MEMISAHKAN DATA DARI *WHITE NOISE* MENGGUNAKAN *WAVELET SHRIKAGE*

Agus Sukmana*)

ABSTRACT

This paper deals with a methodfor cleaning up time series data contaminated with the white noise by using wavelet shrinkage. First, data are transformeded to the wavelet domain then data are discriminated using soft-thresholding to the wavelet coefficients to get wavelet shrinkage coefficients. These wavelet shrinkage coefficients are retransform to the time domain and the result is expected to be free of the white noise. We conducted some experiments [using 2 sets of data] and the results were satisfactory.

I. Pendahuluan

Gagasan penulisan ini berawal ketika proyek model matematika kompresi data dikerjakan di subbagian sistem informasi distribusi GW Amsterdam. Disana ditemukan sederetan data yang dikumpulkan setiap tiga detik dengan tujuan ingin memonitor suatu- proses yang kontinu secermat mungkin. Meskipun pencatatan dilakukan secara otomatis di lapangan dan hasil pengnkuran ditransmisikan melalui alat komunikasi yang memadai ke pusat komputer, ternyata dipastikan selalu ditemui fluktuasi kecil dengan amplitudo beragam yang kemunculannya acak. Deret fluktuasi kecil tersebut diduga adalah white noise, karena karakteristiknya serupa dengan keterangan berikut:

---

*) Dosen tetap Fakultas MIPA jurusan Matematika UNPAR

White noise is aperiodic sound (that is, its wave pattern is not uniform). Its constituent frequencies are of random amplitude and occur at random intervals [Encyclopedia Britannica]

Proses penghilangan white noise dilakukan sebagai berikut. Pertama data (*data times series)* ditransformasikan kedalam domain wavelet menjadi koefisien-koefisien wavelet, dengan tujuan agar mudah dilakukan analisis terhadap frekuensi, kemudian dilakukan pemilihan untuk dipisahkan mana yang merupakan *white noise* dan mana yang bukan. Koefisien wavelet yang kecil mewakili white noise sehingga harus dihilangkan; untuk menentukan koefisien mana yang harus dihilangkan digunakan metoda *soft-thresholding*. Koefisien wavelet yang tersisa disebut koefisien wavelet shrinkage, koefisien ini kembali ditransformasikan ke domain waktu dan hasilnya diharapkan sudah terbebas dari white noise. Pada tulisan ini akan ditunjukkan 2 hasil percobaan yang menunjukkan bahwa dengan wavelet shrinkage data dapat "dimurnikan' dari white noise.

Pembahasan dimulai dengan mengulas singkat tentang wavelet dan komputasinya, lalu metoda thresholding untuk memperoleh koefisien wavelet shrinkage. Dengan menggunakan dua himpunan data akan dibahas hasil yang diperoleh dari pemisahan data dari white noise.

II. Wavelet

Wavelet adalah hasil pengembangan baru dalam matematika terapan, yang merupakan hasil sintesis dari ilmu rekayasa, fisika, dan matematika murni. Gagasan mengenai transformasi wavelet adalah menghampiri suatu fungsi dengan menggunakan kombinasi beberapa fungsi lainnya. Gagasan ini bukanlah hal yang baru, karena pada tahun 1800-an *Jean Baptiste Joseph* ***Fourier*** telah melakukan hal serupa yaitu menyatakan suatu fungsi dengan superposisi fungsi sinus dan kosinus. Wavelet tidak menggunakan fungsi sinus dan kosinus, tetapi menggunakan fungsi yang memiliki sifat *compact support* sehingga memiliki keunggulan dalam menghampiri fungsi yang berfluktuasi tajam (*sharp spikes*) [2], dan fungsi non stasioner pada data *time series* [6]. Pada subbagian ini akan dibahas: basis wavelet dan cara mengkonstruksinya, cara menghampiri sebuah fungsi melalui analisis multi resolusi, teknik komputasi yang digunakan.

