

# ESTIMASI REGRESI NON PARAMETRIK DENGAN METODE WAVELET SHRINKAGE NEURAL NETWORK PADA MODEL RANCANGAN TETAP

Hasbi Yasin

Staf Pengajar Program Studi Statistika FMIPA UNDIP

## Abstract

If  $X$  is a predictor variable and  $Y$  is a response variable of following model  $Y = g(X) + e$  with function  $g$  is a regression which not yet been known and  $e$  is an independent random variable with mean 0 and variant  $\sigma^2$ . The function of  $g$  can be estimated by parametric and nonparametric approach. In this paper,  $g$  is estimated by nonparametric approach that is named wavelet shrinkage neural network method. At this method, the smoothly function estimation is depending on shrinkage parameter's that are threshold value and level of wavelet that be used. It also depending on the number of neuron in the hidden layer and the number of epoch that be used in feed forward neural network. Therefore, it is required to be select the optimal value of threshold, level of wavelet, the number of neuron and the number of epoch to determine optimal function estimation.

**Keywords:** Nonparametric Regression, Wavelet Shrinkage Neural Network

## 1. Pendahuluan

Diketahui  $X$  variabel prediktor,  $Y$  variabel respon dalam analisis regresi dan  $\{(X_i, Y_i)\}_{i=1}^n$  adalah data observasi independen dengan model  $Y_i = g(X_i) + \varepsilon_i$  (1) dengan  $i = 1, 2, \dots, n$ , sementara  $g$  adalah fungsi regresi yang tidak diketahui, dan  $\varepsilon_i$  adalah variabel random independen dengan mean 0 dan varian  $\sigma^2$ .

Berdasarkan variabel prediktornya, maka model rancangannya dibedakan dalam dua kasus. Jika variabel prediktornya merupakan variabel random maka rancangannya dinamakan model rancangan stokastik (model random) tetapi jika variabel prediktornya merupakan variabel tetap maka model rancangannya dinamakan model non stokastik (model tetap)<sup>[1,2]</sup>. Sedangkan berdasarkan bentuk fungsinya maka ada dua pendekatan yang dapat dilakukan dalam mengestimasi fungsi regresi  $g$ , yakni pendekatan parametrik dan pendekatan non parametrik. Pendekatan parametrik dilakukan jika asumsi bentuk  $g$  diketahui, sedangkan pendekatan non parametrik menghubungkan variabel prediktor  $X$  terhadap variabel respon  $Y$  tanpa diketahui model dari fungsi  $g$ . Dalam hal ini, diasumsikan bahwa fungsi  $g$  termuat dalam kelas fungsi mulus; artinya mempunyai turunan yang kontinyu atau dapat diintegrasikan secara kuadrat.

Untuk pendekatan non parametrik, teknik yang sudah populer diterapkan adalah teknik pemulus deret orthogonal seperti deret Fourier. Dari estimator deret Fourier kemudian dikembangkan menjadi estimator wavelet dengan mengganti basis-basis fungsi konstan, sinus dan cosinus pada estimator deret Fourier dengan basis-basis fungsi wavelet<sup>[3,8]</sup>. Beberapa fungsi wavelet yang digunakan sebagai basis dalam estimator wavelet antara lain wavelet Haar, Daublet, Symmlet, dan Coiflet<sup>[5,11]</sup>. Estimator wavelet dibedakan menjadi 2 yaitu estimator wavelet linier dan estimator wavelet shrinkage<sup>[5]</sup>. Kemudian dikembangkan metode baru yang merupakan gabungan dari dekomposisi Wavelet sebagai metode ekstraksi untuk mendapatkan informasi awal berkaitan dengan input model dan Neural Network yang digunakan sebagai proses pelatihan data untuk

mendapatkan output model. Metode ini dikenal dengan WNN (Wavelet Neural Network). Tulisan ini membahas penentuan estimator regresi non parametrik dengan metode wavelet shrinkage neural network (WSNN) pada model rancangan tetap disertai dengan contoh simulasinya. Penelitian ini bertujuan untuk membangun sebuah model baru yang lebih baik yang merupakan kombinasi dari model feed forward neural network dan model wavelet shrinkage.

## 2. Bahan dan Metode

Penelitian ini merupakan kajian literatur dari beberapa model yang telah ada yang kemudian dikembangkan dengan simulasi menggunakan software Matlab 7.1.

