

Unified Discrete Wavelet Transform with Ridge Regression and Principal Component Regression to Predict Concentration of Gingerol Compound in Ginger Crop

Sony Sunaryo¹

Abstract—Multivariate calibration model can be used as an alternative method to predict the concentration of a gingerol compound. The prediction usually are carried out chemically through a long and expensive process using High Performance Liquid Chromatography (HPLC) measurement method. Since the number of samples (n) are less than of the number of independent variables (p), and between the independent variables are correlated, so the development of model using conventional regression is not valid. The combination of Discrete Wavelete Transform (DWT) with Ridge Regression and Principal Component Regression have been adopted in this paper to predict concentration of gingerol, and it showed a promising result.

Keywords—Calibration model, Wavelet, Partial Least Square, gingerol

I. PENDAHULUAN

Chemometrics dapat dipandang sebagai gabungan antara matematika dan statistika dengan kimia. Kalibrasi peubah ganda merupakan bagian dari *Chemometrics* yang bertujuan untuk menemukan hubungan antara sekumpulan ukuran yang relatif mudah atau murah memperolehnya, dengan sekelompok ukuran lain yang relatif sulit (*labour intensive*) atau mahal memperolehnya. Menurut referensi [1] menyebutkan bahwa tujuan kalibrasi peubah ganda adalah menemukan model yang dapat digunakan untuk memprediksi ukuran-ukuran yang mahal dengan tepat dan akurat dari ukuran-ukuran yang murah. Secara umum kalibrasi peubah ganda menggunakan formula matematika untuk menduga informasi pada Y , yaitu ukuran yang mahal, yang tidak diketahui berdasarkan informasi pada X , yaitu ukuran yang murah, yang tersedia [2].

Menurut referensi [1] pembuatan model untuk memprediksi Y dengan kalibrasi peubah ganda, yaitu dengan mempertimbangkan beberapa atau semua pengamatan pada spektrum yang berupa pengamatan absorban, akan memberikan hasil lebih baik dibanding dengan pemodelan kalibrasi peubah tunggal yang hanya mempertimbangkan satu puncak pada masing-masing spektrum. Mengkombinasikan informasi dari beberapa atau bahkan semua peubah spektrum, permasalahan yang muncul

pada pendugaan model kalibrasi ganda adalah kasus multikolinearitas di antara peubah absorban (X) dan banyaknya sampel (n) yang lebih kecil dari banyaknya peubah bebas (p) [1], [2], sehingga metode baku seperti model regresi sering memberikan solusi yang tidak stabil. Oleh karena itu diperlukan suatu metode yang dapat mengatasi masalah tersebut, sehingga diperoleh solusi yang lebih stabil.

Penentuan kadar senyawa gingerol pada rimpang jahe secara kuantitatif dilakukan melalui proses yang panjang meliputi penghancuran bahan, pelarutan, dan pengukuran dengan HPLC (*High Performance Liquid Chromatography*). Proses tersebut memerlukan waktu dan biaya yang relatif mahal [1]. Pengukuran dengan FTIR (*Fourier Transform Infra Red*) relatif lebih mudah dan murah untuk dilakukan daripada pengukuran dengan HPLC. Setiap bentuk spektrum absorban dari FTIR akan mencerminkan gugus fungsi yang terdapat pada senyawa gingerol dari suatu sampel rimpang jahe. Cara alternatif untuk memprediksi kadar gingerol pada rimpang jahe adalah dengan mengembangkan model kalibrasi peubah ganda yang menyatakan hubungan kadar senyawa aktif hasil dari HPLC (sebagai peubah tak bebas Y) dengan data hasil pengukuran bilangan gelombang menggunakan FTIR dari serbuk rimpang jahe yang berupa data spektra absorban (sebagai peubah bebas X).

