangat penting dalam pengambilan keputusan. Peranan peramalan menjelajah ke berbagai bidang, misal fisika, lingkungan, ekonomi, teknik dan kesehatan. Dalam bidang Teknik Sipil, curah hujan mempunyai peranan yang sangat penting dalam perencanaan pembangunan. Untuk itu diperlukan metode untuk meramalkan curah hujan dengan kesalahan yang kecil. Salah satu metode yang digunakan dalam peramalan adalah analisis runtun waktu yang mendasarkan nilai masa kini berbasis oleh nilai – nilai sejenis di masa lalu. Metode analisis runtun waktu yang sering digunakan untuk data musiman adalah SARIMA. Dalam penelitian ini diusulkan metode *wavelet* untuk menganalisis data tersebut. Prosedur pemodelan dengan *wavelet* yaitu data runtun waktu dilakukan transformasi sehingga diperoleh koefisien *detail* dan *smooth.* Dari hasil transformasi ini yang selanjutnya digunakan untuk memodelkan runtun waktu dan digunakan untuk meramalkan satu periode ke depan dengan memandang masing – masing bagian *detail* dan *smooth* sebagai proses autoregresi. Dalam penelitian ini dilakukan metode *wavelet* yang selanjutnya dibandingkan dengan metode SARIMA. Dari hasil analisis data menunjukkan peramalan curah hujan dengan metode *wavelet* lebih baik dibandingkan metode SARIMA yaitu dilihat dari nilai MSE dan MAPE.

Kata kunci: *Wavelet*, SARIMA, MSE, MAPE

PENDAHULUAN

Dalam kehidupan sehari – hari kita sering dihadapkan pada permasalahan yang harus diramalkan terlebih dahulu sebelum diambil keputusan. Seperti di bidang fisika, lingkungan, teknik, ekonomi dan kesehatan. Contohnya adalah untuk merencanakan bangunan di bidang Teknik Sipil seperti jalan raya dan saluran drainase diperlukan data curah hujan serta peramalannya. Salah satu cara meramalkan suatu kejadian di masa yang akan datang adalah dengan analisis runtun waktu.

Dalam analisis runtun waktu , nilai masa kini dipengaruhi oleh nilai – nilai sejenis di masa lalu. Analisis runtun waktu secara umum bertujuan untuk mempelajari atau membuat mekanisme model stokastik yang memberikan reaksi suatu runtun yang diamati dan memprediksi nilai runtun waktu yang akan datang didasarkan pada histori runtun itu sendiri. Komponen data pada runtun waktu terdiri dari trend, siklus dan musiman. Pendekatan dekomposisi runtun waktu merupakan metode untuk memisahkan masing – masing komponen. Sehingga penulisan matematis umum dari pendekatan dekomposisi adalah

$$X_t = f(I_t, T_t, C_t, E_t)$$

Dengan X_t adalah nilai runtun waktu pada periode t

I_t adalah komponen musiman pada periode t

T_t adalah komponen trend pada periode t

C_t adalah komponen siklus pada periode t

E_t adalah komponen galat pada periode t

Wavelet merupakan fungsi dekomposisi dari *wavelet* ayah dan *wavelet* ibu yang masing – masing bagian adalah orthogonal. *Wavelet* ayah mempunyai sifat *smooth* sedangkan *wavelet* ibu mempunyai sifat *detail* yang mengakibatkan data dapat dipisahkan dalam komponen yang berbeda.

Wavelet banyak digunakan di berbagai bidang, seperti pada *signal processing*, kesehatan, kompresi data, analisis numerik, kimia, statistik. Vidakovic(1999) mengungkapkan penggunaan *wavelet* di bidang statistik diantaranya adalah sebagai estimator densitas, analisis runtun waktu dan model Bayes. Sedangkan Abramovich dkk (2000) membahas aplikasi *wavelet* pada regresi non parametrik.