A. Basis wavelet

Basis wavelet diperoleh dari sebuah fungsi ψ, yang disebut *mother-wavelet*, dan fungsi skala ϕ yang disebut *father-wavelet*, melalui *translasi* dan *dilatasi* [2]. Sebuah fungsi ψ merupakan *mother-wavelet* jika memenuhi tiga kondisi berikut [5],[6]:

(1). $\int_{-\infty}^{\infty} \psi(t)dt = 0$

(2). $\int_{-\infty}^{\infty} |\psi(t)|dt < \infty$

(3). $\int_{-\infty}^{\infty} \frac{|\hat{\psi}(\omega)|^2}{\omega} d\omega < \infty$

dengan $\psi(\omega) = \int_{-\infty}^{\infty} \psi(t)e^{-i\omega t}dt$ adalah *tranformasi Fourier* dari fungsi $\psi(t)$.

Jika diberikan sebuah mother wavelet $\psi(t)$, untuk setiap bilangan real a, b $(a \neq 0)$, maka dapat diperoleh barisan wavelet melalui translasi dan dilatasi $\psi(t)$,

$$\psi_{a,b}(t) = |a|^{-1/2} \psi\left(\tfrac{t-b}{a}\right)$$

dengan a merupakan parameter dilatasi dan b parameter translasi.

Wavelet yang paling sederhana dan paling tua (1909) adalah wavelet Haar. Basis Haar merupakan suatu basis ortonormal untuk $L^2(\Re)$. Contoh wavelet yang lain dapat dilihat pada [1],[2],[3].

B.Analisis multi resolusi

Wavelet dapat diperoleh melalui *Analisis Multi Resolusi* (AMR) yang diperkenalkan oleh Mallat (1986). Penjelasan lebih rinci dan bukti matematisnya dapat dilihat pada [2],[9]. Dengan menggunakan AMR, fungsi f dapat dituliskan sebagai kombinasi linear dari wavelet dan fungsi skala ;

$$f(t) = \sum_{i \in Z} \beta(l)\phi_l(t) + \sum_{j=0}^{J-1} \sum_{k \in Z} \alpha(j,k)\psi_{j,k}(t),$$

koefisien skala dan wavelet merupakan hasil kali dalam: $\alpha(j,k) = \langle \psi_{j,k}, f \rangle$ dan $\beta(l) = \langle \phi_l, f \rangle$.

C. *Fast Wavelet Transform* (FWT)

FWT merupakan teknik yang dikembangakan dari *Fast Fourier Transform* (FFT). Keuntungan dari FWT adalah cacah operasi yang dibutuhkan untuk menghitung koefisien wavelet lebih kecil dibandingkan dengan transformasi Fourier 'biasa' dan FFT. Sebagai ilustrasi untuk data sebanyak $n = 2^J$: dengan transformasi fourier dibutuhkan n^2 operasi, dengan FFT dibutuhkan $n \log_2 n$, sedangkan dengan FWT cukup n operasi saja. Sehingga proses komputasi dapat dihemat secara berarti. Penjelasan tentang algoritma ini, lihat [8].

III. Wavelet shrinkage

Ide dari wavelet shrinkage adalah menciutkan banyaknya koefisien wavelet *tak-nol* dengan menggunakan metoda *thresholding* [3],[7],[8]. *Threshold* dipilih sedemikian sehingga penciutan tidak mempengaruhi data secara berarti. Ada dua jenis threshold: yang pertama ditentukan secara subyektif oleh pemakai (*hard threshold*), yang kedua dengan memperhitungkan perilaku data (*soft threshold*). Pada tulisan ini hanya akan dibahas metoda *soft thresholding*.

Metoda *Soft-thresholding* secara matematis dituliskan dalam bentuk:

$$\eta_t(y) = sign(y)\max(0,|y| - t_n)$$

tujuannya adalah mengeliminasi *white noise* dengan level σ [3].

Threshold dipilih sbb:

$$t_n = \sqrt{2\log n}.\varepsilon$$

dengan $\varepsilon = \frac{\gamma}{\sqrt{n}}\sigma$.

nilai ε ditaksir dengan $\hat{\varepsilon} = \frac{\hat{\sigma}}{0,6745}$,

$\hat{\sigma} = median(|y_{n-1,k}|;0 \le k < 2^J - 1)$.

Metoda soft thresholding yang lain yang lebih kompleks antara lain: metoda SURE (***Stein Unbiased Estimated of Risk***) dapat dilihat pada [3][8].