Dimensi peubah bebas X sangat tinggi dan antar peubah saling berkorelasi, maka kasus jumlah pengamatan sampel lebih kecil dari jumlah peubah bebas, dan kasus multikolinearitas sering muncul dalam kalibrasi peubah ganda. Sehingga penanganan dengan metode regresi ganda baku secara langsung kurang valid. Data spektra absorban (X) akan berupa sederetan data vektor $\underline{x} = (x_1, x_2, \dots, x_p)^T$ yang berdimensi tinggi dan saling berkorelasi, sehingga pengembangan model kalibrasi peubah ganda $Y = f(X)$ dengan mengikutkan semua data X menjadi tidak efisien. Melalui reduksi dimensi, diharapkan pengembangan model kalibrasi peubah ganda menjadi lebih efisien. Reduksi dimensi yang digunakan dalam paper ini adalah metode transformasi *wavelet* diskret (TWD), dengan alasan [3] metode ini setelah dikaji ternyata lebih baik dibanding metode lain seperti transformasi Fourier maupun PCA (*Principal Component Analysis*).

Fokus metode *wavelet* yang digunakan hanya sebagai reduksi dimensi, bukan untuk mengatasi kasus multikolinearitas, maka dimungkinkan untuk menggunakan hasil keluaran dari metode *wavelet* sebagai masukan untuk

Naskah diterima 1 Mei 2007; selesai revisi pada 8 Mei 2008

¹ Sony S. Adalah dosen Jurusan Statistika, FMIPA, Institut Teknologi Sepuluh Nopember, Surabaya, INDONESIA (e-mail: sonny_s@statistika.its.ac.id)

metode kalibrasi peubah ganda yang lain, seperti regresi bertatar, regresi komponen utama (PCR) dan PLS, sehingga akan diperoleh model yang lebih baik. Pada paper ini akan dibahas penerapan gabungan metode *wavelet* dengan *Ridge Regression* dan regresi komponen utama untuk memprediksi kadar senyawa gingerol pada rimpang jahe. Paper ini merupakan pengembangan dari referensi [3], [4], [5]. Perhitungan matriks koefisien *wavelet* dengan menggunakan *software* wavetresh 3 seperti yang diterangkan referensi [6], [7].

A. Ridge Regression dan Principal Component Regression

Ridge Regression dan *Principal Component Regression* merupakan metode untuk mengatasi kasus multikolinearitas dalam analisis regresi. Bentuk umum regresi linear berganda adalah :

$$y = 1 b_0 + X_1 \underline{b} + e \tag{1}$$

atau bisa ditulis :

$$y = X \underline{\beta} + e,$$

dengan $E(y) = X \underline{\beta}$, $E(e) = 0$ dan $Var(e) = I \sigma^2$,

dengan $\underline{b}^T = [b_1, b_2, \dots, b_p]$, $\underline{\beta}^T = [b_0, b_1, \dots, b_p]$,

$$X_1 = \begin{bmatrix} x_{11} & x_{12} & \dots & x_{1p} \\ x_{21} & x_{22} & \dots & x_{2p} \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ x_{n1} & x_{n2} & \dots & x_{np} \end{bmatrix},$$

$$X = \begin{bmatrix} 1 & x_{11} & x_{12} & \dots & x_{1p} \\ 1 & x_{21} & x_{22} & \dots & x_{2p} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 1 & x_{n1} & x_{n2} & \dots & x_{np} \end{bmatrix}, \quad \underline{y} = \begin{bmatrix} y_1 \\ y_2 \\ \cdot \\ \cdot \\ \cdot \\ y_n \end{bmatrix} \quad \text{dan} \quad \underline{e} = \begin{bmatrix} e_1 \\ e_2 \\ \cdot \\ \cdot \\ \cdot \\ e_n \end{bmatrix}.$$

Dugaan kuadrat terkecil dari persamaan (2) adalah : $\hat{\underline{\beta}} = (X^T X)^{-1} X^T \underline{y}$,

Jika terjadi kasus multikolinearitas maka $(X^T X)$ susah dihitung inversnya secara akurat. Penambahan konstanta k yang cukup besar, yang berakibat $((X^T X) + k I)$ dapat dihitung inversnya secara lebih akurat. Hal inilah yang mendasari pendugaan koefisien regresi dalam metode *Ridge Regression* berbentuk :

$$\hat{\underline{\beta}}_R = (X^T X + k I)^{-1} X^T \underline{y}$$

Yang perlu dicatat bahwa $\hat{\underline{\beta}}_R$ merupakan penduga berbias bagi $\underline{\beta}$ tetapi $\hat{\underline{\beta}}_R$ memiliki varian yang lebih kecil dari $\hat{\underline{\beta}}$. Pemilihan k didasarkan pada kestabilan nilai $\hat{\underline{\beta}}_R$ yang diperoleh.