Penelitian ini mengkaji mengenai peranan *wavelet* untuk menganalisis data curah hujan di Samarinda, sekaligus membuat peramalan satu periode ke depan berdasarkan model yang diperoleh. Jenis metode *wavelet* yang digunakan dalam penelitian ini adalah *Maximal Overlap Discrete Wavelet Transform* (MODWT) dengan keluarga *wavelet* Haar. Penggunaan MODWT diusulkan untuk mengatasi keterbatasan dari *Discrete Wavelet Transform* (DWT) yang mensyaratkan N=2^J dengan J bilangan bulat positif, padahal data runtun waktu tidak selalu banyak data berkelipatan 2^J

Definisi 1: **X** *vektor berdimensi N adalah elemen analisis runtun waktu bernilai real* $\{X_t : t = 0,1,\cdots,N-1\}$.

Definisi 2 : *Untuk sebarang bilangan bulat positif Jo merupakan level MODWT dari* **X** *adalah transformasi yang terdiri dari vektor sebanyak* $Jo+1$ *yaitu* $\tilde{\mathbf{W}}_1,\tilde{\mathbf{W}}_2,\cdots,\tilde{\mathbf{W}}_{Jo},\tilde{\mathbf{V}}_{Jo}$ *yang masing – masing dengan dimensi N. Vektor* $\tilde{\mathbf{W}}_j$ *memuat koefisien wavelet MODWT ,sedangkan* $\tilde{\mathbf{V}}_{Jo}$ *memuat koefisien skala MODWT yang didefinisikan sebagai*

$$\tilde{\mathbf{W}}_j = \tilde{\mathcal{W}}_j \mathbf{X} \text{ dan } \tilde{\mathbf{V}}_{Jo} = \tilde{\mathcal{V}}_{J0}\mathbf{X}$$

Dengan $\tilde{\mathcal{W}}_j, \tilde{\mathcal{V}}_{J0}$ *matrik NxN* .

Runtun waktu **X** dapat ditemukan kembali dari MODWT melalui

$$\mathbf{X} = \sum_{j=1}^{Jo} \tilde{\mathcal{W}}_j^T \tilde{\mathbf{W}}_j + \tilde{\mathcal{V}}_{Jo}^T \tilde{\mathbf{V}}_{Jo}$$

$$\equiv \sum_{j=1}^{Jo} \tilde{\mathcal{D}}_j + \tilde{\mathcal{S}}_{Jo}$$

Yang merupakan pendefinisien MODWT dengan dasar analisis multiresolusi (MRA) dari **X** dalam bentuk *detail* dan *smooth* .

Wavelet filter dan skala filter yang digunakan pada MODWT diperoleh dengan modifikasi wavelet filter dan skala filter yang didefinisikan sebagai berikut;

Definisi 3: *Wavelet filter MODWT dan skala filter dibangkitkan dari filter dasar yaitu:*

$$\tilde{h}_{1,l} \equiv \tilde{h}_l \equiv \frac{h_l}{\sqrt{2}} \text{ dan } \tilde{g}_{1,l} \equiv \tilde{g}_l = \frac{g_l}{\sqrt{2}}$$

$$\tilde{g}_{1,l} \equiv \tilde{g}_l \equiv (-1)^{l+1}\, \tilde{h}_{L-1-l}$$

Filter $\{h_{j,l}\}$ menghubungkan $\{\tilde{W}_{j,t}\}$ ke $\{X_t\}$ dibentuk dengan filter sebanyak j, dengan filter ke j yaitu $\{\tilde{h}_{j,l}\}$ adalah $\tilde{h}_0, \underbrace{0,\cdots,0}_{2^{j-1}-1}, \tilde{h}_1, \underbrace{0,\cdots,0}_{2^{j-1}-1}, \cdots, \underbrace{0,\cdots,0}_{2^{j-1}-1}, \tilde{h}_{L-1}$. Sedangkan filter sebelumnya yaitu $\{\tilde{g}_{k,l}\}$ dengan k=1,2,…,j-1

Untuk sampel ukuran N sebarang, didefinisikan level ke j koefisien wavelet MODWT dan koefisien skala MODWT adalah vektor berdimensi N yaitu $\widetilde{\mathbf{W}}_j$ dan $\widetilde{\mathbf{V}}_j$ dengan $\widetilde{\mathbf{W}}_j = \widetilde{\mathcal{W}}_j \mathbf{X}$ dan $\widetilde{\mathbf{V}}_j = \widetilde{\mathcal{V}}_j \mathbf{X}$, sehingga $\widetilde{\mathcal{D}}_j \equiv \widetilde{\mathcal{W}}_j^T \widetilde{\mathbf{W}}_j$ dan $\widetilde{\mathcal{S}}_j \equiv \widetilde{\mathcal{V}}_j^T \widetilde{\mathbf{V}}_j$