IV. Data

Dilakukan dua percobaan untuk melihat kemampuan metoda wavelet shrinkage untuk memisahkan white noise dari data. Percobaan pertama dilakukan terhadap data buatan, sedangkan percobaan kedua dilakukan pada data nyata.

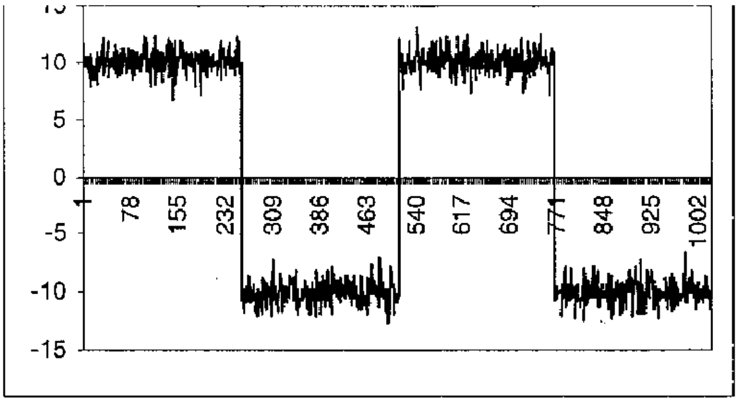
A. Himpunan data 1

Himpunan data 1 merupakan data buatan sebanyak 1024 data, sampel yang diambil dari suatu proses:

$$f(x) = \begin{cases} 1 & \text{jika } 0 \le x \le 265 \text{ atau } 513 \le x \le 768 \\ -1 & \text{jika } 266 \le x \le 512 \text{ atau } 761 \le x \le 1024 \end{cases}$$

Himpunan data itu kemudian diganggu dengan 1024 data, sampel acak yang diambil dari populasi berdistribusi normal dengan rataan (*mean*) 0 dan variansi 1. (*white noise level 1*)

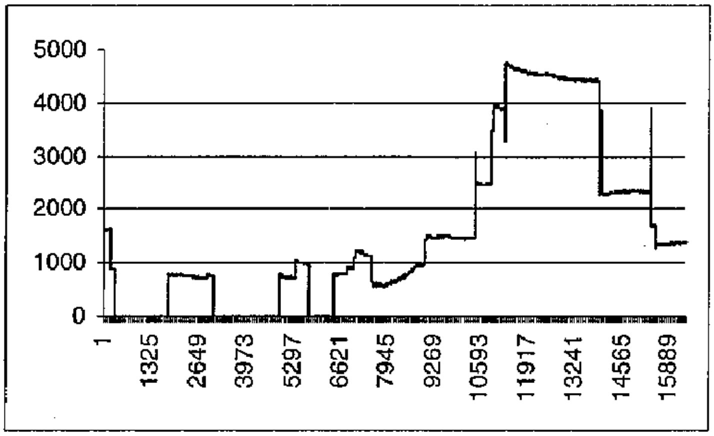
Grafik dari sampel tersebut dapat dilihat pada gambar 1.



Gambar 1

B. Himpunan data 2

Himpunan data 2 diambil dari [8], yang merupakan hasil pengukuran terhadap aliran air yang keluar dari *Haarlemmeerweg Reservoir* milik Amsterdam Water Supply, dilakukan setiap tiga detik untuk16384 kali pengamatan.

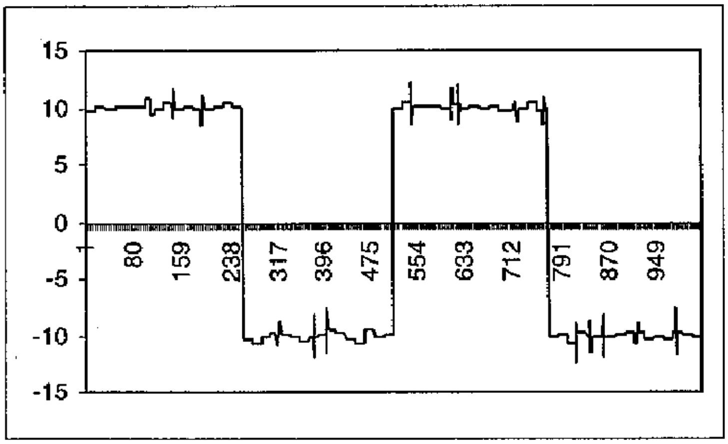


Gambar 2

V. Analisis data

Komputasi dilakukan dengan menggunakan program *wavelett* ver 1.3 yang dibuat oleh penulis bersama Vincent Jansen. Basis yang dipilih adalah basis Haar karena paling sesuai dengan karakteristik data dan resolusi yang digunakan adalah resolusi medium (L=7 , [8]).

