ISBN: 978-602-14387-0-1

PEMODELAN TIME SERIES DENGAN MAXIMAL OVERLAP DISCRETE WAVELET TRANSFORM

Budi Warsito[1], Subanar[2] dan Abdurakhman[3]
[1)] Jurusan Statistika FSM UNDIP
[2,3)]Jurusan Matematika FMIPA UGM

Abstrak

Penggunaan dekomposisi wavelet untuk pemodelan statistika khususnya pada data time telah mengalami perkembangan yang pesat. Transformasi wavelet yang dipandang lebih sesuai untuk data time series adalah Maximal Overlap Discrete Wavelet Transform (MODWT) karena dalam setiap level dekomposisi terdapat koefisien wavelet dan skala sebanyak panjang data. Kelebihan ini mereduksi kelemahan pemfilteran dengan Discrete Wavelet Transform (DWT) yang tidak dapat dilakukan pada sebarang ukuran sampel. Penentuan level dekomposisi dan koefisien yang digunakan sebagai input model menggunakan dekomposisi multi skala. Pengembangan yang dilakukan pada tulisan ini adalah penyempurnaan pada teknik komputasi sehingga level dekomposisi dan banyaknya koefisien pada setiap level dapat terpilih secara otomatis berdasarkan nilai prediksi yang meminimalkan error.
Kata Kunci: MODWT, time series, dekomposisi

1. Pendahuluan

Analisis dekomposisi wavelet merupakan fungsi basis yang memberikan alat baru sebagai pendekatan yang dapat digunakan dalam merepresentasikan data atau fungsi-fungsi yang lain (Banakar dan Azeem, 2006). Algoritma wavelet mampu memproses data pada skala atau resolusi yang berbeda. Beberapa kajian yang berkaitan dengan transformasi wavelet telah banyak dibahas, diantaranya oleh Khashman dan Dimililer (2008) dan Mallat (1998). Beberapa kajian tentang transformasi wavelet pada data time series juga telah dilakukan, diantaranya oleh Murguia dan Canton (2006) serta Kozlowski (2005). Transformasi Wavelet akan menghasilkan himpunan koefisien Wavelet yang dihitung dari titik (lokasi) observasi pada level (skala) dan lebar range yang berbeda (Kozlowzki, 2005). Penghitungan koefisien wavelet dapat dilakukan dengan Discrete Wavelet Transform (DWT) sebagaimana dikemukakan oleh Mallat (1998) atau Maximal Overlap Discrete Wavelet Transform (MODWT) seperti dalam Percival dan Walden (2000).

Penelitian yang dilakukan oleh Renaud dkk (2003) membahas aplikasi wavelet pada time series dengan menentukan koefisien wavelet yang digunakan untuk prediksi pada skala yang berbeda berdasarkan level dekomposisinya. Transformasi wavelet yang

ISBN: 978-602-14387-0-1

digunakan adalah "*à trous*" *wavelet transform* yang identik dengan MODWT. Prosedur yang diusulkan telah dibuktikan konvergen ke prosedur optimal dan estimator yang dihasilkan juga konvergen ke parameter yang menghasilkan prediksi terbaik jika prosesnya mengikuti proses autoregresif. Pemilihan level dekomposisi dan banyaknya koefisien pada setiap level bersifat opsional. Pada tulisan ini disusun program komputasi untuk melengkapi metode yang diusulkan oleh Renaud dkk (2003) sehingga level dekomposisi dan banyaknya koefisien pada setiap level dapat terpilih secara otomatis berdasarkan nilai prediksi yang meminimalkan error. Prosedur penulisan makalah ini diatur sebagai berikut. Bagian 2 membahas pengertian dasar dekomposisi wavelet, bagian 3 memuat penghitungan koefisien wavelet dengan MODWT, sedangkan bab 4 berisi contoh penerapan pada time series dan bagian 5 penutup.