### 2.1 Fungsi 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. Fungsi wavelet dibedakan atas dua jenis, yaitu wavelet ayah ( $\phi$ ) dan wavelet ibu ( $\psi$ )

yang mempunyai sifat  $\int_{-\infty}^{\infty} \phi(x)dx = 1$  dan  $\int_{-\infty}^{\infty} \psi(x)dx = 0$ . Dengan dilatasi diadik dan translasi integer, wavelet ayah dan wavelet ibu melahirkan keluarga wavelet yaitu  $\phi_{j,k}(x) = (p2^j)^{1/2} \phi(p2^j x - k)$  dan  $\psi_{j,k}(x) = (p2^j)^{1/2} \psi(p2^j x - k)$  untuk suatu skalar  $p > 0$ , dan tanpa mengurangi keumuman dapat diambil  $p=1$ , sehingga  $\phi_{j,k}(x) = 2^{j/2} \phi(2^j x - k)$  dan  $\psi_{j,k}(x) = 2^{j/2} \psi(2^j x - k)$ . Fungsi  $\phi_{j,k}(x)$  dan  $\psi_{j,k}(x)$

mempunyai sifat  $\int_{-\infty}^{\infty} \phi_{j,k}(x) \phi_{j,k'}(x) dx = \delta_{k,k'}$ ,  $\int_{-\infty}^{\infty} \psi_{j,k}(x) \phi_{j,k'}(x) dx = 0$ , dan

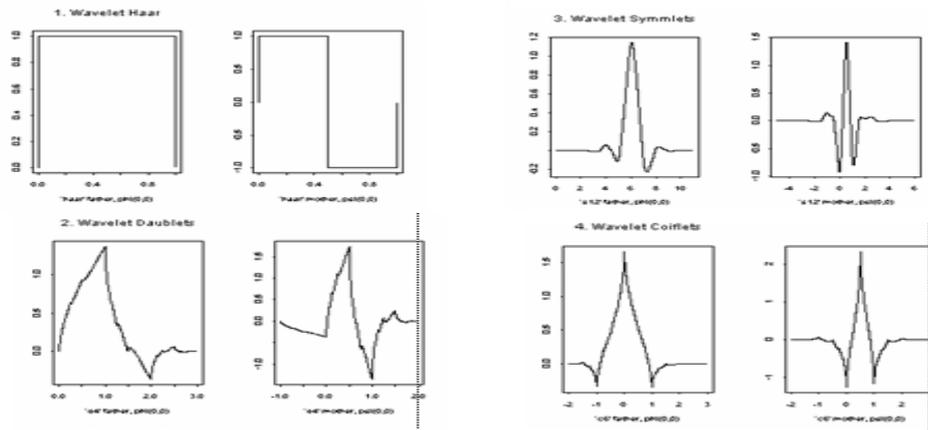
$$\int_{-\infty}^{\infty} \psi_{j,k}(x) \psi_{j,k'}(x) dx = \delta_{j,j'} \delta_{k,k'}$$

dengan  $\delta_{i,j} = \begin{cases} 1 & \text{jika } i = j \\ 0 & \text{jika } i \neq j. \end{cases}$

Contoh wavelet paling sederhana adalah wavelet Haar yang mempunyai rumus

$$\psi(x) = \begin{cases} 1 & , 0 \leq x < 1/2 \\ -1 & , 1/2 \leq x < 1 \\ 0 & , x \text{ yang lain} \end{cases} \text{ dan } \phi(x) = \begin{cases} 1 & , 0 \leq x < 1 \\ 0 & , x \text{ yang lain.} \end{cases} \quad (2)$$

Beberapa contoh wavelet selain wavelet haar diantaranya adalah wavelet Daubechies (Daublet), symmetris (Symmlet), dan Coifman (Coiflet)<sup>[5]</sup>. Visualisasi beberapa wavelet dapat ditunjukkan pada Gambar 1 berikut:



**Gambar 1.** Visualisasi beberapa Wavelet

Jika  $g \in L^2(\mathbb{R})$  maka  $g$  dapat dinyatakan dalam deret wavelet ortonormal

$$g(x) = \sum_{k \in \mathbb{Z}} c_{j_0,k} \phi_{j_0,k}(x) + \sum_{j \geq j_0} \sum_{k \in \mathbb{Z}} d_{j,k} \psi_{j,k}(x) \quad (3)$$

dengan  $c_{j_0,k} = \langle g, \phi_{j_0,k} \rangle = \int_{-\infty}^{\infty} g(x) \phi_{j_0,k}(x) dx$  dan  $d_{j,k} = \langle g, \psi_{j,k} \rangle = \int_{-\infty}^{\infty} g(x) \psi_{j,k}(x) dx$ .

Dengan mengambil  $J$  cukup besar maka deret wavelet (3) dapat didekati oleh

$$g_J(x) = \sum_{k \in \mathbb{Z}} c_{j_0,k} \phi_{j_0,k}(x) + \sum_{j \geq j_0} \sum_{k \in \mathbb{Z}} d_{j,k} \psi_{j,k}(x)$$

$$g_J(x) = \sum_{k \in \mathbb{Z}} c_{J,k} \phi_{J,k}(x) \quad (4)$$

dimana  $g_J$  merupakan pendekatan  $g$  pada level ke  $J$ .

