Regresi Kontinum dengan Prapemrosesan Transformasi Wavelet Diskret (*Continum Regression with Discrete Wavelet Transformation Preprocessing*)

Setiawan [1)] dan Khairil Anwar Notodiputro [2)]

[1)] *Staf Pengajar Jurusan Statistik FMIPA Institut Teknologi Sepuluh Nopember*

[2)] *Staf Pengajar Departemen Statistika FMIPA, Institut Pertanian Bogor*

ABSTRACT

In the multiple regression modeling, serious problems will be occurred if independent variables are correlated, to be named ill conditioned problem, and the number of observations is much less than the number of independent variables, it is a singularity problem. Continum Regression (CR) approach, it's better to overcome the problem of ill conditioned, but if the number of observations is much less than the number of independent variables usually facing the problem in computing. So the first step, it needes dimension reduction of independent variables (known as a preprocessing method). Discrete wavelet transformation (DWT) is one of a good method handle the problem of singularity. The research we have studied combination of CR and DWT as a preprocessing method can solved the problems of ill conditioned and singularity. The result of empirical research with simulation data has concluded that performance of CRDWT have very good potency to overcome the problems of the number of observations much less than the number of independent variables and ill conditioned.

Keywords : continum regression, singularity, discrete wavelet transformation, ill conditioned

PENDAHULUAN

Pada pemodelan regresi ganda permasalahan serius akan muncul jika banyaknya pengamatan (n) jauh lebih kecil dari pada banyaknya peubah bebas (p), serta terjadinya korelasi yang tinggi di antara peubah bebas (Naes, 1985). Regresi klasik (regresi kuadrat terkecil yang biasa, RKT) tidak dapat mengatasi masalah tersebut karena bila korelasi di antara peubah bebas tinggi (kolinearitas ganda), maka akan terjadi kondisi yang sakit (*ill conditioned).* Sedangkan bila banyaknya peubah bebas lebih besar dari pada banyaknya pengamatan (p>n), maka struktur matriks peubah bebas **X** menjadi singular (masalah singularitas).

Beberapa metode statistika yang telah dikembangkan untuk mengatasi masalah tersebut antara lain : Regresi Komponen Utama (RKU), Regresi Kuadrat Terkecil Parsial (RKTP), Regresi Ridge (RR), pendekatan Bayes, Regresi atas Koefisisen Fourier (RKF), Jaringan Syaraf Tiruan (JST), serta Transformasi *Wavelet*. Metode lain yang diperkenalkan Stone dan Brooks (1990) adalah Regresi Kontinum (RK) yang merupakan pengembangan dari RKT, RKU, serta RKTP. Hasil kajian Setiawan dan Notodiputro (2005a, 2005b) menunjukkan bahwa RK lebih unggul dari pada RKU dan RKTP dalam mengatasi masalah kolinearitas ganda.

RK merupakan pendekatan yang baik untuk mengatasi masalah *ill conditioned* dan singularitas, tetapi bila banyaknya pengamatan jauh lebih kecil dari pada banyaknya peubah bebas $(n << p)$ akan mengalami kendala komputasi. Sehingga pada tahap awal diperlukan pereduksian dimensi peubah bebas. Pemampatan dimensi peubah yang semula berdimensi tinggi $\mathbf{X}_{(nxp)}$ menjadi peubah baru, misalkan $\mathbf{Z}_{(nxk)}$ sehingga $k < (n-1) < p$ disebut metode prapemrosesan. Beberapa metode prapemrosesan yang sudah dipelajari adalah : analisis komponen utama (AKU), transformasi Fourier, transformasi *wavelet* diskret (TWD), serta pursuit proyeksi (Naes *et al*., 2002). Hasil penelitian Sunaryo (2005) menunjukkan bahwa dari kajian empirik TWD merupakan metode prapemrosesan yang sangat baik untuk mereduksi dimensi matriks peubah bebas.

Makalah ini bertujuan untuk mengkaji kombinasi antara RK dengan TWD dan disebut Regresi Kontinum dengan Transformasi *Wavelet* Disktret (RKTWD), untuk mengatasi masalah *ill conditioned* dan singularitas. Kajian dilakukan secara empirik dengan data bangkitan hasil simulasi.