2. Dekomposisi Wavelet

Fungsi wavelet adalah suatu fungsi matematika yang mempunyai sifat-sifat tertentu diantaranya berosilasi di sekitar nol (seperti fungsi sinus dan cosinus) dan terlokalisasi dalam domain waktu, artinya pada saat nilai domain relatif besar, fungsi wavelet berharga nol. Wavelet merupakan fungsi basis yang dapat digunakan dalam merepresentasikan data atau fungsi-fungsi yang lain. Fungsi Wavelet mempunyai nilai yang berbeda dari nol dalam interval waktu yang relatif pendek. Dalam hal ini wavelet berbeda dengan fungsi normal, ataupun fungsi gelombang seperti sinusoida, yang semuanya ditentukan dalam suatu domain waktu (-1,1). Wavelet dibedakan menjadi dua jenis, yaitu wavelet ayah (ϕ) dan wavelet ibu (ψ) yang mempunyai sifat:

$$\int_{-\infty}^{\infty} \phi(x)\, dx = 1 \text{ dan } \int_{-\infty}^{\infty} \psi(x)\, dx = 0 \quad (1)$$

Keluarga wavelet dihasilkan dari wavelet ayah dan wavelet ibu melalui dilatasi diadik dan translasi integer yaitu :

$$\phi_{j,k}(x) = \left(2^j\right)^{1/2} \phi\left(2^j x - k\right) \quad (2)$$

$$\psi_{j,k}(x) = \left(2^j\right)^{1/2} \psi\left(2^j x - k\right) \quad (3)$$

dengan j dan k masing-masing adalah parameter dilatasi dan parameter translasi.

Wavelet dengan bentuk dilatasi dan translasi dengan j = 0 dan k = 0 dapat dipandang sebagai wavelet dasar. Indeks dilatasi j dan translasi k berpengaruh terhadap perubahan support dan range dari wavelet dasar. Indeks translasi k berpengaruh terhadap pergeseran posisi pada sumbu mendatar tanpa mengubah lebar support sedangkan pada indeks dilatasi j, jika support menyempit maka range akan melebar.

Karena wavelet terlokalisasi dalam domain waktu (artinya pada saat nilai domain relatif besar, fungsi wavelet berharga nol) maka representasi fungsi dengan wavelet menjadi lebih efisien. Hal ini dikarenakan banyaknya koefisien wavelet yang tidak nol dalam rekonstruksi fungsi dengan wavelet relatif sedikit (Suparti dan Subanar, 2005). Selain itu, wavelet juga mampu merepresentasikan fungsi yang bersifat tidak mulus maupun fungsi dengan lonjakan atau volatilitas tinggi. Pada bagian fungsi yang tidak mulus, representasi wavelet akan menggunakan panjang support yang sempit.

Fungsi wavelet dapat dinyatakan sebagai kombinasi linear dari basis yang dibangun oleh wavelet atau dapat dituliskan dalam persamaan berikut.

$$f(x) = \sum_{k\in Z} c_{J,k}\phi_{J,k}(x) + \sum_{j<J}\sum_{k\in Z} d_{j,k}\psi_{j,k}(x) \quad (4)$$

dengan

$$c_{J,k} = \int f(x)\phi_{J,k}(x)dx$$

$$d_{j,k} = \int f(x)\psi_{j,k}(x)dx$$

Transformasi pada persamaan (4) merupakan transformasi wavelet kontinu atau *Continue Wavelet Transform* (CWT) dimana koefisien-koefisien wavelet diperoleh melalui proses integrasi sehingga nilai wavelet harus terdefinisi pada setiap $x \in \mathbb{R}$. Bentuk transformasi yang lain adalah *Discrete Wavelet Transform* (DWT) dimana nilai-nilai wavelet hanya terdefinisi pada titik-titik diskret. Vektor yang memuat nilai-nilai wavelet disebut filter wavelet. Ada dua jenis filter pada DWT yaitu filter wavelet (filter detil) dinotasikan dengan h dan filter skala yang dinotasikan dengan g. Panjang suatu filter dinotasikan dengan L. Suatu filter wavelet harus memenuhi tiga sifat dasar berikut:

$$\sum_{l=0}^{L-1} h_l = 0 \quad (5a)$$

$$\sum_{l=0}^{L-1} h_l^2 = 1 \quad (5b)$$

$$\sum_{l=0}^{L-1} h_l h_{l+2n} = \sum_{l=-\infty}^{\infty} h_l h_{l+2n} = 0 \quad (5c)$$