### 2.2 Estimator Wavelet Linier

Jika terdapat sekumpulan data independen  $\{(X_i, Y_i)\}_{i=1}^n$  yang mempunyai model (1) dan  $n = 2^m$  dengan  $m$  bilangan bulat positif. Jika  $X_i$  rancangan titik reguler pada interval  $[0,1]$  dengan  $X_i = i/n$ , maka estimator  $g$  pada level  $J$  adalah

$$\hat{g}_J(x) = \sum_{k \in \mathbb{Z}} \hat{c}_{j_0,k} \phi_{j_0,k}(x) + \sum_{j \geq j_0} \sum_{k \in \mathbb{Z}} \hat{d}_{j,k} \psi_{j,k}(x) \quad (5)$$

dengan  $\hat{c}_{j_0,k} = \frac{1}{n} \sum_{i=1}^n Y_i \phi_{j_0,k}(X_i)$ ,  $\hat{d}_{j,k} = \frac{1}{n} \sum_{i=1}^n Y_i \psi_{j,k}(X_i)$ , dan  $\hat{c}_{J,k} = \frac{1}{n} \sum_{i=1}^n Y_i \phi_{J,k}(X_i)$

yang merupakan estimator tak bias dari  $c_{j_0,k}$ ,  $d_{j,k}$  dan  $c_{J,k}$  [13]. Estimator wavelet (5) dinamakan estimator wavelet linier.

### 2.3 Estimator Wavelet Shrinkage.

Jika diberikan data  $\{(X_i, Y_i)\}_{i=1}^n$  dengan model (1),  $n = 2^m$  dan  $X_i = i/n$ , maka  $Y_i \sim$

$N(g(i/n), \sigma^2)$ . Selanjutnya,  $\sqrt{n}(\hat{d}_{j,k} - d_{j,k}) \sim N(0, \sigma^2)$  atau  $\hat{d}_{j,k} = d_{j,k} + \left(\frac{1}{\sqrt{n}}\right) z_{j,k}$

dengan  $d_{j,k}$  adalah dugaan koefisien  $g$  yang nyata dan  $z_{j,k}$  adalah  $n$  himpunan yang tidak teramati berdistribusi  $N(0, \sigma^2)$ . Jadi, koefisien wavelet empiris  $\hat{d}_{j,k}$  memuat sejumlah

noise dan hanya relatif sedikit yang memuat sinyal signifikan. Karena itu, dapat direkonstruksi wavelet dengan menggunakan sejumlah koefisien terbesar.

Dengan ide demikian, berkembang suatu metode yang menekankan rekonstruksi wavelet dengan menggunakan sejumlah koefisien wavelet terbesar, yakni hanya koefisien yang lebih besar dari suatu nilai tertentu yang diambil, sedangkan koefisien selebihnya diabaikan, karena dianggap  $0^{[7,8,9]}$ . Nilai tertentu tersebut dinamakan nilai threshold dan estimator waveletnya dinamakan estimator wavelet thresholding. Misal tersedia nilai threshold  $\lambda$ , maka estimator “thresholding” dari fungsi regresi  $g$  dapat dituliskan sebagai

$$\hat{g}_\lambda(x) = \sum_k \hat{c}_{j_0,k} \phi_{j_0,k}(x) + \sum_{j \geq j_0} \sum_{k=0}^{2^j-1} \partial_\lambda(\hat{d}_{j,k}) \psi_{j,k}(x) \quad (6)$$

Thresholding ini merupakan operator non linier pada vektor koefisien wavelet  $d$  yang diestimasi dengan  $\hat{d}$ . Dengan demikian, estimator thresholding tersebut dinamakan juga estimator wavelet non linier (shrinkage). Karena thresholding ini dirancang untuk membedakan antara koefisien wavelet empiris mana yang masuk dan mana yang keluar dari rekonstruksi wavelet, sedangkan untuk membuat keputusan ada 2 faktor yang mempengaruhi ketepatan estimator, yaitu ukuran sampel  $n$  dan tingkat noise  $\sigma^2$ , maka setiap koefisien merupakan calon kuat masuk didalam rekonstruksi wavelet jika ukuran sampel besar atau tingkat noise kecil. Karena  $\sqrt{n} \hat{d}_{j,k} / \sigma$  berdistribusi normal dengan varian 1 untuk seluruh  $n$  dan  $\sigma$ , maka estimator thresholding dari  $d_{j,k}$  adalah

$$\tilde{d}_{j,k} = \frac{\sigma}{\sqrt{n}} \partial_\lambda \left( \frac{\sqrt{n} \hat{d}_{j,k}}{\sigma} \right)$$

sehingga estimator wavelet shrinkage adalah

$$\hat{g}_\lambda(x) = \sum_k \hat{c}_{j_0,k} \phi_{j_0,k}(x) + \sum_{j \geq j_0} \sum_{k=0}^{2^j-1} \frac{\sigma}{\sqrt{n}} \partial_\lambda \left( \frac{\sqrt{n} \hat{d}_{j,k}}{\sigma} \right) \psi_{j,k}(x) \quad (7)$$

dengan

- $\hat{c}_{j_0,k}$  : penduga koefisien fungsi skala  $c_{j_0,k}$
- $\hat{d}_{j,k}$  : penduga koefisien wavelet  $d_{j,k}$
- $\lambda$  : parameter nilai threshold
- $\partial_\lambda$  : fungsi threshold