Regresi Kontinum

Misalkan **X** matriks data yang sudah dipusatkan (*centred*) berukuran nxp dan disebut peubah bebas, sedangkan **y** adalah vektor peubah respon berukuran nx1 yang sudah dipusatkan, $\boldsymbol{\beta}$ vektor parameter regresi berukuran $p \times 1$, serta $\boldsymbol{\varepsilon}$ adalah vektor galat berukuran $n \times 1$. Regresi Kontinum dikembangkan berdasarkan model regresi linear klasik sebagai berikut :

$$\mathbf{y} = \mathbf{X\beta} + \boldsymbol{\varepsilon}$$

……….(1)

Pada model regresi linear terboboti formula matematis dapat ditulis sebagai berikut, maksimumkan

$$r_w^2 = \frac{\left(\sum_{i=1}^{n} y_i \mathbf{w}^T \mathbf{x}_i\right)^2}{\left(\sum_{i=1}^{n} y_i^2\right)\sum_{i=1}^{n} (\mathbf{w}^T \mathbf{x}_i)^2} = \frac{\left(\mathbf{w}^T \mathbf{s}\right)^2}{\|\mathbf{y}\|^2 \mathbf{w}^T \mathbf{S}\mathbf{w}} .$$

.............(2)

dengan $\mathbf{x}_i$ adalah vektor pengamatan peubah bebas ke-i (i=1,2, ..., n) berukuran (px1), $\mathbf{s} = \mathbf{X}^T\mathbf{y}$ dan $\mathbf{S} = \mathbf{X}^T\mathbf{X}$.

Regresi komponen utama pada prinsipnya adalah memaksimumkan :

$$S_w = \sum_{i=1}^{n} \left(\mathbf{w}^T\mathbf{x}_i\right)^2 = \mathbf{w}^T\mathbf{S}\mathbf{w}$$

..........(3)

Dari formula (3) tersebut dapat dijelaskan bahwa prinsip dasar dalam RKU adalah memaksimumkan keragaman dari peubah bebas **X** sehingga dibentuk peubah baru berupa beberapa komponen utama yang merupakan kombinasi linear dari peubah-peubah asal (X). Selanjutnya data peubah respon **y** diregresikan dengan beberapa komponen utama tersebut dengan menggunakan teknik regresi ganda.

RKTP prinsipnya adalah memaksimumkan :

$$S_w = \left(\sum_{i=1}^{n} y_i \mathbf{w}^T \mathbf{x}_i\right)^2 = \left(\mathbf{w}^\mathbf{T}\mathbf{s}\right)^2$$

..........(4)

Dari formula (4) tersebut dapat dilihat bahwa prinsip RKTP adalah memaksimumkan koragam antara peubah bebas dengan peubah respon.

Pada RK peubah baru diformulasikan dalam model sebagai berikut

$$\mathbf{y} = \mathbf{T_h}\boldsymbol{\xi} + \boldsymbol{\varepsilon}$$

..........(5)

dengan : $\mathbf{T}_h = \mathbf{X}\mathbf{W}_h$

..........(6)

dan $\mathbf{W}_h = \left(\mathbf{w_1}, \mathbf{w_2}, ..., \mathbf{w_h}\right)$ matriks berisi h kolom peubah dengan h<p dan disebut matriks pembobot.

Stone & Brooks (1990) memformulasikan matriks pembobot tersebut sebagai berikut :

$$\mathbf{w}_i = \arg\max_{w} \left\{Cov\left(\mathbf{Xw}, \mathbf{y}\right)^2 Var\left(\mathbf{Xw}\right)^{[\delta/(1-\delta)]-1}\right\}$$

.............(7)

dengan kendala $\|\mathbf{w}_i\| = 1$ dan $Cov\left(\mathbf{Xw}_i, \mathbf{Xw}_j\right) = 0$ untuk $i < j$ sedangkan parameter penyesuaian δ merupakan bilangan real $0 \le \delta \le 1$. Alternatif lain adalah formula yang dikembangkan oleh Malpass (1996) sebagai berikut :

$$\mathbf{w}_i = \arg\max_{w} \left\{Cov\left(\mathbf{Xw}, \mathbf{y}\right)^{(2+2\delta-4\delta^2)} Var\left(\mathbf{Xw}\right)^{(-1+2\delta)}\right\}$$

..........(8)

Dari formula 7 dibuat suatu formula yang umum sebagai berikut :

$$G = \left(\mathbf{w}^\mathbf{T}\mathbf{X}^\mathbf{T}\mathbf{y}\right)^2 \left(\mathbf{w}^\mathbf{T}\mathbf{X}^\mathbf{T}\mathbf{X}\mathbf{w}\right)^{[(\delta/(1-\delta))-1]}$$

...........(9)

selanjutnya disebut metode Stone. Formula 8 dapat diubah menjadi :

$$G = \left(\mathbf{w}^\mathbf{T}\mathbf{X}^\mathbf{T}\mathbf{y}\right)^{(2+2\delta-4\delta^2)} \left(\mathbf{w}^\mathbf{T}\mathbf{X}^\mathbf{T}\mathbf{X}\mathbf{w}\right)^{(-1+2\delta)}$$

............(10)

selanjutnya disebut metode Portsmouth (Malpass, 1996).