Jika diberikan filter wavelet $\{h_l\}$ maka filter skala didefinisikan sebagai berikut :

$$g_l \equiv (-1)^{l+1} h_{L-1-l} \quad (6)$$

Filter skala diasumsikan memenuhi kondisi berikut :

$$\sum_{l=0}^{L-1} g_l = \sqrt{2} \quad (7a)$$

$$\sum_{l=0}^{L-1} g_l^2 = 1 \quad (7b)$$

ISBN: 978-602-14387-0-1

$$\sum_{l=-\infty}^{\infty} g_l g_{l+2n} = 0 \text{ dan } \sum_{l=-\infty}^{\infty} g_l h_{l+2n'} = 0 \quad (7c)$$

Syarat yang harus dipenuhi untuk memenuhi sifat-sifat tersebut adalah panjang filter L bernilai genap. Misalkan diberikan filter wavelet h = (h_0, h_1, ..., h_{L-1}) dan $\mathcal{X} = (\mathcal{X}_1, \dots, \mathcal{X}_n)$ adalah nilai fungsi $\mathcal{X}$ pada x_1, ... , x_n. Syarat yang harus dipenuhi adalah $n=2^J$ dengan J suatu bilangan bulat positif. Transformasi wavelet dengan DWT dapat dituliskan sebagai :

$$W = \mathcal{W}\mathcal{X} \quad (8)$$

dengan W = hasil transformasi dengan DWT dan $\mathcal{W}$ = matriks transformasi berukuran nxn. Dalam hal ini elemen-elemen dari vektor W didekomposisi menjadi J+1 sub vektor. Transformasi dengan DWT akan memetakan vektor $\mathcal{X} = (\mathcal{X}_1, \dots, \mathcal{X}_n)$ ke vektor koefisien W = (W_1, W_2, ..., W_J, V_J) dengan W_j, j = 1, 2, ..., J memuat koefisien wavelet $d_{j,k}$ dan V_J memuat koefisien skala $c_{J,k}$. Koefisien wavelet yang bernilai besar mempunyai kontribusi besar dalam rekonstruksi fungsi sedangkan koefisien yang kecil mempunyai kontribusi yang kecil sehingga dapat diabaikan (dianggap nol). Dengan mengabaikan koefisien-koefisien wavelet yang dianggap kecil, transformasi dengan DWT dapat digunakan untuk proses denoising.

3. Maximal Overlap DWT

Pemfilteran dengan DWT sebagaimana pada persamaan (8) tidak dapat dilakukan jika sampel yang diamati berukuran sebarang yang tidak dapat dinyatakan dalam bentuk 2^J dengan J bilangan bulat positif. Sebagai alternatif, penghitungan koefisien $d_{j,k}$ dan $c_{J,k}$ dapat dilakukan dengan *Maximal Overlap Discrete Transform* (MODWT). Keuntungan MODWT adalah dapat mengeliminasi reduksi data menjadi setengahnya (*down-sampling*) sehingga dalam setiap level akan terdapat koefisien wavelet dan skala sebanyak panjang data (Percival dan Walden, 2000). Misalkan data time series dengan panjang N, transformasi MODWT akan memberikan vektor kolom w_1, w_2, ..., w_{J0} dan v_{J0} masing-masing dengan panjang N.

Misalkan dipunyai filter wavelet MODWT $\tilde{h}_l$ dengan $\tilde{h}_l \equiv h_l/\sqrt{2}$ dan filter skala $\tilde{g}_l$ dengan $\tilde{g}_l \equiv g_l/\sqrt{2}$, maka filter wavelet dan filter skala MODWT harus memenuhi kondisi berikut :

$$\sum_{l=0}^{L-1} \tilde{h}_l = 0\,,\ \sum_{l=0}^{L-1} \tilde{h}_l^2 = \frac{1}{2}, \quad \text{dan} \quad \sum_{l=-\infty}^{\infty} \tilde{h}_l \tilde{h}_{l+2n} = 0 \quad (9)$$

ISBN: 978-602-14387-0-1

$$\sum_{l=0}^{L-1} \tilde{g}_l = 1\,,\ \sum_{l=0}^{L-1} \tilde{g}_l^2 = \frac{1}{2}\,,\quad \text{dan}\quad \sum_{l=-\infty}^{\infty} \tilde{g}_l \tilde{g}_{l+2n} = 0 \tag{10}$$

Hubungan antara $\tilde{g}_l$ dan $\tilde{h}_l$ dapat dirumuskan sebagai berikut :

$$\sum_{l=-\infty}^{\infty} \tilde{g}_l \tilde{h}_{l+2n} = 0 \tag{11}$$

Pada MODWT koefisien wavelet pada setiap level selalu sama sehingga lebih sesuai untuk pemodelan pada time series dibandingkan dengan DWT. Prediksi data time series satu langkah ke depan dimodelkan secara linear berdasarkan koefisien wavelet hasil dekomposisi pada waktu-waktu sebelumnya.