Jika didefinisikan $\widetilde{V}_{0,t} = X_t$, persamaan di atas menghasilkan level pertama *koefisien wavelet* MODWT dan *koefisien skala* MODWT. Transformasi MODWT dari $\widetilde{\mathbf{V}}_{j-1}$ ke $\widetilde{\mathbf{W}}_j$ dan $\widetilde{\mathbf{V}}_j$ diperoleh.

$$\begin{bmatrix} \widetilde{\mathbf{W}}_j \\ \widetilde{\mathbf{V}}_j \end{bmatrix} = \widetilde{\mathcal{P}} \widetilde{\mathbf{V}}_{j-1}$$

$$= \begin{bmatrix} \widetilde{\mathcal{B}}_j \\ \widetilde{\mathcal{A}}_j \end{bmatrix} \widetilde{\mathbf{V}}_{j-1}$$

Selanjutnya sintesis $\widetilde{\mathbf{V}}_{j-1}$ dari $\widetilde{\mathbf{W}}_j$ dan $\widetilde{\mathbf{V}}_j$ dapat juga dinyatakan dalam bentuk $\widetilde{\mathcal{B}}_j$ dan $\widetilde{\mathcal{A}}_j$ yaitu $\widetilde{\mathbf{V}}_{j-1} = \widetilde{\mathcal{B}}_j^{\,T} \widetilde{\mathbf{W}}_j + \widetilde{\mathcal{A}}_j^{\,T} \widetilde{\mathbf{V}}_j$.

Mengingat $\widetilde{\mathbf{V}}_0 \equiv \mathbf{X}$, secara rekursif untuk tingkat Jo menghasilkan

$$\mathbf{X} = \widetilde{\mathcal{B}}_1^T \widetilde{\mathbf{W}}_1 + \widetilde{\mathcal{A}}_1^T \widetilde{\mathcal{B}}_1^T \widetilde{\mathbf{W}}_2 + \widetilde{\mathcal{A}}_1^T \widetilde{\mathcal{A}}_2^T \widetilde{\mathcal{B}}_3^T \widetilde{\mathbf{W}}_3 + \cdots +$$
$$\widetilde{\mathcal{A}}_1^T \cdots \widetilde{\mathcal{A}}_{Jo-1}^T \widetilde{\mathcal{B}}_{Jo}^T \widetilde{\mathbf{W}}_{Jo} + \widetilde{\mathcal{A}}_1^T \cdots \widetilde{\mathcal{A}}_{Jo-1}^T \widetilde{\mathcal{A}}_{Jo}^T \widetilde{\mathbf{V}}_{Jo}$$

Suatu sinyal $\mathbf{X} = (X_i : i = 1, 2, \cdots, t)$ stasioner. Suatu proses AR dengan orde p dapat ditulis $X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \cdots + \phi_p X_{t-p} + a_t$. Dalam penggunaan dekomposisi maka proses AR menjadi proses *Multiscale Autoregressive (MAR)* yang diberikan oleh Renaud dkk (2002). Pada proses MAR ini, masing- masing koefisien wavelet dan koefisien skala mengikuti proses *autoregressive*, yang dinyatakan sebagai :

$$\hat{X}_{t+1} = \sum_{j=1}^{J} \sum_{k=1}^{A_j} \hat{a}_{j,k} \tilde{w}_{j,t-2^j(k-1)} + \sum_{k=1}^{A_j} \hat{a}_{J+1,k} \tilde{v}_{J,t-2^J(k-1)}$$