Principal Component Regression adalah meregresikan \underline{y} dengan komponen-komponen utama yang dibentuk dari kovarian matriks X. Komponen utama yang terbentuk merupakan kombinasi linear dari variabel asli (X). Antar komponen utama secara matematis dapat ditunjuk-

kan saling bebas, sehingga kasus multikolinearitas jelas teratasi.

B. Transformasi Wavelet Diskret (TWD)

Di dalam statistika biasanya ingin diperoleh dekomposisi *wavelet* dari suatu fungsi yang diamati pada sekumpulan data. Misalkan $\underline{x} = (x_0, x_1, \dots, x_{2^M-1})^T$ adalah vektor data berukuran 2^M , M bilangan bulat positif. Maka vektor data tersebut dapat dihubungkan dengan potongan-potongan fungsi konstan pada interval [0,1] yang biasa disebut fungsi tangga, dengan persamaan :

$$f(t) = \sum_{k=0}^{2^M-1} x_k I_{\left\{ \frac{k}{2^M} \leq t < \frac{(k+1)}{2^M} \right\}}$$

dimana fungsi indikator,

$$I_{\left\{ \frac{k}{2^M} \leq t < \frac{(k+1)}{2^M} \right\}} = \begin{cases} 1, & \frac{k}{2^M} \leq t < \frac{(k+1)}{2^M} \\ 0, & t \text{ selainya} \end{cases}$$

Fungsi tangga f(t) pada persamaan (3) termasuk dalam $L^2([0,1])$, artinya f(t) fungsi yang terintegralkan kuadrat pada interval [0,1], sehingga dekomposisi *wavelet* dari f(t) adalah [12]:

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

Persamaan (4) disebut transformasi *wavelet* diskret, karena nilai j hanya diambil pada bilangan bulat positif saja. Bilangan j pada persamaan (4) 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.

Fungsi $\psi_{j,k}(t)$ dan $\phi(t)$ masing-masing disebut fungsi *mother wavelet* dan *father wavelet*. Penentuan nilai $\psi_{j,k}(t)$ dan $\phi(t)$ untuk berbagai t, berakibat persamaan (4) dapat dituliskan dengan notasi matriks,

$$\underline{x} = W^T \underline{d}$$

dimana W^T adalah matriks yang elemen-elemen kolomnya adalah nilai dari $\phi(t)$ dan $\psi_{j,k}(t)$ untuk berbagai $t \in [0, 1]$ yang saling ortonormal (bukti di referensi [13]). Matriks W^T yang ortonormal menyebabkan (5) dapat ditulis

$$\underline{d} = W \underline{x}$$

dimana $\underline{d} = (c_{0,0}, d_{0,0}, d_{1,1}, d_{1,0}, \dots, d_{n-1,0})^T$ dan Sifat-sifat menarik dari matriks W^T , selain ortonormal, adalah kolom pertama bernilai sama, jumlah unsur tiap kolom yang lain sama dengan nol.

Contoh bentuk matriks W^T dari Haar *wavelet* untuk $2^M = 8$ adalah :

$$\begin{bmatrix} 0.353553 & 0.707107 & 0.000000 & 0.000000 & 0.000000 & 0.5 & 0.0 & 0.353553 \\ 0.353553 & -0.707107 & 0.000000 & 0.000000 & 0.000000 & 0.5 & 0.0 & 0.353553 \\ 0.353553 & 0.000000 & 0.707107 & 0.000000 & 0.000000 & -0.5 & 0.0 & 0.353553 \\ 0.353553 & 0.000000 & -0.707107 & 0.000000 & 0.000000 & -0.5 & 0.0 & 0.353553 \\ 0.353553 & 0.000000 & 0.000000 & 0.707107 & 0.000000 & 0.0 & 0.5 & -0.353553 \\ 0.353553 & 0.000000 & 0.000000 & -0.707107 & 0.000000 & 0.0 & 0.5 & -0.353553 \\ 0.353553 & 0.000000 & 0.000000 & 0.000000 & 0.707107 & 0.0 & -0.5 & -0.353553 \\ 0.353553 & 0.000000 & 0.000000 & 0.000000 & -0.707107 & 0.0 & -0.5 & -0.353553 \end{bmatrix}$$

Jika ukuran vektor data \underline{x} sangat besar, maka perhitungan dengan cara matriks akan memerlukan komputasi yang tinggi, sehingga menjadi kurang praktis. Mallat (1989) menemukan algoritma cepat untuk menghitung koefisien *wavelet* dan koefisien pemulusan pada persamaan (4), yaitu melalui analisis multiresolusi. Algoritma tersebut disebut algoritma piramida.