Misalkan dipunyai sinyal $X=(X_1, ..., X_t)$. Prediksi satu langkah ke depan dari proses autoregresif order p atau AR(p) dapat dituliskan sebagai $\hat{X}_{t+1} = \sum_{k=1}^{p} \hat{\phi}_k X_{t-(k-1)}$. Pada pemodelan wavelet untuk proses ini, Renaud dkk (2003) dan Murtagh dkk (2004) menyusun prosedur penentuan lag-lag yang menjadi variabel input untuk prediksi multiskala autoregresif. Koefisien wavelet (detil) dan koefisien skala hasil transformasi MODWT yang dianggap mempunyai pengaruh untuk prediksi pada waktu t+1 akan berbentuk $w_{j,t-2^j(k-1)}$ dan $c_{J,t-2^J(k-1)}$ atau dapat dituliskan dalam persamaan (12) berikut :

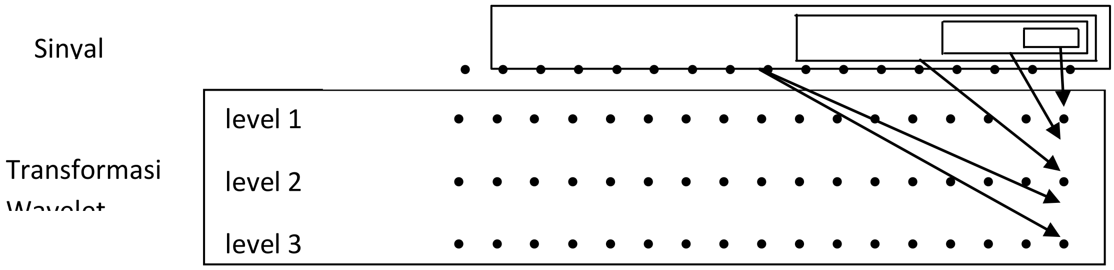
$$\hat{X}_{t+1} = \sum_{j=1}^{J} \sum_{k=1}^{A_j} \hat{a}_{j,k} w_{j,t-2^j(k-1)} + \sum_{k=1}^{A_{J+1}} \hat{a}_{J+1,k}\, c_{J,t-2^J(k-1)} \tag{12}$$

Simbol J menyatakan level dekomposisi sedangkan Aj menjelaskan banyaknya koefisien yang terpilih pada setiap level dekomposisi. Misalkan jika dipilih $A_j = 1$ untuk semua level resolusi j, maka bentuk persamaan (12) akan menjadi :

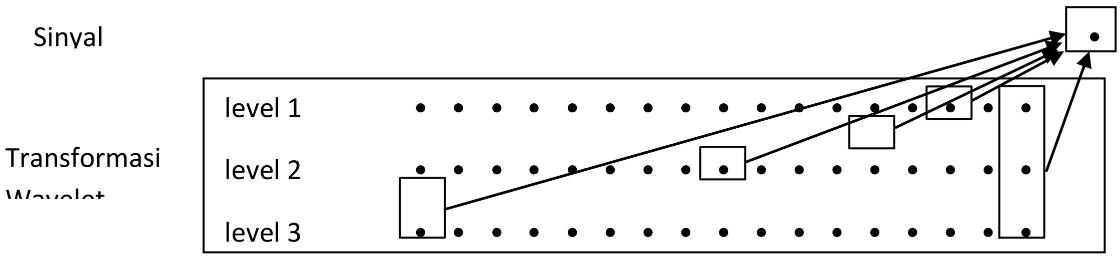
$$\hat{X}_{t+1} = \sum_{j=1}^{J} \hat{a}_j w_{j,t} + \hat{a}_{J+1} c_{J,t} \tag{13}$$

Gambar 1 menunjukkan pixel dari sinyal input yang digunakan untuk menghitung koefisien wavelet terakhir pada skala yang berbeda sedangkan gambar 2 memperlihatkan koefisien wavelet yang digunakan untuk prediksi menggunakan $A_j = 2$ untuk semua level resolusi j, dan J = 4 atau transformasi wavelet dengan lima skala (empat koefisien wavelet + koefisien skala). Pada kasus ini dapat dilihat bahwa hanya terdapat sepuluh koefisien yang digunakan.