Formula tersebut merupakan generalisasi dari RKT, RKU serta RKTP dengan bentuk keterkaitan sebagai berikut :

Untuk $\delta = 0$, maka $G = \left(\mathbf{w}^\mathbf{T}\mathbf{s}\right)^2 \left(\mathbf{w}^\mathbf{T}\mathbf{S}\mathbf{w}\right)^{-1}$ formula ini ekivalen dengan persamaan 2, artinya pada $\delta = 0$ RK merupakan RKT.

Untuk $\delta = 0.5$, maka $G = \left(\mathbf{w}^\mathrm{T}\mathbf{s}\right)^2$ formula ini ekivalen dengan persamaan 4, sehingga pada $\delta = 0.5$ RK merupakan RKTP

Untuk $\delta = 1$, maka $G = \left(\mathbf{w}^T\mathbf{S}\mathbf{w}\right)^2$ formula ini ekivalen dengan persamaan 3, sehingga pada $\delta = 1$ RK merupakan RKU.

Dengan kata lain RK, RKU serta RKTP merupakan bentuk khusus dari RK.

Pendugaan parameter regresi $\boldsymbol{\xi}$ pada persamaan (5) dilakukan dengan menggunakan metode kuadrat terkecil yang diformulasikan sebagai berikut :

$$\hat{\boldsymbol{\xi}}_{\delta,h} = \left(\mathbf{T}_h^T\mathbf{T}_h\right)^{-1}\mathbf{T}_h^T\mathbf{y}$$

............(11)

$$\hat{\mathbf{y}}_{\delta,h} = \mathbf{X}\mathbf{W}_h\hat{\boldsymbol{\xi}}_{\delta,h}$$

$$\hat{\boldsymbol{\beta}}_{\delta,h} = \mathbf{W}_h\left(\mathbf{T}_h^T\mathbf{T}_h\right)^{-1}\mathbf{T}_h^T\mathbf{y}$$

..........(12)

dengan δ merupakan parameter penyesuaian dan *h* banyaknya komponen.

Transformasi Wavelet

Ada dua jenis fungsi *wavelet*, yaitu *mother wavelet* (ψ), serta *father wavelet* (ϕ). Suatu fungsi dikatakan *wavelet* jika memenuhi dua syarat, yaitu :

$$\int_{-\infty}^{\infty}\phi(x)\,dx = \int_{-\infty}^{\infty}\psi^2(x)\,dx = 1$$

(*father wavelet*)

$$\int_{-\infty}^{\infty}\psi(x)\,dx = 0\ .$$

(*mother wavelet*)

……..(13)

Wavelet yang sederhana yaitu *wavelet* Haar yang dikenalkan oleh Alferd Harr pada tahun 1909 (Vidocovic dan Meuller, 1991). Bila dilihat dari bilangan dilatasi dan translasi, terdapat dua jenis fungsi *wavelet*, yaitu : transformasi *wavelet* kontinu (TWK) bila bilangan tersebut real, serta transformasi *wavelet* diskret (TWD) bila bilangan tersebut bulat.

Fungsi basis diperoleh dengan dilatasi dan translasi fungsi *father wavelet* dan *mother wavelet* (Nason dan Silverman, 1994). Dari fungsi *wavelet* $\psi(x)$ dapat dibangkitkan fungsi basis dalam suatu ruang fungsi $L^2(\Re)$ dengan cara translasi dan dilatasi dari $\psi(x)$. Bentuk umum fungsi-fungsi basis tersebut adalah

$$\psi_{a,b}(x) = \left\{\frac{1}{\sqrt{|a|}}\psi\left(\frac{x-b}{a}\right);\ a > 0, -\infty < b < \infty\right\}.$$

..........(14)

Pada nilai khusus $a = 2^{-j}, b = k2^{-j};\ j,k \in Z = \{0,\pm1,...\}$ maka diperoleh sekumpulan fungsi basis yang saling ortogonal sehingga grafiknya tidak saling tumpang tindih.