II. METODE PENELITIAN

Dari 20 sampel masing-masing untuk serbuk rimpang jahe dan 40 sampel serbuk rimpang temulawak dengan FTIR dihasilkan data spektra absorban yang diamati 1866 titik, pada bilangan gelombang 4000–400 cm^{-1} yang mencerminkan kadar gingerol. Karena jumlah sampel dipandang mencukupi maka sampel-sampel dibagi menjadi 2 kelompok. Untuk rimpang jahe 15 sampel untuk kalibrasi dan 5 sampel untuk validasi, sedangkan untuk temulawak 30 sampel untuk kalibrasi dan 10 sampel sisanya untuk validasi. Pemilihan pada masing-masing kelompok dilakukan secara acak. Karena metode *wavelet* mensyaratkan jumlah titik harus 2^M , untuk M bilangan bulat positif, maka dari 1866 titik diambil 1024 titik dengan memperhatikan daerah identifikasi spektra infra merah gingerol yang memberikan informasi.

Dari 1024 titik yang terpilih dilakukan transformasi *wavelet* diskret (TWD), dengan melihat berbagai kemungkinan resolusi yang menghasilkan koefisien-koefisien *wavelet* yang jumlahnya lebih kecil dari jumlah sampel untuk kelompok data kalibrasi, serta berbagai fungsi *mother wavelet* keluarga Daubechies. Alasan pemilihan *mother wavelet* keluarga Daubechies karena sering dipakai dalam aplikasi dan memberikan hasil pemodelan yang baik [8], [9], [10]. Koefisien-koefisien *wavelet* yang dihasilkan digunakan untuk pengembangan model kalibrasi peubah ganda. Perhitungan matriks *wavelet* pada penelitian ini menggunakan *software wawetresh 3* seperti yang diterangkan pada referensi [6], [7]. Karena lama penyimpanan berpengaruh terhadap kadar gingerol yang dihasilkan, maka dalam pencarian model prediksi yang lebih baik diikuti peubah *dummy* yang mencerminkan kelompok lama penyimpanan (untuk data pada penelitian ini 3 bulan dikode 0 dan 10 bulan dikode 1).

Secara matematis TWD tidak menjamin bahwa antara koefisien-koefisien *wavelet* yang dihasilkan tidak saling berkorelasi, sehingga metode selanjutnya dalam penelitian ini digunakan *Ridge Regression* dan Regresi Komponen Utama untuk menghilangkan kasus multikolinearitas yang selanjutnya disebut TWD-Ridge dan TWD-PCR.

Langkah-langkah analisis dengan metode TWD-Ridge dan TWD-PCR dapat dijelaskan sebagai berikut :

Dari data spektrum absorban dapat dituliskan matriks $X_{(n \times p)}$, dimana n adalah banyaknya sampel dan p adalah banyaknya titik absorban yang diteliti pada masing-masing bilangan gelombang. Konsentrasi senyawa aktif dituliskan dalam $\underline{y}_{(n \times 1)}$. Misalkan matriks X berukuran $(n \times p)$ dan X yang terkoreksi terhadap nilai rata-ratanya adalah seperti dinyatakan dalam persamaan (8).

$$\mathcal{X} = X - \frac{1}{n} \bar{x} \bar{x}^T = \begin{bmatrix} \underline{x}_1^T \\ \underline{x}_2^T \\ \vdots \\ \underline{x}_n^T \end{bmatrix} \quad (8)$$

Dengan menggunakan transformasi *wavelet* diskret $\underline{x}_j^T = \underline{d}_j^T W$ dan W ditentukan oleh *mother wavelet* tertentu, maka diperoleh matriks koefisien *wavelet* $D_{(n \times p)} = \mathcal{X}_{(n \times p)} W^T_{(p \times p)}$. Kemudian dengan memilih level-level resolusi tertentu yang jumlah koefisien *wavelet* yang dihasilkan lebih kecil dari $n-1$, maka akan diperoleh $D^*_{(n \times m)} = \mathcal{X}_{(n \times p)} W^{*T}_{(p \times m)}$, yang mereduksi pengamatan dari p titik tiap-tiap sampel menjadi m titik koefisien *wavelet* yang terpilih.