PROSIDING SEMINAR NASIONAL STATISTIKA
UNIVERSITAS DIPONEGORO 2013
ISBN: 978-602-14387-0-1



Gambar 1. Pixel dari sinyal input yang digunakan untuk menghitung koefisien wavelet terakhir pada skala yang berbeda



Gambar 2. Koefisien wavelet yang digunakan untuk prediksi data time series

4. Terapan pada Data Time Series

Data yang digunakan dalam makalah ini adalah data kuartalan U.S. Gross Domestic Product (dalam juta US dolar) dari tanggal 1 Januari 1947 sampai 1 Januari 2013 sebanyak 265 data yang bisa diakses secara bebas di website http://www.bea.gov/national/pdf/nipaguid.pdf. Pengolahan data dilakukan dengan menggunakan software Matlab dengan package toolkit wmtsa. Jenis wavelet yang digunakan adalah wavelet Haar. Untuk mendapatkan hasil yang optimal, dipilih level maksimal yang dihitung adalah 4 sedangkan banyaknya koefisien pada setiap level adalah 8. Dasar pertimbangan pemilihan ini adalah bahwa banyaknya input dari model awal ini adalah 40 dianggap sudah cukup besar dan untuk menghindari terlalu banyak data yang terpotong. Level dekomposisi yang terlalu tinggi juga akan meningkatkan nilai residual data hasil transformasi pada level ini, yang dipandang kurang memberikan pengaruh yang besar terhadap data. Dengan mengambil J = 4, banyaknya koefisien pada setiap level adalah $2^4 = 16$, sehingga dengan mengambil nilai maksimal 8 dianggap sudah cukup merepresentasikan semua koefisien yang kemungkinan cukup penting. Hasil penghitungan secara statistik disajikan pada tabel 1, sedangkan gambar 3

ISBN: 978-602-14387-0-1

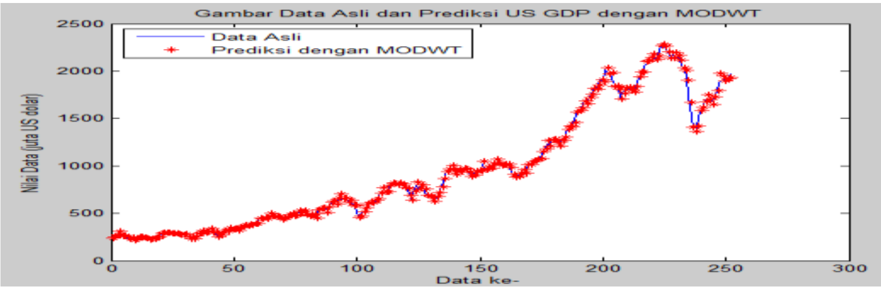
menyajikan plot data asli dan prediksi model menggunakan MODWT dengan level dekomposisi dan banyaknya koefisien pada setiap level dipilih secara optimal.

Tabel 1. Hasil perhitungan statistik nilai prediksi data US GDP dengan MODWT

Kriteria	Nilai
MSE minimal	1.486
Level dekomposisi (J) optimal	1
Banyaknya koefisien pada setiap level (Aj)	7
R square	0,9963
Banyak koefisien	14

Dari tabel 1 nampak bahwa nilai MSE minimal diperoleh pada dekomposisi level 1 dan banyaknya koefisien pada level tersebut adalah 7. Dengan demikian terdapat 14 koefisien yang menjadi input model, masing-masing 7 koefisien wavelet pada dekomposisi level 1 dan 7 koefisien skala. Model yang diperoleh dapat dituliskan dalam persamaan berikut :

$$\hat{X}_{t+1} = \sum_{k=1}^{7} \hat{a}_{1,k} w_{1,t-2(k-1)} + \sum_{k=1}^{7} \hat{a}_{2,k}\, c_{1,t-2(k-1)}$$



Gambar 3. Plot data asli dan prediksi GDP USA menggunakan MODWT dengan level dekomposisi dan banyaknya koefisien pada setiap level optimal

Plot data asli dan prediksi memperlihatkan bahwa secara visual model telah dapat mendekati nilai aslinya, terlihat dari pola data prediksi yang relatif berimpit dengan data asli. Hasil ini juga menunjukkan bahwa transformasi wavelet menggunakan MODWT dengan level dekomposisi dan koefisien pada setiap level yang dipilih secara optimal dapat digunakan untuk prediksi data time series.