Bentuk fungsi basis ortogonal yang diperoleh dengan cara dilatasi dan translasi dari fungsi *mother wavelet* $\psi(t)$,

$$\{\psi_{j,k}(x) = 2^{j/2}\ \psi(2^j x - k);\ j,k \in Z\},$$

.............(15)

Dari fungsi *father wavelet* $\phi(t)$ dapat dibangkitkan fungsi basis ortonormal yang menyusun ruang $L^2(\Re)$.

Bentuk umum fungsi basis dalam ruang $L^2(\Re)$ adalah

$$\{\phi_{jo,k}\ ,\psi_{j,k}\ ,\ j \ge j_o\ ,\ k \in Z\},$$

.............(16)

dengan $\phi_{0,0}(t)$ disebut fungsi skala, yang berhubungan dengan $\psi_{j,k}(t)$. Himpunan $\{\phi_{jo,k}\ ,\ k \in Z\}$ membentuk anak ruang yang sama seperti $\{\psi_{j,k}\ ,\ j \ge j_o\ ,\ k \in Z\}$.

Fungsi skala atau *father wavelet* , ϕ, merupakan penyelesaian dari persamaan

$$\phi(t) = \sqrt{2}\sum_k \ell_k\ \phi(2t-k)\ .$$

.............(17)

Fungsi $\phi(t)$ dapat membangkitkan suatu keluarga ortonormal pada ruang $L^2(\Re)$,

$$\phi_{j,k}(t) = 2^{j/2}\ \phi(2^j t - k)\ ,\ j,k \in Z$$

.............(18)

selanjutnya didapatkan *mother wavelet* ψ dari fungsi ϕ melalui persamaan :

$$\psi(t) = \sqrt{2}\sum_k h_k\ \phi(2t-k)\ ,$$

...........(19)

dengan $h_k = (-1)^k\ \ell_{1-k}$ (Vidacovic dan Meuller, 1991) dan disebut *quadrature mirror filter relation*. Sedangkan ℓ_k dan h_k merupakan koefisien-koefisien dari *low pass* dan *high pass filters* dan disebut *quadrature mirror filters* dan dapat digunakan menghitung TWD (Morettin, 1997). Koefisien-koefisien tersebut didefinisikan sebagai berikut :

$$\ell_k = \sqrt{2}\int_{-\infty}^{\infty}\phi(t)\ \phi(2t-k)\ dt\ \ ;$$

$$h_k = \sqrt{2}\int_{-\infty}^{\infty}\psi(t)\ \phi(2t-k)\ dt\ .$$

............(20)

Persamaan (17) dan (19) disebut persamaan dilatasi.

Berdasarkan sistem persamaan ortonormal : $\{\phi_{j,k}(t),\psi_{j,k}(t), j,k \in Z\}_{j\ge jo,k}$, maka $f(t) \in L^2(\Re)$ dapat didekomposisi menjadi :

$$f(t) = \sum_k c_{jo,k}\ \phi_{jo,k}(t) + \sum_{j\ge jo}\sum_k d_{j,k}\ \psi_{j,k}(t)\ .$$

.............(21)

dengan : $c_{jo,k} = \int_{-\infty}^{\infty} f(t)\,\phi_{jo,k}(t)\,dt$ dan $d_{j,k} = \int_{-\infty}^{\infty} f(t)\,\psi_{j,k}(t)\,dt$.

Transformasi *Wavelet* Diskret

Misalkan terdapat vektor data $\mathbf{x} = (x_0\,, x_1\,, \ldots, x_{q-1})^T$ dengan $q = 2^M$, M>0 integer. TWD didefinisikan sebagai berikut :

$$d_{j,k}^{(\psi)} = \sum_{t=0}^{q-1} x_t \psi_{j,k}(t)$$

.............(22)

$j = 0,1,2,..., M-1$ dan $k = 0,1,...,2^j - 1$ sehingga diperoleh $(q-1)$ koefisien dan satu koefisien $c_{0,0}$ sama dengan dimensi matriks peubah **X**.