Persamaan regresi antara $\underline{y}_{(n \times 1)}$ terhadap $D^*_{(n \times m)}$ dapat ditulis seperti dalam persamaan (9) berikut.

$$\underline{y} = \frac{1}{n} q_0 + D^* \underline{q} + \underline{e} \quad (9)$$

Jika $m < n$ maka dugaan kuadrat terkecil dari persamaan (9) adalah :

$$\hat{\underline{q}} = (D^{*T} D^*)^{-1} D^{*T} \underline{y}$$

Persamaan regresi prediksi linear yang berbentuk seperti persamaan (10):

$$\hat{y}_{pred} = b_0 + b_1 x_1 + \dots + b_p x_p = b_0 + \underline{x}^T \underline{b}$$

dapat dihitung dengan menggunakan $\underline{b} = W^* \hat{\underline{q}}$ dan

$$b_0 = \bar{y} - \bar{x}^T \underline{b}$$

Jika multikolinearitas masih terjadi antar koefisien *wavelet*, yang terdeteksi dengan statistik *Variance Inflation Factors* (VIF) lebih besar 10, maka langkah yang bisa diambil adalah menghitung skor komponen utama dari D^* . Kemudian meregresikan ulang antara $\underline{y}_{(n \times 1)}$ dengan skor-skor komponen utama. Pemilihan model terbaik dapat dilakukan dengan memperhatikan beberapa ukuran kebaikan model prediksi seperti R^2 dan S.

Hasil persamaan \hat{y}_{pred} dalam persamaan (10) akan digunakan untuk memprediksi kadar senyawa aktif kelompok sampel data validasi. RMSEP (*Root Mean Square Error of prediction*) merupakan salah satu ukuran yang dapat digunakan sebagai ukuran kebaikan hasil prediksi [9], [10], [11] seperti dinyatakan dalam persamaan (11).

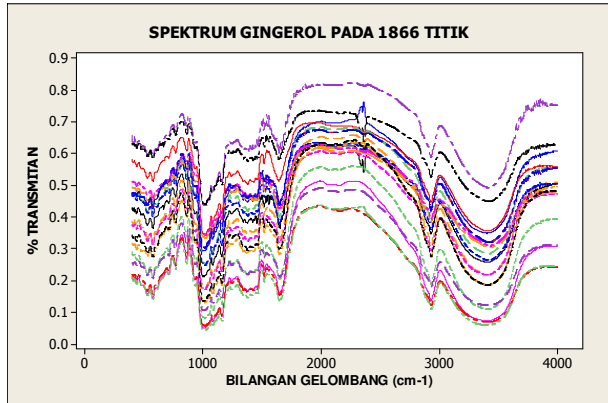
$$RMSEP = \sqrt{\frac{\sum_{i=1}^{N_{pred}} (\hat{y}_{i pred} - y_i)^2}{N_{pred}}} \quad (11)$$

N_{pred} adalah banyaknya sampel untuk validasi.

Semakin kecil RMSEP, semakin baik prediksi model yang dihasilkan.

III. ANALISIS DAN PEMBAHASAN

Gambar spektra absorban serbuk jahe pada 1866 titik dan 1024 titik terpilih, untuk 20 sampel serbuk rimpang jahe bisa dilihat pada Gambar 1 dan Gambar 2. Dengan mengambil 11 koefisien *wavelet* (untuk *mother wavelet* Daubechies - 20) pada resolusi 0, 1 dan 3 serta 1 koefisien untuk fungsi skala hasil transformasi *wavelet* diskret dilakukan pencarian model terbaik untuk prediksi kadar gingerol. Alasan diambil *mother wavelet* Daubechies- 10 dan level resolusi 0, 1 dan 3, karena perilakunya yang relatif lebih baik dibanding yang lain, dalam arti dapat menangkap ukuran-ukuran kebaikan model seperti R^2 dan S yang relatif lebih baik.