Metode wavelet shrinkage mampu mengestimasi fungsi baik mulus maupun tidak karena wavelet shrinkage mempunyai panjang support yang adaptif secara lokal. Artinya pada titik diskontinyu dan bagian fungsi tak mulus, metode wavelet shrinkage akan menggunakan support yang lebih sempit. Sedangkan untuk bagian fungsi mulus estimasi wavelet shrinkage akan menggunakan support yang lebih lebar dari data.

## 2.4 Langkah-langkah Thresholding

Langkah-langkah thresholding terdiri dari :

### 1. Pemilihan Fungsi Thresholding

Ada dua jenis fungsi thresholding  $\partial_\lambda$ , yaitu Hard Thresholding,

$$\partial_\lambda^H(x) = \begin{cases} x, & |x| > \lambda \\ 0, & x \text{ yang lain} \end{cases} \text{ dan Soft Thresholding, } \partial_\lambda^S(x) = \begin{cases} x - \lambda, & x > \lambda \\ 0, & x \leq \lambda \\ x + \lambda, & x < -\lambda \end{cases}$$

dengan  $\lambda$  merupakan parameter thresholding.

Fungsi Hard thresholding lebih dikenal karena terdapat diskontinyu dalam fungsi thresholding sehingga nilai  $x$  yang berada diatas threshold  $\lambda$  tidak disentuh. Sebaliknya, fungsi soft thresholding kontinyu yaitu sejak nilai  $x$  berada diatas threshold  $\lambda$ . Motivasi penggunaan soft thresholding berasal dari prinsip bahwa noise mempengaruhi seluruh koefisien wavelet. Juga kekontinyuan dari fungsi soft shrinkage membuat kondisi yang lebih baik untuk alasan statistik.

**2. Estimasi  $\sigma$**

Dalam merekonstruksi fungsi wavelet biasanya nilai  $\sigma$  tidak diketahui. Oleh karena itu,  $\sigma$  harus diestimasi dari data. Ogden<sup>[10]</sup> memberikan estimasi  $\sigma$  berdasarkan koefisien wavelet empiris pada level resolusi tertinggi dengan fungsi Median Deviasi Absolut (MAD), yaitu:

$$\hat{\sigma} = \frac{\text{median} \left( \left| \hat{d}_{J-1,k} - \text{median} \left( \hat{d}_{J-1,k} \right) \right| \right)}{0,6745} \tag{8}$$

**3. Pemilihan Parameter Thresholding**

Pada estimasi wavelet thresholding, tingkat kemulusan estimator ditentukan oleh level resolusi  $J$ , fungsi thresholding  $\partial_\lambda$  dan parameter threshold  $\lambda$ . Namun pemilihan  $J$  dan  $\partial_\lambda$  tidak sedominan  $\lambda$ . Nilai  $\lambda$  yang terlalu kecil memberikan estimasi fungsi yang sangat tidak mulus (under smooth) sedangkan nilai  $\lambda$  yang terlalu besar memberikan estimasi yang sangat mulus (over smooth). Oleh karena itu perlu dipilih parameter threshold yang optimal untuk mendapatkan fungsi yang optimal. Untuk memilih nilai threshold optimal, ada dua kategori pemilihan yaitu memilih satu harga threshold untuk seluruh level resolusi (pemilihan secara global) dan pemilihan threshold yang tergantung pada level resolusi (dependent level thresholding).

Untuk pemilihan threshold global, Ogden<sup>[10]</sup> memberikan 2 pemilihan threshold yang hanya bergantung pada banyaknya data pengamatan  $n$  yaitu threshold universal ( $\lambda_j = \sqrt{2 \log n}$ ) dan threshold minimax yang telah ditabelkan oleh Donoho dan Johnstone<sup>[6]</sup>. Nilai-nilai threshold minimax selalu lebih kecil dibandingkan dengan nilai threshold universal untuk ukuran sampel yang sama. Pemilihan threshold yang tergantung pada level resolusi berarti memilih  $\lambda_j$  bergantung level resolusi  $j$ . Dengan demikian ada kemungkinan perbedaan nilai threshold  $\lambda_j$  yang dipilih untuk tiap level wavelet  $j$ . Ada beberapa cara level-dependent thresholding diantaranya yaitu threshold Adapt dan threshold Top. Threshold adapt didasarkan pada prinsip untuk meminimalkan *Stein Unbiased Risk Estimator* (SURE) pada suatu level resolusi. Threshold adapt untuk himpunan koefisien detail  $d_j$  yang beranggotakan  $K$  koefisien didefinisikan sebagai,  $\lambda_j = \arg \min_{t \geq 0} \text{SURE}(d_j, t)$ , dengan