Dengan j adalah level wavelet $(j = 1, 2, \cdots, J)$

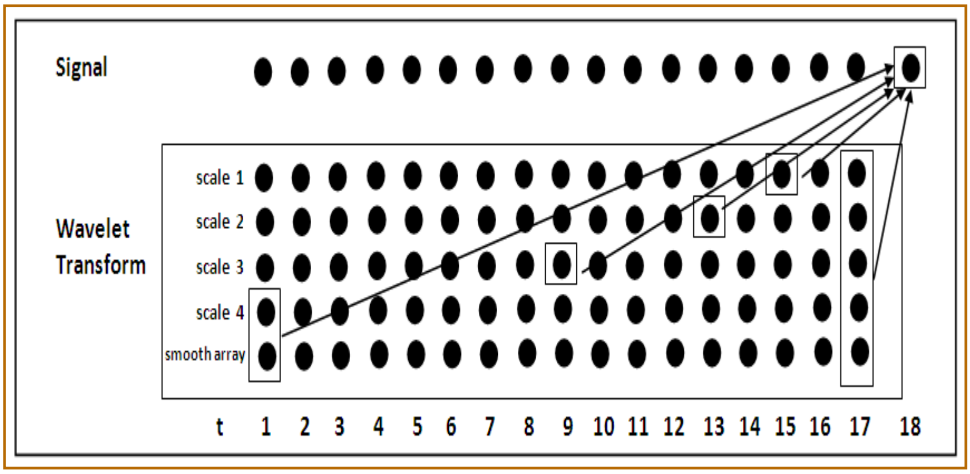
A_j adalah orde MAR $(k = 1, 2, \cdots, A_j)$

$\tilde{w}_{j,t}$ adalah koefisien wavelet dari data

$\tilde{v}_{j,t}$ adalah koefisien skala dari data

$a_{j,k}$ adalah koefisien MAR

Untuk menjelaskan input dan prosedur peramalan data ke t+1 dapat dilihat pada gambar di bawah yang mewakili untuk model dengan level J=4, order MAR $A_j = 2$.



Gambar 1 Model MAR

Berdasarkan gambar tersebut menunjukkan bahwa untuk melakukan peramalan data ke 18 dengan model MAR orde 2, maka variabel input yang digunakan adalah koefisien wavelet level 1 pada t=17 dan t=15, koefisien wavelet level 2 pada t=17 dan t=13, koefisien wavelet level 3 pada t=17 dan t=9, koefisien wavelet level 4 pada t=17 dan t=1 dan koefisien skala level 4 pada t=17 dan t=1.

Metode MAR diusulkan untuk mengikuti proses AR untuk masing – masing skala dari transformasi multiresolusi. Berikut akan diberikan teorema yang menunjukkan apakah model tersebut prosesnya AR, melalui konvergensi prosedur peramalannya terhadap prosedur optimal dan secara asimtotik akan ekivalen ke peramalan terbaik.

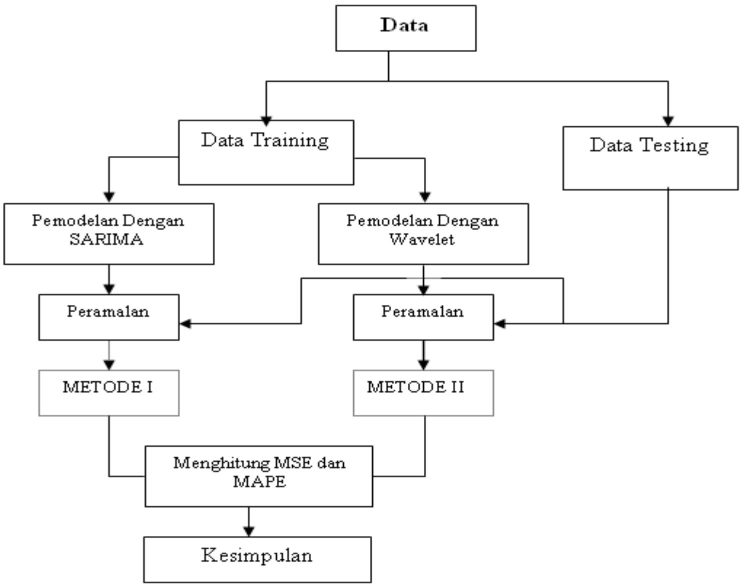
Teorema 1 *Misalkan* $\{X_t, t\in\mathbb{Z}\}$ *mengikuti proses kausal AR dengan orde p dengan parameter* $\phi'=\{\phi_1,\phi_2,\cdots,\phi_p\}$, *misal* $X_t=\phi_1X_{t-1}+\phi_2X_{t-2}+\cdots+\phi_pX_{t-p}+a_t$ *dengan* $\{a_t\}\sim WN(0,\sigma^2)$. *Jika orde* A_j *dipilih pada masing – masing skala lebih besar atau sama dengan* $p/2^j$ *untuk* $j=1,2,\cdots,J$ *maka model multiresolusi yaitu*