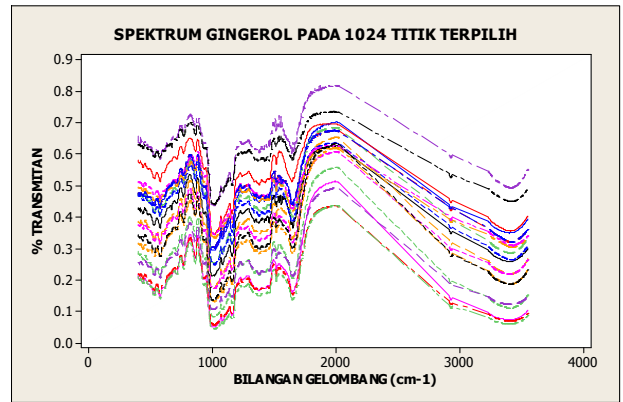


Gambar 1. Spektra absorban 1866 titik, untuk 20 sampel serbuk rimpang jahe

Permasalahan yang timbul ternyata korelasi antar koefisien *wavelet* masih banyak yang tinggi. Sehingga pemodelan dengan regresi ganda biasa dengan peubah respon kadar gingerol dan peubah bebas koefisien *wavelet* menjadi kurang valid, karena masih terjadi kasus multikolinearitas. Hal ini terlihat dari analisis regresi antara kadar gingerol dengan 12 koefisien *wavelet* dan peubah *dummy* waktu penyimpanan diperoleh R^2 yang tinggi (99,9%) sedang semua peubah bebas tidak signifikan, selain itu nilai VIF (*Variance Inflation Factor*) masing-masing peubah berkisar antara 43,4-681,2. Metode yang digunakan untuk mengatasi kasus multikolinear antar koefisien *wavelet* dalam penelitian ini adalah *Ridge Regression* dan Regresi Komponen Utama. Hasil ukuran-ukuran kebaikan model terbaik untuk TWD-Ridge dan TWD-PCR dapat dilihat pada Tabel 1. Dari Tabel 1 ternyata dengan PCR 7 komponen pertama dari 12 komponen serta satu peubah *dummy* menghasilkan model yang relatif lebih baik dibandingkan *Ridge Regression*.

Hasil dari TWD-Ridge dan TWD-PCR digunakan untuk menduga 5 sampel kelompok data validasi. Ringkasan prediksi untuk kelompok sampel data kalibrasi dan kelompok sampel data validasi model TWD-PCR dapat dilihat pada Tabel 2 sedangkan untuk TWD-Ridge disajikan pada Tabel 3.

Plot Y dengan \hat{Y} untuk kelompok kalibrasi (15 sampel) dapat dilihat pada Gambar 3. Sedangkan plot antara Y dengan \hat{Y} untuk kelompok sampel data validasi dapat dilihat pada Gambar 4.



Gambar 2 Spektra absorban 1024 titik, untuk 20 sampel serbuk rimpang jahe

TABEL 1
NILAI KEBAIKAN MODEL GINGEROL DENGAN TWD-RIDGE DAN TWD-PCR

Metode	R^2	RMSEP
TWD-Ridge (k=0,02)	94,15%	0,1114
TWD-PCR (PC1-PC7)	96,70%	0,1072

Dari Tabel 1, Tabel 2, Tabel 3, Gambar 3 dan Gambar 4 menunjukkan bahwa model terbaik untuk prediksi gingerol diperoleh dari TWD-PCR dengan *mother wavelet* Daubechies - 10 dan level resolusi 0, 1 dan 3. Ukuran kebaikan model yang diperoleh adalah $R^2 = 96,7\%$ dan $RMSEP = 0,1072$. Hasil analisis dengan TWR-PCR menunjukkan hasil yang lebih memuaskan dibanding TWD-Ridge.