Dengan notasi matriks transformasi pada persamaan (22) dapat ditulis :

$\mathbf{d} = \mathbf{Bx}$ ....(23)

karena **B** ortogonal, maka dapat ditulis :

$\mathbf{x} = \mathbf{B}^T\,\mathbf{d}$ ....(24)

dengan

$\mathbf{d} = (c_{0,0}\,, d_{0,0}\,, d_{1,0}\,, d_{1,1}\,, d_{2,0}, d_{2,1}, d_{2,2}, d_{2,3}, \ldots, d_{n-1,0}, \ldots)'$

dan $\mathbf{B}^T$ adalah matriks yang elemen-elemen kolomnya adalah nilai dari $\phi(t)$ dan $\psi_{j,k}(t)$ untuk berbagai t $\in [0, 1]$. Sifat-sifat khusus dari matriks $\mathbf{B}^T$ adalah : ortonormal, kolom pertama bernilai sama, serta jumlah unsur tiap kolom yang lain sama dengan nol.

Vektor data **x** dapat dihubungkan dengan fungsi f pada interval [0,1) dan didefinisikan sebagai :

nnn.

..........(25)

Fungsi ini dikenal dengan fungsi tangga dan termasuk dalam $L^2([0,1])$, sehingga dekomposisi *wavelet* dari $f(t)$ adalah :

$$f(t) = c_{0,0}\;\phi(t) + \sum_{j=0}^{M-1} \sum_{k=0}^{2^j-1} d_{j,k}\;\psi_{j,k}(t)\,.$$

..........(26)

untuk $\phi(t) = 1$ disebut fungsi skala untuk wavelet Haar.

Persamaan (26) disebut transformasi *wavelet* diskret, karena nilai j hanya diambil pada bilangan bulat positif saja. Bilangan j pada persamaan (26) disebut level resolusi, dan f(t) dapat diperoleh secara tepat, jika diambil semua level resolusi untuk dekomposisi, yaitu level resolusi 0 sampai dengan (M-1). Koefisien $c_{0,0}$ disebut koefisien pemulusan atau bagian pendekatan dari suatu fungsi, sedang $d_{j,k}$ disebut koefisien *wavelet* atau juga disebut bagian detail suatu fungsi. Contoh bentuk matriks $\mathbf{B}^T$ dari Haar *wavelet* untuk M=3 ($2^M = 8$) adalah :

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \\ x_7 \\ x_8 \end{bmatrix} = \begin{bmatrix} \frac{1}{2\sqrt{2}} & \frac{1}{2\sqrt{2}} & \frac{1}{2} & 0 & \frac{1}{\sqrt{2}} & 0 & 0 & 0 \\ \frac{1}{2\sqrt{2}} & \frac{1}{2\sqrt{2}} & \frac{1}{2} & 0 & -\frac{1}{\sqrt{2}} & 0 & 0 & 0 \\ \frac{1}{2\sqrt{2}} & \frac{1}{2\sqrt{2}} & -\frac{1}{2} & 0 & 0 & \frac{1}{\sqrt{2}} & 0 & 0 \\ \frac{1}{2\sqrt{2}} & \frac{1}{2\sqrt{2}} & -\frac{1}{2} & 0 & 0 & -\frac{1}{\sqrt{2}} & 0 & 0 \\ \frac{1}{2\sqrt{2}} & -\frac{1}{2\sqrt{2}} & 0 & \frac{1}{2} & 0 & 0 & \frac{1}{\sqrt{2}} & 0 \\ \frac{1}{2\sqrt{2}} & -\frac{1}{2\sqrt{2}} & 0 & \frac{1}{2} & 0 & 0 & -\frac{1}{\sqrt{2}} & 0 \\ \frac{1}{2\sqrt{2}} & -\frac{1}{2\sqrt{2}} & 0 & -\frac{1}{2} & 0 & 0 & 0 & \frac{1}{\sqrt{2}} \\ \frac{1}{2\sqrt{2}} & -\frac{1}{2\sqrt{2}} & 0 & -\frac{1}{2} & 0 & 0 & 0 & -\frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} c_{0,0} \\ d_{0,0} \\ d_{1,0} \\ d_{1,1} \\ d_{2,0} \\ d_{2,1} \\ d_{2,2} \\ d_{2,3} \end{bmatrix}$$

Regresi Kontinum Dengan Transformasi *Wavelet* Diskret

RKTWD pada dasarnya adalah regresi kontinum antara peubah respon Y dengan peubah bebas **D** hasil TWD. Matriks **D** hasil transformasi *Wavelet* tidak dijamin saling bebas. Sunaryo (2005) menunjukkan bahwa peubah baru hasil TWD masih mempunyai korelasi yang cukup tinggi walaupun nilainya lebih kecil dari pada korelasi antar peubah asal. Oleh karena itu pemodelan regresi antara peubah respon (Y) dengan peubah hasil TWD tidak dapat menggunakan RKT karena masih ada masalah kolinearitas ganda.