$$\hat{X}_{t+1}=\sum_{j=1}^{J}\sum_{k=1}^{A_j}\hat{a}_{j,k}\tilde{w}_{j,t-2^j(k-1)}+\sum_{k=1}^{A_j}\hat{a}_{J+1,k}\tilde{v}_{J,t-2^J(k-1)}$$

, dengan bertambahnya ukuran sampel, $\hat{\alpha}$ *mempunyai sifat asimtotik :* $n^{1/2}\left(\hat{\alpha}-\alpha\right)\Rightarrow N\left(0,\sigma^2\left(R'\mathcal{W}_B{}'\Gamma_B\mathcal{W}_BR_B\right)^{-1}\right)$ *dengan* α *sebanding dengan* $\Omega\phi$, *dan* $\Gamma_B=\left[\gamma\left(t-k\right)\right]_{t,k=1,2,\cdots B}$ *adalah matrik autokovarian dengan* $\gamma\left(l\right)$ *autokovarians dari runtun waktu pada lag l.*

PEMBAHASAN

Metode yang digunakan dalam penelitian ini adalah dengan studi kasus yaitu curah hujan di Samarinda. Metode perhitungan dengan menggunakan metode *Wavelet* yang selanjutnya hasil peramalan dibandingakan dengan metode SARIMA. Langkah – langkah perhitungan ditunjukkan seperti pada gambar di bawah ini



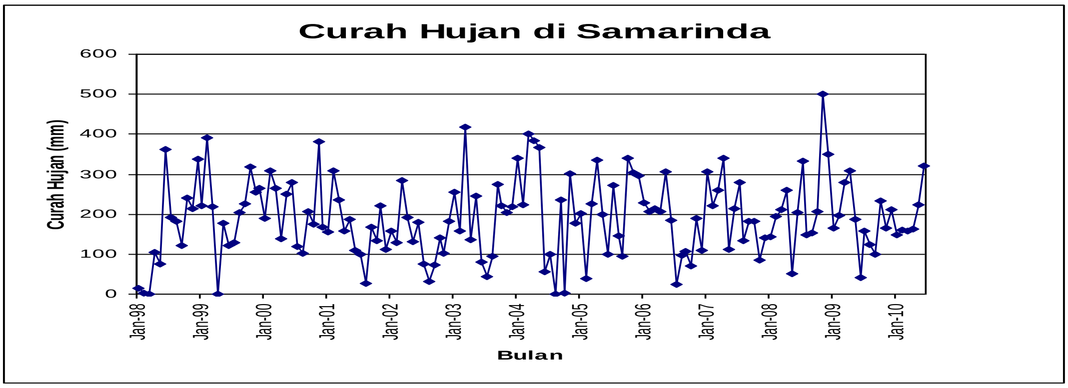
Gambar 2 Diagram alir langkah perhitungan

Identifikasi model dengan *wavelet* digunakan metode MODWT dengan variant *wavelet* Haar order 2 pada level 4, dan selanjutnya dengan menggunakan *Multiscale Autoregressive* (MAR) seperti pada Renaud dkk (2002) ditambah dengan lag musiman

Keakuratan peramalan diukur dengan beberapa nilai ukuran keakuratan yaitu MSE (*Mean Square Error*), dan MAPE (*Mean Absolute Percentage Error* = rata – rata prosentase kesalahan mutlak peramalan) relatif terhadap data asli masing – masing dengan formula sebagai berikut (Wei, 1990).

$$MSE = \left(\frac{1}{M} \sum_{l=1}^{M} a_l^2 \right), \quad MAPE = \left(\frac{1}{M} \sum_{l=1}^{M} \left| \frac{a_l}{X_{n+l}} \right| \right) 100\%$$