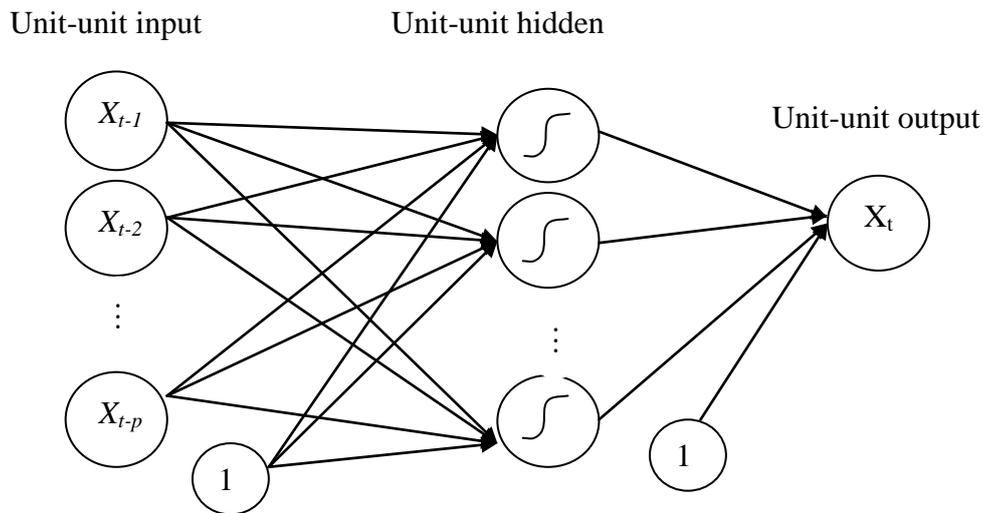
$$\text{SURE}(d_j, t) = K - 2 \sum_{k=1}^K \mathbb{1}_{\{|d_{j,k}| \leq t\}} + \sum_{k=1}^K \min \left\{ (d_{j,k} / \sigma_j)^2, t^2 \right\}.$$

Sedangkan nilai threshold top ditentukan berdasarkan besar prosentase koefisien yang akan digunakan dari keseluruhan koefisien wavelet yang ada<sup>[3]</sup>.

**2.5 Model Neural Network**

Model Neural Network yang diamati secara khusus dalam tulisan ini adalah model Feed Forward Neural Network (FFNN). Arsitektur jaringan FFNN untuk estimasi fungsi regresi nonparametrik dengan konfigurasi unit input  $X_1$  sampai  $X_p$  dan satu unit konstan

(bias), sebuah *hidden layer* yang terdiri dari n neuron dan 1 unit output diilustrasikan pada Gambar 2.



**Gambar 2.** Arsitektur FFNN untuk peramalan data time series dengan satu hidden layer yang terdiri n neuron dan variabel input nilai pada lag  $X_{t-1}, X_{t-2}, \dots, X_{t-p}$

Model FFNN dengan satu hidden layer dengan input  $X_{t-1}, X_{t-2}, \dots, X_{t-p}$  ditulis dalam bentuk

$$\hat{X}_t = \psi_o \left\{ w_{co} + \sum_n w_{no} \psi_n \left( w_{cn} + \sum_i w_{in} X_{t-j_i} \right) \right\} \quad (9)$$

dimana  $\{w_{cn}\}$  adalah bobot antara unit konstan dan neuron dan  $w_{co}$  adalah bobot antara unit konstan dan output.  $\{w_{in}\}$  dan  $\{w_{no}\}$  masing-masing menyatakan bobot koneksi input dengan neuron dan antara neuron dengan output. Kedua fungsi  $\psi_n$  dan  $\psi_o$  masing-masing fungsi aktivasi yang digunakan pada neuron dan output. Notasi yang digunakan untuk model neural network adalah  $NN(j_1, \dots, j_k, n)$  untuk menyatakan NN dengan variabel input pada lag  $j_1, \dots, j_k$  dan n neuron dalam satu hidden layer.

Algoritma pembelajaran yang akan digunakan untuk jaringan feed forward neural network adalah pembelajaran backpropagation yang meliputi tiga tahap yaitu umpan maju (*feedforward*) dari pola input, penghitungan dan propagasi balik dari error dan penyesuaian bobot-bobot. Pada tahap umpan maju setiap unit input (sensor) menerima sebuah sinyal input ( $x_i$ ) dan menyebarkan sinyal tersebut pada setiap unit-unit tersembunyi  $z_1, \dots, z_p$ . Setiap unit tersembunyi kemudian menghitung aktivasinya dan menghitung jumlah terboboti dari input-inputnya dalam bentuk