5. Penutup

Pada tulisan ini masih terdapat beberapa kekurangan diantaranya adalah banyaknya koefisien yang dipilih pada setiap level dekomposisi masih selalu sama, tidak menutup kemungkinan terdapat koefisien yang kurang bermanfaat dan bisa direduksi untuk lebih mengefisienkan model. Model yang dikembangkan juga masih terbatas pada model linear autoregresif, untuk berikutnya bisa dicoba dengan model-model nonlinear seperti neural network. Selanjutnya juga belum dilakukan studi komparasi dengan model-model yang lain, baik model tradisional seperti ARIMA maupun model nonlinear.

Daftar Pustaka

Khashman, A. and Dimililer, K., 2008, Image Compression using Neural Networks and Haar Wavelet, Wseas Transactions On Signal Processing, ISSN: 1790-5022, 330 Issue 5, Volume 4, May

Kozlowski, B., 2005, Time Series Denoising with Wavelet Transform, Journal of Telecommunications and Information Technology, Warsawa, Polandia

Mallat, S., 1998, A Wavelet Tour of Signal Processing. New York: Academic Press

Murguia, J.S. and Canton, E.C., 2006, Wavelet Analysis of Chaotic Time Series, Revista Mexicana de Fisica 52 (2) 155162

Murtagh, F., Stark, J.L., and Renaud, O., 2004, On Neuro-Wavelet Modelling, Decision Support System, 37, 475-484

Percival, D.,B. and Walden, A.,T., 2000,Wavelet Methods for Time Series Analysis, Cambridge University Press, Cambridge, United Kingdom

Renaud, O., Starcx, J.L., and Murtagh, F., 2003, Prediction Based on a Multiscale Decomposition, Int. Journal of Wavelets, Multiresolution and Information Processing, Vol. 1., No. 2, pp 217-232

Suparti dan Subanar, H., Estimasi Regresi dengan Metode Wavelet Shrinkage, Jurnal Sains dan Matematika, 2000, 8/3:105-113

Lampiran 1 Program Utama

```
function modwttimeseriesloop(JJ,Ajj)
data=xlsread('D:\SEKOLAHhh\DATA excel\GPDIC96.xlsx');
min_mse=9999;
for J=1:JJ
    for Aj=1:Ajj
       w=modwt_filter('haar');
       detil=[];
       skala=[];
       for j=1:J
        [m,n] = modwtj(data,w.h,w.g,j);
        detil=[detil,m];
        skala=[skala,n];
       end
       cdetil=[];
       for k=1:Aj
         for i=1:k
         c=detil(1+(2^J)*(Aj-i):length(data)-1-(2^J)*(i-1),J);
         end
         cdetil=[cdetil,c];
       end
       sdetil=size(cdetil);
       inputdetil=cdetil;
       input=cdetil;
       wskala=[];
       for j=1:J
         for k=1:Aj
           for i=1:k
             w=skala(1+(2^J)*(Aj-i):length(data)-1-(2^J)*(i-1),j);
           end
           wskala=[wskala,w];
         end
       end
       inputskala=wskala;
       inputall=[inputdetil,inputskala];
       t=data(1+(2^J)*(Aj-1)+1:length(data));
       rs=regstats(t,inputall,'linear');
       m_se=rs.mse;
       R_square=rs.rsquare;
       y_asli=t;
       y_hat=rs.yhat;
       if m_se < min_mse
           min_mse=m_se;
           J_nya=J;
           Aj_nya=Aj;
           R_s=R_square;
           y_topi=y_hat;
           yy=y_asli;
       end
    end
end
mse_minimal = min_mse
J_optimal = J_nya
Aj_optimal = Aj_nya
Rsquare_optimal = R_s
Banyak_koefisien = Aj_optimal*(J_optimal+1)
ytopi=y_topi; yasli=yy;
x=1:length(yasli);
figure(1)
plot(x,yasli,'b-',x,ytopi,'r*')
title('Gambar Data Asli dan Prediksi US GDP dengan MODWT')
xlabel('Data ke-')
ylabel('Nilai Data (juta US dolar)')
legend('Data Asli','Prediksi dengan MODWT','Location','NorthWest')
```