Pada penelitian ini data yang digunakan adalah data curah hujan di Samarinda yang diperoleh dari BMKG Temindung Samarinda selama 13 (tiga belas) tahun yaitu mulai Januari 1998 sampai dengan Desember 2010. Untuk membuat model curah hujan di Samarinda digunakan data curah hujan bulanan mulai bulan Januari 1998 sampai dengan bulan Juni 2010 yaitu sebanyak 150 data. Sedangkan untuk data testing digunakan data curah hujan bulan Juli 2010 sampai dengan Desember 2010. Plot data training ditunjukkan pada gambar di bawah ini :



Gambar 3 Plot Curah Hujan di Samarinda

Hasil pengujian serentak dan individu seperti pada Lampiran 4 untuk model wavelet menunjukkan bahwa secara serentak dan individu terdapat minimal satu variabel yang membentuk model. Untuk menentukan variabel yang optimal membentuk model digunakan metode *step wise* . **Tabel 1**. menggambarkan pengujian secara serentak yang dimaksud sedangkan **Tabel 2** merupakan hasil pengujian individu setelah menggunakan metode *step wise*.

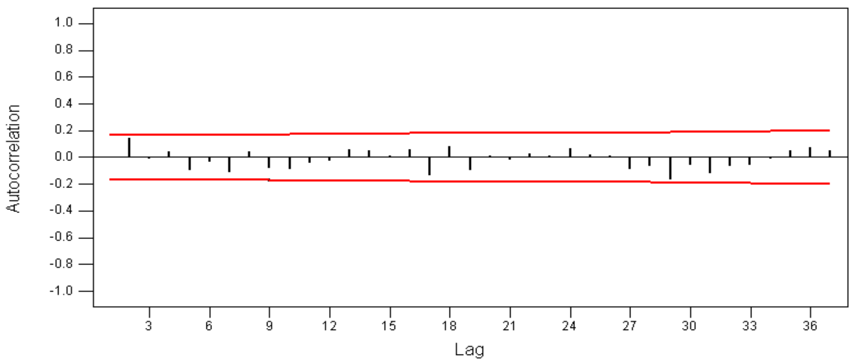
Tabel 1 ANOVA metode wavelet untuk data curah hujan di Samarinda

MAR(1,0)-Haar						
Sumber	db	SS	MS	F	P	Ket
Regression	20	4337723	216886	11.21	0.000	* signifikan pada α =5%
Residual	130	2514513	19342			
Total	150	6852235				

Tabel 2 Hasil pengujian individu metode wavelet untuk data curah hujan di Samarinda setelah menggunakan metode *step wise*.

Prediktor	Koef	T-Hitung	P-Value	Ket
$\tilde{w}_{1,t-1}$	0.7312	-3.47	0.001	* signifikan pada α =5%
$\tilde{w}_{2,t-1}$	0.1550	-2.53	0.013	
$\tilde{v}_{4,t-1}$	0.4112	-2.27	0.026	
$\tilde{v}_{4,t-12}$	0.4312	-2.75	0.007	
$\tilde{w}_{3,t-12}$	0.3465	-3.29	0.001	
$\tilde{v}_{4,t-24}$	0.3482	-2.10	0.039	

Berdasarkan *white noise* residual dapat disimpulkan memenuhi asumsi tersebut karena nilai ACF untuk semua lag berada di dalam garis Bartlet seperti ditunjukkan pada gambar di bawah. Dengan demikian model *wavelet* dapat digunakan untuk membentuk model curah hujan di Samarinda