Tahapan-tahapan dalam RKTWD dapat dijelaskan sebagai berikut :

Tahap 1. Mendapatkan matriks **D** dengan menggunakan TWD.

Jika matriks $\mathbf{X}_{(nxp)} = [\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_p]^T$; dengan p=q=2^M (M=bilangan bulat positip) maka TWD dapat ditulis :

$$\mathbf{D}_{(nxp)} = \mathbf{X}_{(nxp)}\mathbf{B}^T_{(pxp)} \qquad \text{...........(27)}$$

Karena dimensi dari **B** sangat besar, maka dipilih level-level resolusi tertentu sedemikian hingga banyaknya koefisien wavelet yang terpilih sebesar p’ dengan $p' < (n-1) < p$ maka akan diperoleh

$$\mathbf{D}^*_{(nxp')} = \mathbf{X}_{(nxp)}\mathbf{B}^{*T}_{(pxp')} \qquad \text{..........(28)}$$

yang mereduksi pengamatan dari p titik tiap-tiap contoh menjadi p’ titik koefisien *wavelet* yang terpilih.

Tahap 2. Meregresikan antara peubah respon $\mathbf{y}_{(nx1)}$ terhadap peubah bebas $\mathbf{D}^*_{(nxp')}$ dengan menggunakan metode regresi kontinum.

$$\mathbf{y} = \mathbf{D}^*_{(\mathbf{nxp'})}\boldsymbol{\beta} + \boldsymbol{\xi} \qquad \text{...........(29)}$$

Walaupun dimensi dari $\mathbf{D}^*_{(nxp')}$ sudah kecil, tetapi korelasi diantara koefisien-koefisien Wavelet masih tinggi (Sunaryo 2005). Sehingga dari matriks peubah bebas $\mathbf{D}^*_{(nxp')}$ dilakukan transformasi menjadi :

$$\mathbf{T}_h = \mathbf{D}^*\mathbf{W}_h \qquad \text{............(30)}$$

dengan $\mathbf{W}_h = (\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_h)$ matriks berukuran p’xh dimana $h < p' < (n-1)$ dan disebut matriks pembobot, sedangkan $\mathbf{T}_h$ adalah matriks peubah baru berukuran nxh.

Dengan demikian persamaan (29) dapat diubah menjadi sebagai berikut :

$$\mathbf{y} = \mathbf{T_h}\boldsymbol{\xi} + \boldsymbol{\varepsilon} \qquad \text{...........(31)}$$

Karena pada matriks $\mathbf{T}_h$ sudah tidak ada masalah singularitas atau *ill conditioned*, maka pendugaan parameter $\boldsymbol{\xi}$ pada persamaan (31) dapat dilakukan dengan menggunakan metode kuadrat terkecil dengan menggunakan persamaan 11.

METODE

Pada penelitian digunakan kajian empirik dengan menggunakan data bangkitan simulasi dengan menggunakan program SAS. Ada dua data, yaitu : (1) n=20 dan p=128, serta (2) n=20 dan p = 256. Kedua jenis data tersebut mempunyai tingkat korelasi antar peubah bebas yang sangat tinggi (r= 0,99). Selanjutnya dari data hasil bangkitan tersebut dilakukan penduguaan model dengan menggunakan metode RKTWD. Hasil pendugaan tersebut dibandingkan dengan hasil metode RKUTWD dan RKTPTWD.

Untuk mengevalusi kinerja RKTWD ada beberapa kriteria yang dibandingkan, antara lain : R^2, s, serta plot antara y dengan $\hat{y}$. Model dikatakan lebih baik jika R^2 lebih tinggi, s lebih kecil, serta hasil pengepasan lebih dekat ke garis lurus dengan gradien 45^o melalui pusat koordinat.