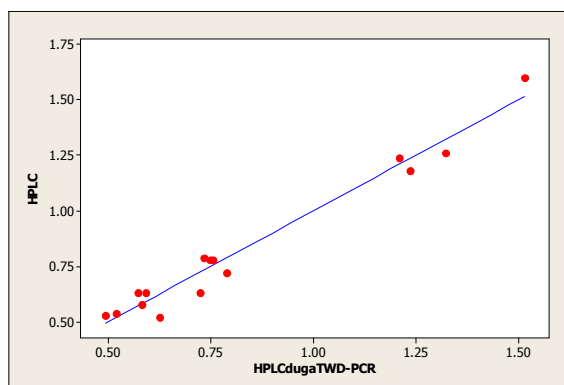
TABEL 2
NILAI Y DAN \hat{Y} KADAR GINGEROL DENGAN TWD-PCR

Kelompok Data Kalibrasi		Kelompok Data Validasi	
Kadar Gingerol dari HPLC (%)	Dugaan (%)	Kadar Gingerol dari HPLC (%)	Dugaan (%)
0.63	0.723	0.53	0.583
0.72	0.789	0.78	0.755
0.58	0.582	0.79	0.715
0.53	0.494	1.14	1.304
0.52	0.624	1.07	1.217
0.54	0.520		
0.79	0.733		
0.78	0.755		
0.63	0.572		
0.63	0.592		
0.78	0.748		
1.26	1.322		
1.60	1.513		
1.18	1.236		
1.24	1.209		

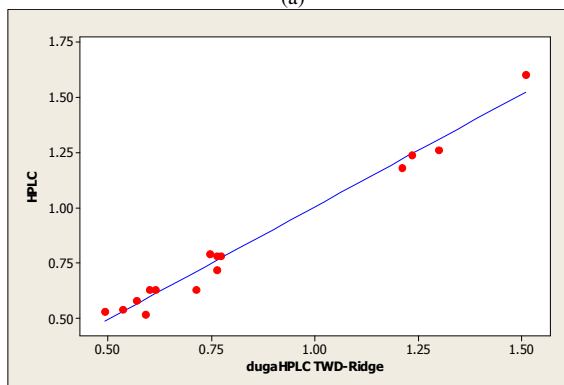
TABEL 3

NILAI Y DAN \hat{Y} KADAR GINGEROL DENGAN TWD-RIDGE ($k=0,02$)

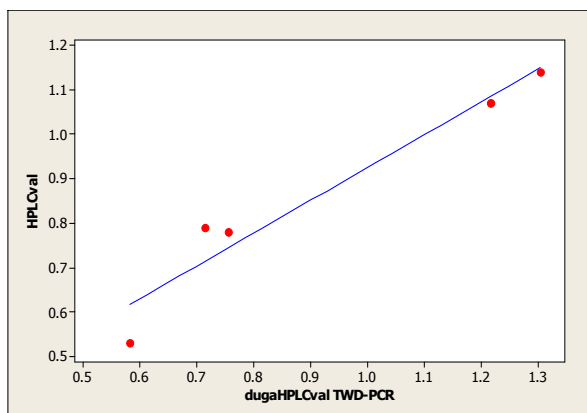
Kelompok Data Kalibrasi		Kelompok Data Validasi	
Kadar Gingerol dari HPLC (%)	Dugaan (%)	Kadar Gingerol dari HPLC (%)	Dugaan (%)
0.63	0.713	0.53	0.516
0.72	0.763	0.78	0.762
0.58	0.568	0.79	0.720
0.53	0.493	1.14	1.310
0.52	0.589	1.07	1.237
0.54	0.536		
0.79	0.747		
0.78	0.762		
0.63	0.600		
0.63	0.613		
0.78	0.772		
1.26	1.300		
1.60	1.509		
1.18	1.211		
1.24	1.235		



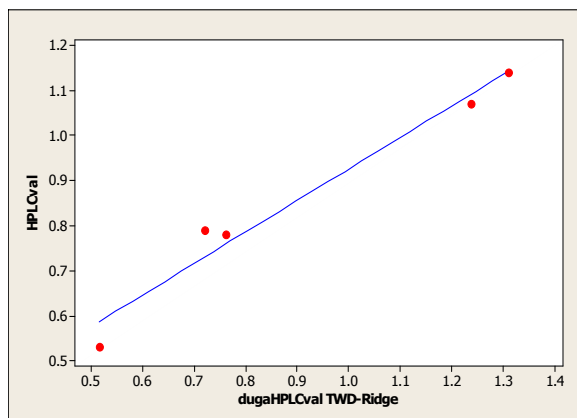
(a)