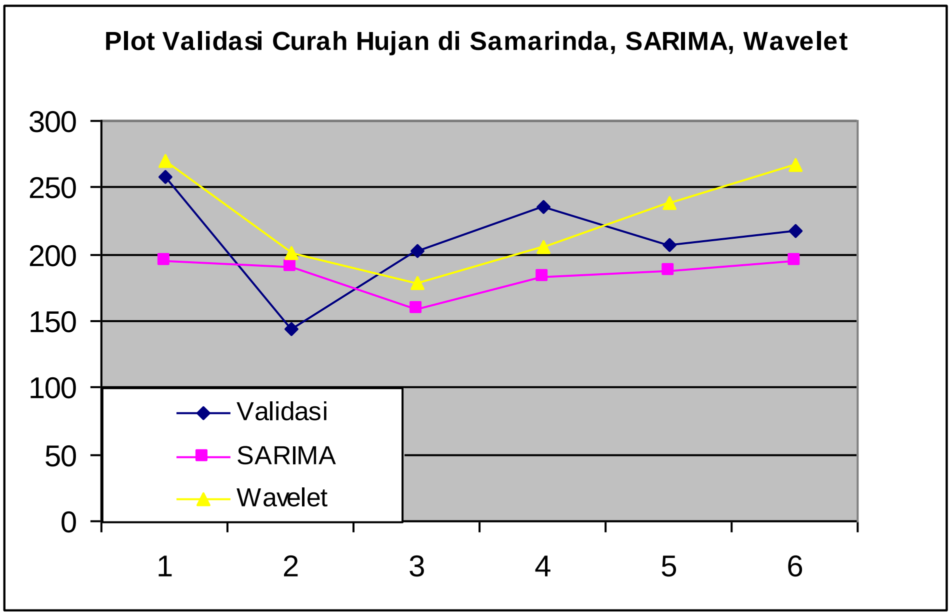
Gambar 4. Plot ACF Residual

Perbandingan hasil peramalan satu langkah ke depan dengan metode SARIMA (Garini , 2011) dan wavelet dirangkum pada pada **Tabel 3** berikut.

Tabel 3 Peramalan satu periode ke depan SARIMA dan wavelet untuk data curah hujan di Samarinda.

t	VALIDASI	Peramalan	
		SARIMA	Wavelet
1	258.7	194.36	270.1046
2	144.1	190.168	201.1052
3	202	159.669	178.1102
4	235.1	182.26	204.8945
5	207.1	187.581	237.9865
6	216.9	194.901	266.5674
	MSE	1951	1381
	MAPE	19%	18%

Secara visual diperlihatkan plot dari data testing, peramalan wavelet dan SARIMA pada Gambar 4.14 di bawah. Dari gambar di bawah menunjukkan peramalan dengan metode wavelet lebih mendekati data validasi dibandingkan metode SARIMA, sedangkan dari Tabel 4.8 di atas menunjukkan bahwa model wavelet mempunyai kesalahan peramalan lebih kecil dibandingkan metode SARIMA berdasarkan kriteria MSE dan MAPE.



Gambar 5 Plot data validasi, SARIMA dan Wavelet data Curah Hujan di Samarinda

KESIMPULAN

Metode wavelet dapat digunakan untuk membuat model curah hujan di Samarinda dengan kesalahan peramalan curah hujan dengan menggunakan metode *wavelet* lebih kecil dibandingkan metode SARIMA berdasarkan kriteria MSE dan MAPE.

Perlu dikembangkan penelitian lanjutan berkaitan dengan *wavelet* untuk membentuk model analisis runtun waktu musiman secara analitis.

DAFTAR PUSTAKA

Anonim. 2011. Data Curah Hujan Samarinda Tahun 1998 – 2010, Samarinda : BMKG Temindung Samarinda

Abramovich, Felix. Bailey, Trevor C and Sapatinas, Theofanis. 2000. *Wavelet analysis and its Statistical Applications.* Royal Statistical Society , 0039-0526/00/49001

Brockwell, Peter J dan Richard A Davis. 1991. *Time Series Theory & Methods*. Springer Verlag . New York.

Widosari, Garini. 2011. *Prediksi Curah Hujan di Samarinda dengan Metode Runtun Waktu*. Jurnal Teknologi Media Perspektif vol 11 no. 01

Percival, Donald B dan Andrew T Walden. 2000. *Wavelet Methods for Time Series Analysis,* Cambridge University Press. America

Renaud, O. ,J.L Starck and F. Murtagh. 2002. *Wavelet-based Forecasting of Short and Long Memory Time Series.* Universite de Geneve. Geneve

Vidakovic, Brani. 1999. *Statistical Modeling by Wavelets*. John Wiley & Sons. Inc. New York.

Wei, William W.S. 1990. *Time Series Analysis*. Addison Wisley. New York.