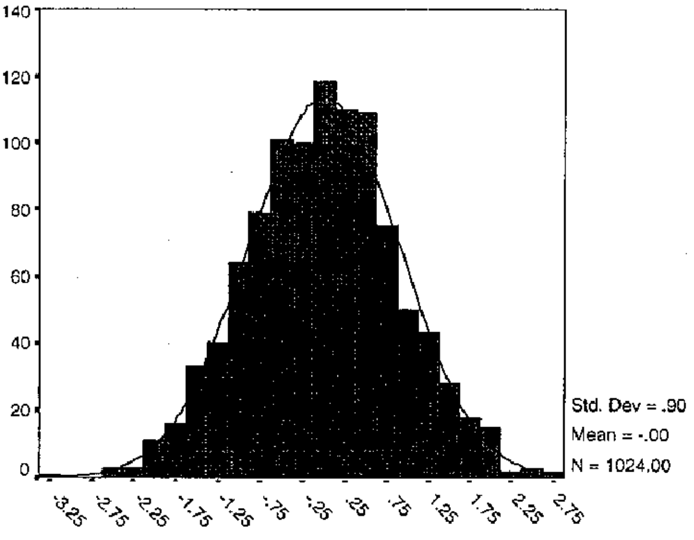
Penciutan wavelet memisahkan Himpunan data 1 menjadi dua kelompok data: yaitu data hasil "pembersihan" dari *white noise* (gambar 3) ,dan data yang dianggap sebagai *white noise* oleh *soft thresholding*. (distribusi data dapat dilihat pada gambar 4).



Gambar 3

Gambar 3 seharusnya memiliki bentuk seperti fungsi $f(x)$, meskipun masih ada penyimpangan tetapi 'kesan umum' fungsi f 'tertangkap' pada Gambar 3.

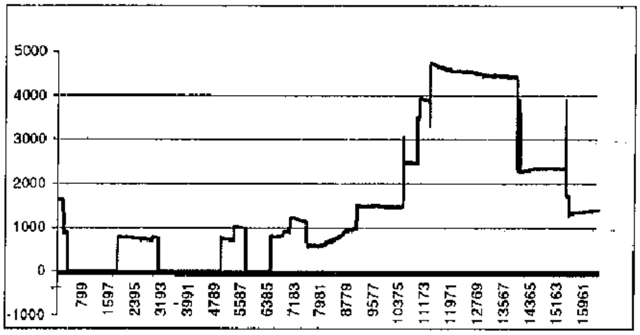
Gangguan yang kita buat untuk himpunan data 1 memiliki distribusi $N(0,1)$, jika *soft thresholding* dapat mengenalinya maka data yang dianggap *white noise* akan memiliki distribusi $N(0,1)$. Secara eksploratif dapat kita lihat dari histogram Gambar 4 bahwa kita memiliki cukup alasan untuk mengatakan data tsb hampir berdistribusi $N(0,1)$, tentu saja secara inferensi (menggunakan *metoda goodness of fit*) dapat dilakukan pengujian secara statistika.



Gambar 4

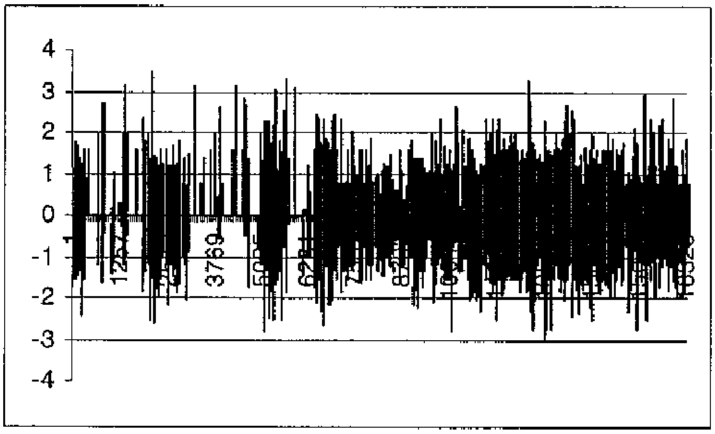
Dari percobaan pertama menunjukan bahwa untuk himpunan data 1, wavelet shrinkage dengan soft thresholding memiliki kemampuan untuk mendeteksi white noise.