(b)

Gambar 3. Plot Y dengan \hat{Y} kelompok data kalibrasi gingerol dengan TWD-PCR (a) TWD-Ridge (b)

(a)



(b)

Gambar 4. Plot Y dengan \hat{Y} kelompok data validasi gingerol dengan TWD-PCR (a) TWD-Ridge (b)

IV. KESIMPULAN

Metode Transformasi *Wavelet* Diskret (TWD) mampu melakukan reduksi dimensi dengan baik, tetapi tidak ada jaminan, bahwa kasus multikolinearitas teratasi. Sehingga TWD sebaiknya digabung dengan metode lain yang mampu mengatasi multikolinearitas dalam pembuatan model kalibrasi peubah ganda.

Gabungan Transformasi *wavelet* Diskret (TWD) dengan *Principal Component Regression*, untuk menduga model prediksi kadar senyawa gingerol pada rimpang jahe, ternyata menghasilkan model yang cukup memuaskan. Hasil analisis dengan TWR-PCR ini lebih memuaskan dibanding gabungan Transformasi *wave-*

let Diskrit (TWD) dengan *Ridge Regression* (TWD-Ridge).

V. DAFTAR PUSTAKA

- [1] Naes, T., Isaksson, T., Fearn, T., Davies, "A User Friendly Guide to Multivariate Calibration and Classification," *NIR publications*, UK. 2002.
- [2] Martens, H., Naes, T., *Multivariate Calibration*. John Wiley & Sons. Chichester, England. 1989.
- [3] Sunaryo, S., Notodiputro, K.A., "Penerapan Metode Transformasi *Wavelet* Diskret untuk Menentukan Kadar Senyawa Kurkuminoid pada Rimpang Temulawak," *Prosiding Seminar Nasional Matematika 2005*, halaman 100-107, Jurusan Matematika UNS, Surakarta, 7 Mei 2005.

- [4] Sunaryo, S., "Model Kalibrasi dengan Transformasi Wavelet sebagai Metode Pra-Pemrosesan," Disertasi Program Pascasarjana, Institut Pertanian Bogor. 2006.
- [5] Sunaryo, S., Notodiputro, K.A., "Penerapan Metode Transformasi Wavelet Diskret untuk Menentukan Kadar Senyawa Gingerol pada Rimpang Jahe." *Statistika - Forum Teori dan Aplikasi Statistika* 4: 181-185, Jurusan Statistika FMIPA UNISBA. 2004.
- [6] Nason, G.P., Silverman, B.W., "The discrete wavelet transform in S." *J.comp graph. Stat.* 3: 163-191. 1994.
- [7] Nason, G.P. (1998), *Wavethresh 3 software*. Department of Mathematics, University of Bristol, UK. Available: [<http://www.stats.bris.ac.uk/~wavethresh>] (tanggal akses : 12 April 2007).
- [8] Brown, P.J., Fearn, T., Vanucci, M., "Bayesian Wavelet Regression on Curves with Application to a Spectroscopic Calibration Problem." *J Amer Statist Assoc* 96: 398-408. 2001.
- [9] McNulty, S.C., Ganapati, M., "Application of Wavelet Analysis Determining Glucose Concentration of Aqueous Solution Using NIR Spectroscopy." *Hewlett-Packard comp.* 1998.
- [10] Yi-yu, C., Chen M.J., "A New Computing Multivariate Spectral Analysis Method Based on Wavelet Transform." *Journal of Zhejiang University Science* 1: 15-19. 2000.
- [11] Shao, X., Yadong, Z., "Determining of Chlorogenic Acid in Plant Samples by Using Near-Infrared Spectrum with Wavelet Transform Preprocessing." *Analytical Sciences* 20. 2004.
- [12] Vidacovic, B., Meuller, P., "Wavelets for Kids. A Tutorial Introduction", *AMS Subject Classification*, Duke University 1991.
- [13] Percival, D.B. (2005), *Wavelets : Data Analysis, Algorithms and Theory*. University Washington. Available: [<http://www.ms.washington.edu/~s530/>] (tanggal akses : 18 april 2005).