$$z_{in_j} = \sum_i w_{ji} x_i + w_{bj} \quad (10)$$

dimana  $x_i$  adalah aktivasi dari unit input ke-i yang mengirimkan sinyalnya ke unit hidden ke j dan  $w_j$  adalah bobot dari sinyal yang terkirim tersebut dan  $j = 1, 2, \dots, q$  adalah jumlah unit hidden. Sedangkan  $w_{bj}$  adalah bobot dari bias ke unit hidden ke-j. Hasil penjumlahan tersebut ditransformasi dengan fungsi aktivasi nonlinear  $f(\cdot)$  untuk memberikan aktivasi  $z_j$  dari unit j dalam bentuk

$$z_j = f(z_{in_j}) \quad (11)$$

Setelah semua unit tersembunyi menghitung aktivasinya kemudian mengirimkan sinyalnya ( $z_j$ ) pada unit output. Kemudian unit output menghitung aktivasinya untuk memberikan respon dari jaringan atas pola input yang dimasukkan dalam bentuk

$$g(w, z) = \sum_j w_{jo} z_j + w_{bo} \quad (12)$$

Fungsi pada (11) merupakan nilai output dari jaringan yaitu

$$y = \sum_j w_{jo} f(a_j) + w_{bo} \quad (13)$$

dimana  $w_{bo}$  adalah bobot dari bias ke unit output. Selama pelatihan unit output membandingkan aktivasi terhitung  $y_k$  dengan nilai targetnya  $t$  untuk menentukan error pada pola dengan unit tersebut<sup>[12]</sup>.

### 3. Hasil dan Pembahasan

#### 3.1 Model Wavelet Neural Network untuk Estimasi Regresi Non Parametrik

Wavelet terjadi di dalam keluarga fungsi-fungsi dan masing-masing didefinisikan melalui dilatasi yang mengontrol parameter skala dan translasi yang mengontrol posisi dari suatu fungsi tunggal, yang dinamakan wavelet ibu  $\psi(\mathbf{x})$ . Dengan fungsi pemetaan dalam domain frekuensi dan waktu, Wavelet Neural Network (WNN) diharapkan dapat merefleksikan fungsi menjadi lebih akurat daripada Neural Network. Misalkan diberikan suatu himpunan training dengan  $n$  elemen, respon dari WNN dituliskan sebagai :

$$y(\hat{w}) = \sum_{i=1}^{N_p} w_i \psi_i \left( \frac{x - t_i}{a_i} \right) \quad (14)$$

Dimana  $N_p$  adalah jumlah unit wavelet dalam hidden layer dan  $w_i$  adalah bobot sinapsis dari WNN. Suatu WNN dapat dipandang sebagai pendekatan fungsi (1). Ada beberapa pendekatan untuk konstruksi WNN, diantaranya diusulkan oleh Zhang<sup>[14]</sup>, yang mengkonstruksi WNN dalam dua tahap: Pertama, konstruksi wavelet  $W$  dari dilatasi dan translasi fungsi wavelet ibu  $\psi$  secara diskrit :

$$W = \left\{ \psi_i : \psi_i(x) = \alpha_i \psi(a_i(x - t_i)), \quad a_i = \left[ \sum_{k=1}^n [\psi(a_i(x_k - t_i))]^2 \right]^{\frac{1}{2}} \quad i = 1, \dots, L \right\} \quad (15)$$

Dimana  $x_k$  adalah input sampel and  $L$  adalah jumlah wavelet dalam  $W$ . Kemudian sejumlah  $M$  wavelet diseleksi berdasarkan data training untuk membangun regresi :

$$f_M(x) = \sum_{i \in I} u_i \psi_i(x) \quad (16)$$

dimana  $I$  adalah subset  $M$ -elemen dari himpunan indeks  $\{1, 2, \dots, L\}$  dan  $M \leq L$ . Yang kedua, untuk meminimumkan fungsi berikut :

$$J(I) = \min_{u_i, i \in I} \frac{1}{n} \sum_{k=1}^n \left( y_k - \sum_{i \in I} u_i \psi(x_k)_i \right)^2 \quad (17)$$

Zhang<sup>[14]</sup> menurunkan dua algoritma yaitu seleksi stepwise melalui orthogonalisasi wavelet dalam unit hidden dan eliminasi backward untuk pemilihan jumlah unit hidden. Jumlah wavelet yang digunakan yaitu  $M$ , dipilih sedemikian sehingga dapat meminimumkan nilai Akaike's final prediction error (FPE) :