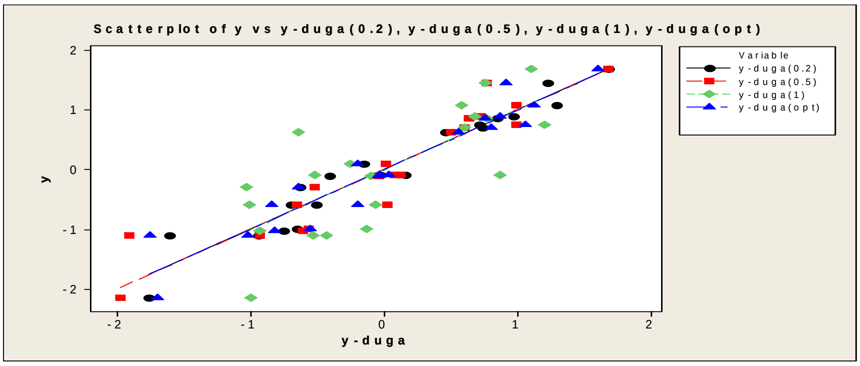
Untuk menghitung matriks koefisien wavelet digunakan perangkat lunak (software) wavetresh 3 seperti yang dijelaskan oleh Nason (1998). Sedangkan untuk RKTWD, RKUTWD serta RKTP TWD digunakan perangkat lunak SAS.

HASIL DAN PEMBAHASAN

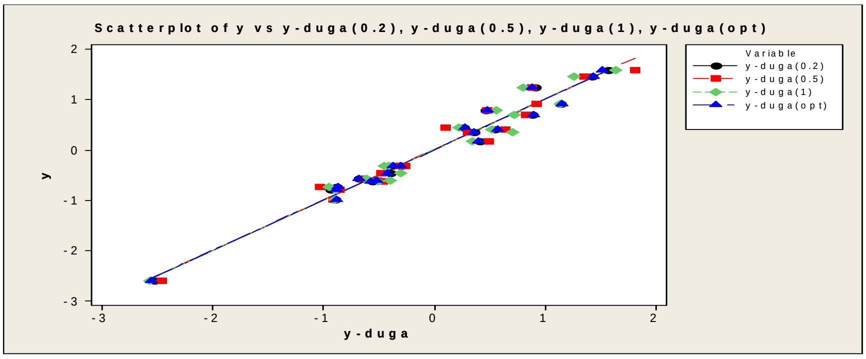
Ringkasan ukuran kebaikan model hasil pendugaan RK, RKTP, RKU dengan metode pra-pemrosesan TWD disajikan pada Tabel 1. Sedangkan diagram pencar antara y dengan $\hat{y}$ hasil pendugaan metode RKTWD, RKUTWD, serta RKTPTWD disajikan pada Gambar 1 dan 2.

Tabel 1 Ringkasan ukuran kebaikan model hasil pendugaan RK, RKTP, RKU dengan metode pra-pemrosesan TWD

Data	Kriteria	δ						
		0.15	0.2	0.3	0.5 RKTPTWD	0.6	0.7	1 RKUTWD
Simulasi 1 N=20, p=128	h	2	3	4	4	4	5	4
	R^2	0.929	0.942	0.941	0.838	0.848	0.828	0.575
	$\bar{R}^2$	0.926	0.934	0.929	0.816	0.808	0.783	0.495
	s	0.281	0.264	0.274	0.389	0.141	0.161	0.734
Simulasi 2 N=20 P=256	h	2	3	3	3	4	4	7
	R^2	0.973	0.975	0.973	0.926	0.966	0.906	0.926
	$\bar{R}^2$	0.971	0.971	0.970	0.907	0.960	0.903	0.945
	s	0.175	0.173	0.180	0.231	0.207	0.252	0.244



Gambar 1 Diagram pencar antara y dengan $\hat{y}$ hasil pendugaan metode RKTWD, RKUTWD, RKTPTWD untuk data simulasi 1



Gambar 2 Diagram pencar antara y dengan $\hat{y}$ hasil pendugaan metode RK-TWD, RKU-TW+D, RKTP-TWD untuk data simulasi 2

Hasil pendugaan pada data simulasi 1 (n=20, p=128) dengan mengambil 16 koefisien *wavelet* pada resolosi 0, 1, 2, 3 serta 1 koefisien untuk fungsi skala, transformasi *wavelet* diskret memberikan hasil yang memuaskan. Hal ini dapat dilihat pada Tabel 1 yang tampak bahwa pada $\delta = 0.2$ (RKTWD) menghasilkan koefisien determinasi paling besar (R^2=94,2%) serta simpangan baku paling kecil (s=0.264) dibandingkan dengan $\delta = 0.5$ (RKTPTWD) serta $\delta = 1$ (RKUTWD) yang menghasilkan R^2 dan *s* masing-masing 83,8%, 0,389 serta 57,5% dan 0,734. Hal ini ditunjang dengan Gambar 1 yang terlihat bahwa pada $\delta = 0.2$ titik-titik yang ada lebih dekat ke garis 45^o dari pada $\delta = 0.5$ serta $\delta = 1$.