Percobaan kedua melakukan hal yang serupa dengan percobaan 1, tetapi mengunakan data yang berbeda (Hmpunan Data 2). Gambar 5 merupakan data yang 'bebas *white noise*' menurut *soft tresholding*, memiliki bentuk umum serupa dengan Gambar 2 tetapi telah dibebaskan dari gangguan yang tampak pada Gambar 6.



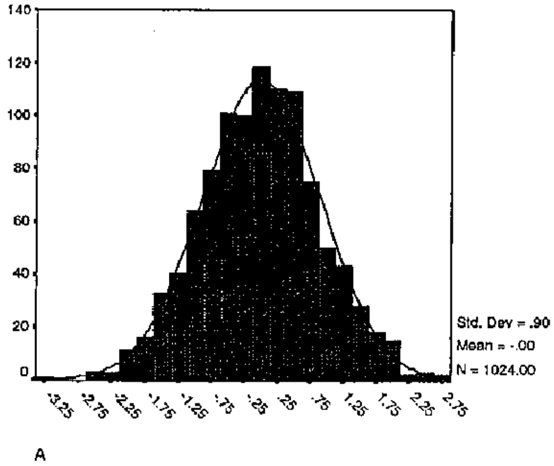
Gambar 5

Gangguan pada Gambar 6 memiliki distribusi seperti ditunjukkan oleh histogram pada Gambar 7.



Gambar 6

Dari pemeriksaan Gambar 6 dan 7, kita memiliki cukup alasan bahwa gangguan bersifat acak dan memiliki distribusi mendekati normal dengan rataan 0 dengan variansi 1. Dengan kata lain gangguan tsb merupakan *white noise.*



Gambar 7

VI. Kesimpulan

Dari hasil pembahasan diatas, wavelet shrinkage dengan *soft-thresholding* memiliki kemampuan cukup baik untuk memisahkan data dari white noise. Manfaat yang dapat diperoleh adalah bahwa dengan metoda tsb kita dapat 'memurnikan data' sebelum melakukan analisis lebih lanjut terhadap data tersebut.

Daftar Referensi

[1]. Benedetto, J.J and Frazier, M.W. (1994), *Wavelets: Mathematics and Applications,* CRC Press, Boca Raton,.

[2]. Daubechies, I. (1992), *Ten Lectures on Wavelet,* SIAM, Philadelphia.

[3]. Donoho, D.L. and Johnstone, I.M. (1995) Adapting to Unknown Smoothness via Wavelet Shrinkage, *Journal of the American Statistical Association,* vol. 90, p. 1200-1224.

[4]. Graps, Amara, (1995) An Introduction to Wavelets, *IEEE Computational Science and Engineering*, vol. 2, no. 2.

[5]. Morretin, P.A. (1997), Wavelets in Statistics, *Resenhas IME-USP,* vol.3, no.2, pp 211-272.

[6]. Morretin, P.A and Chiann, C., (1998) A Wavelet Analysis for Time Series, *Nonparametric Statistics,* vol.10, pp 1-46.

[7]. Sukmana, A. (1999), *Modeling Time Series Data via Wavelet Shrinkage,* GMU-SEAMS International Conference on Mathematics and Its Applications.

[8]. Sukmana, A. (1999), *Wavelet Shrinkage for modeling Time Series Data: a Data Compression Method,* MSc. Thesis Faculty of Mathematics Sciences, Twente University.

[9]. Sondjaja, P. D., (2000), Konstruksi Wavelet dengan Menggunakan Analisis Multi Resolusi, Integral, vol.5, no. 1, hlm 17-23.