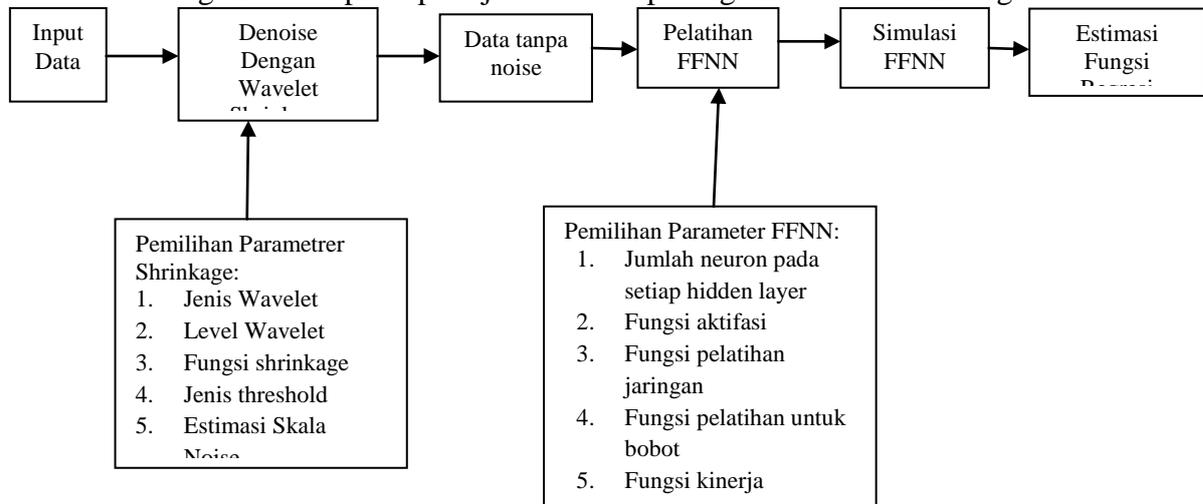
$$J_{FPE}(\hat{f}) = \frac{1 + n_{pa}/n}{1 - n_{pa}/n} \frac{1}{2n} \sum_{k=1}^{2n} (\hat{f}_k - y_k)^2 \quad (18)$$

Dimana  $n_{pa}$  adalah jumlah parameter dalam estimator. Setelah konstruksi awal dari WNN dibangun, kemudian ditraining dengan algoritma gradient descent seperti least mean squares (LMS) untuk meminimumkan mean-squared error:

$$J(w) = \frac{1}{n} \sum_{i=1}^n (y_i - y(\hat{w}))^2 \quad (19)$$

yang merupakan output real dari WNN yang telah ditraining pada vektor bobot akhir  $w^{[4]}$ .

Secara garis besar prinsip kerja WSNN dapat digambarkan dalam diagram berikut:

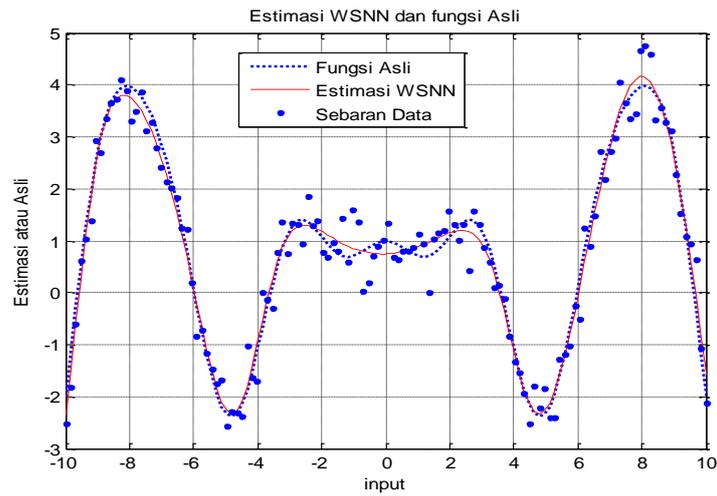


**Gambar 3.** Diagram Alur Estimasi Regresi dengan Metode WSNN

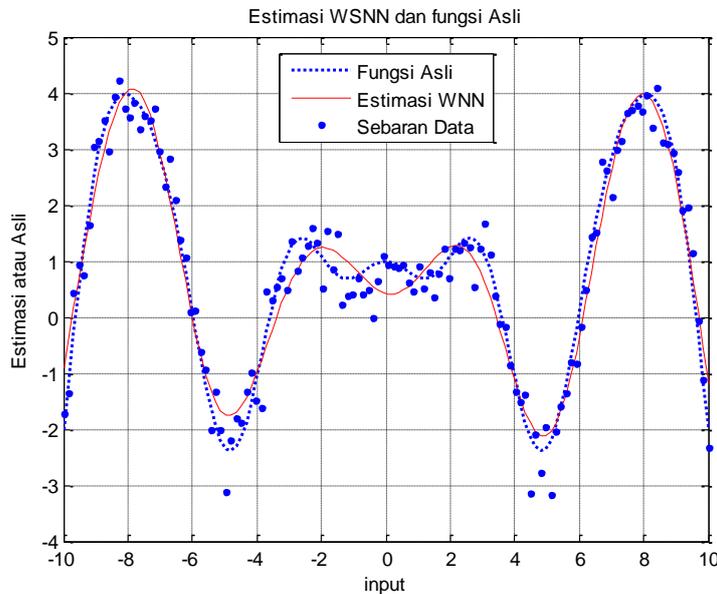
### 3.2 Contoh Simulasi

Data yang digunakan dibangkitkan dari sebuah fungsi  $g_i = 0.5X_i \sin(X_i) + \cos^2(X_i) + \varepsilon_i$ , ( $i= 1, 2, \dots, 128$ ) dengan  $X_i$  berada pada interval  $[-10, 10]$  dan  $\varepsilon_i$  berdistribusi  $N(0, 0.4)$ . Dari data tersebut, dilakukan denoising (penghilangan noise) dengan menggunakan wavelet Symmlet8, fungsi shrinkage Soft dan threshold Adapt. Kemudian setelah diperoleh data baru dilakukan estimasi regresi dengan FFNN menggunakan Fungsi aktifasi 'logsig' dan 'purelin', fungsi pelatihan jaringan 'trainlm' dan fungsi kinerja 'mse'.

Dengan menggunakan level 1 dan jumlah neuron 10 diperoleh  $mse = 0.0284$  dan dengan menggunakan level 5 dan jumlah neuron 5 diperoleh  $mse = 0.1200$ . Visualisasi dari estimasi fungsi regresi yang diperoleh sebagai berikut.



**Gambar 4.** Kurva Estimasi WSNN dengan level dekomposisi 1 dan jumlah neuron 10



**Gambar 5.** Kurva Estimasi WSNN dengan level dekomposisi 5 dan jumlah neuron 5

#### 4. Penutup

Estimasi fungsi regresi dengan wavelet shrinkage neural network merupakan estimasi fungsi dengan pendekatan non parametrik yang tidak mengasumsikan bahwa fungsi yang akan diestimasi harus mulus. Dalam estimasi fungsi wavelet shrinkage neural network, tingkat kemulusan fungsi ditentukan oleh parameter shrinkage yang digunakan pada proses denoising. Semakin besar nilai threshold dan level wavelet, akan semakin mulus hasil estimasi fungsinya. Selain itu banyaknya neuron pada unit hidden layer dan jumlah epoch yang optimal akan mempengaruhi hasil estimasi regresi.

## DAFTAR PUSTAKA

1. Abramovich, F. and Benjamini, Y., *Thresholding of Wavelet Coefficients as Multiple Hypothesis Testing Procedure*. In *Wavelets and Statistics*. Antoniadis, A., and Oppenheim, G. (eds.), Springer -Verlag, New York, 1995: 5-14.
2. Antoniadis, A., Gregore, G., and Mckeague, W., Wavelet Methods for Curve Estimation, *JASA*, 1994, Vol. 89, No. 428: 1340-1353.
3. Bruce, A. and Gao, H.Y., *Applied Wavelet Analysis with S-PLUS*, Springer-Verlag, New York, 1996.
4. Chen, S.C. and Li, S.T., Function Approximation using Robust Wavelet Neural Network. *NSC89-2213-E-327-007*, Taiwan, ROC.
5. Daubechies, I. , *Ten Lectures on Wavelets*, Capital City Press, Philadelphia, 1992.
6. Donoho, D.L. and Johnstone, I.M., Ideal Spatial Adaptation by Wavelet Shrinkage, *Biometrika*, 1994, Vol. 81, No. 3: 425-455.
7. Donoho, D.L. and Johnstone, I.M., Wavelet Shrinkage: Asymptopia?, *J.R.Statist.Soc.B*, 1995, Vol. 57, No.2: 301-369.
8. Hall, P. and Patil, P., On Wavelet Methods for Estimating Smooth Functions, *Bernoulli*, 1995, Vol. 1, No. 2: 041-058.
9. Hall, P. and Patil, P., On the Choice of Smoothing Parameter, Threshold and Truncation in Nonparametric Regression by Non-linear Wavelet Methods, *J.R.Statist.Soc.B*, 1996, Vol. 58, No.1: 370-397.
10. Ogden, R.T., *Essential Wavelets for Statistical Applications and Data Analysis*, Birkhauser, Boston, 1997.
11. Vetterli, M. and Kovacevic, J., *Wavelets And Subband Coding*, Prentice Hall PTR, New Jersey, 1995.
12. Warsito, B. dan Sumiyati, S., Prediksi Curah Hujan kota Semarang dengan Feedforward Neural Network Menggunakan Algoritma Quasi Newton BFGS dan Levenberg-Marquardt, *Jurnal PRESIPITASI UNDIP Semarang*, Edisi September 2007, Vol. 3, No 2.
13. Yasin, H., Estimasi Regresi Non Parametrik dengan Metode Wavelet Shrinkage pada Model Rancangan Tetap, *Skripsi Jurusan Matematika FMIPA Undip*, Semarang, 2005.
14. Zhang, Q., Using Wavelet Network in Nonparametric Estimation, *IEEE Trans. Neural Network*, 1997, Vol. 8: 227-236.