Untuk data simulasi 2 (*n*=20, p=256) hasil TWD terbaik pada 8 koefisein wavelet dengan resolusi 0, 1, 2 serta 1 koefisien fungsi skala. Pada Tabel 1 dapat dilihat bahwa pada $\delta = 0.2$ (RKTWD) menghasilkan koefisien determinasi paling besar (R^2=97,5 %) serta simpangan baku paling kecil (s=0,173) dibandingkan dengan $\delta = 0.5$ (RKTPTWD) serta $\delta = 1$ (RKUTWD). Hal ini ditunjang dengan Gambar 2 yang terlihat bahwa pada $\delta = 0.2$ titik-titik yang ada lebih dekat ke garis 45^o dari pada $\delta = 0.5$ (RKTPTWD) serta $\delta = 1$ (RKUTWD).

Simulasi dilakukan berulang-ulang dengan cara *resampling*, hasil ini merupakan salah satu hasil simulasi yang telah dicobakan. Hasil yang diperoleh dengan *resampling* menunjukkan bahwa pendugaan dengan RKTWD lebih baik dari pada RKUTWD maupun RKTPTWD (Setiawan dan Notodiputro, 2005b).

KESIMPULAN

Dari hasil kajian empirik dengan data simulasi ternyata regresi kontinum dengan metode pra-pemrosesan transformasi *wavelet* (RKTWD) mempunyai potensi yang sangat baik untuk mengatasi masalah kondisi yang sakit dan singularitas. $R^2, \overline{R}^2$ serta *s* belum cukup sebagai kriteria untuk melihat kebaikan suatu model. Perlu dilakukan penelitian lebih lanjut dengan menggunakan *Root Mean Squares Error of Prediction* (RMSEP) sebagai kriteria kebaikan model.

Ucapan Terimakasih

Terimakasih penulis sampaikan kepada DITJEN DIKTI DEPDIKNAS yang telah mendukung penelitian ini melalui Hibah Penelitian Pascasarjana IPB yang diketuai oleh Prof. Dr. Ir. H. Khairil Anwar Notodiputro, MS.

DAFTAR PUSTAKA

Mallpass J. 1996. *Improved Mathematical Methods for Drugs Design : Continum Regression SAS Macro*. University of Portsmouth.

Morettin PA. 1997. *Wavelets in Statistics. Institute of Math. And Stat*., University of Sao Paulo, Brazil.

Naes T. 1985. Multivariate Calibration When the Error Covariance Matrix is Structured. *J Technometrics* **27**: 301-311.

Naes T, Isaksoon T, Fearn T, Davis T. 2002. *A User Friendly Guide to Multivariate Calibration and Classification*. NIR Publication, UK.

Nason GP, Silverman BW. 1994. The discrete wavelet transform in S. *J Comp graph Stat* **3**: 163-191.

Nason GP. 1998. *Wavethresh 3 software*. Department of Mathematics, University of Bristol, UK. http://www.stats.bris.ac.uk/~wavethresh [20 juni 2003]

Percival DB. 2005. *Wavelets : Data Analysis, Algorithms and Theory*. University Washington. http://www.ms.washington.edu/~s530/ [18 april 2005].

Setiawan, Notodiputro KA. 2005a. Regresi Kontinum sebagai Bentuk Umum dari RKT, RKU, serta RKTP. *Prosiding Seminar Nasional Statistika VII*. Jurusan Statistika FMIPA ITS, Surabaya tanggal 26 Nopember 2005.

Setiawan, Notodiputro KA. 2005b. Regresi Kontinum dengan Prapemrosesan Transformasi Wavelet dalam Model Kalibrasi. *Prosiding Seminar Nasional MIPA. FMIPA UNESA*, Surabaya tanggal 17 Desember 2005.

Stone M, & Brooks RJ. 1990. Continum Regression:cross-validated equentially constructed prediction embracing ordinary least squares, partial least squares, and principal component regression (with discussion). *Journal of the Royal Statistical Society Series B* **52**: 237-269.

Sunaryo S. 2005. *Model Kalibrasi dengan Transformasi Wavelet sebagai Metode Pra-pemrosesan*. Disertasi yang tidak dipublikasikan. Bogor : Sekolah Pascasarjana, Institut Pertanian Bogor.

Vidacovic B, Meuller P. 1991. *Wavelets for Kids. A Tutorial Introduction. AMS Subject Classification*, Duke University.