# MODEL KELUARGA SPLINE POLINOMIAL *TRUNCATED* DALAM REGRESI SEMIPARAMETRIK

## (Model of Truncated Polynomial Spline family in Semiparametric Regression)

**I Nyoman Budiantara**

Jurusan Statistika FMIPA-ITS
Kampus ITS, Keputih, Sukolilo, Surabaya

### ABSTRAK

**Diberikan data berpasangan $(x_i, t_i, y_i)$ dan hubungan antara $x_i$, $t_i$, dan $y_i$ diasumsikan mengikuti model regresi semiparametrik $y_i = x_i' b + m(t_i) + e_i, t_i \in [a,b]$, $i$ = 1,2,...,n. Variabel respon $y_i$ berhubungan parametrik dengan variabel prediktor $x = (x_1, x_2,...,x_p)'$ dan berhubungan nonparametrik dengan variabel prediktor $t_i$. Bentuk kurva regresi $m$ diasumsikan tidak diketahui, termuat di dalam ruang *Sobolev* $W_2^q[a,b]$ dan $b = (b_0, b_1,..., b_p)' \in R^{p+1}$ parameter yang tidak diketahui. Estimasi kurva regresi $\hat{m}$ dan $\hat{b}$ diperoleh dari meminimumkan *Penalized Least Square* (PLS) :**

$$n^{-1}\sum_{i=1}^{n}(y_i - x_i'b - m(t_i))^2 + g\int_a^b (m^{(q)}(t))^2\,dt \text{, untuk setiap } m \in W_2^q[a,b],$$

**dengan $g > 0$ parameter penghalus. Beberapa metode seperti *Reproducing Kernel Hilbert Space* (RKHS) dan *Gateaux* sering digunakan untuk menyelesaikan optimasi PLS. Walaupun metode RKHS dan *Gateaux* memiliki kelebihan, tetapi metode ini memerlukan pengetahuan matematika yang relatif tinggi dan memiliki interprestasi Statistik khusus, yang sulit dipahami oleh banyak pengguna Statistika. Dalam tulisan ini, diberikan pendekatan keluarga spline polinominal *truncated* untuk mengestimasi $\hat{m}$ dan $\hat{b}$. Pendekatan ini mempunyai interprestasi Statistik yang mudah dan sederhana, disamping tidak memerlukan pengetahuan matematika yang tinggi. Selanjutnya, diberikan aplikasi model keluarga spline polinominal *truncated* untuk menduga pola hubungan semiparametrik antara produksi Billet pada suatu perusahaan besi baja dengan variabel proses produksi dan bahan baku Billet yang berpola parametrik linear, serta dengan lama waktu mengerjakan Billet yang berpola nonparametrik.**

***Kata kunci:*** Spline, Keluarga Spline, Regresi Semiparametrik, regresi Nonparametrik.

### Abstract

**We are given data points $(x_i, t_i, y_i)$ and the relation between $x_i$, $t_i$, and $y_i$ assumed to follow semiparametric regression model $y_i = x_i' b + m(t_i) + e_i, t_i \in [a,b]$, $i$ = 1,2,...,n. Response variable $y_i$ has parametric relation with predictor variable $x = (x_1, x_2,...,x_p)'$ and non-parametric relation with predictor variable $t_i$. Regression curve $m$ is assumed unknown in Sobalev space $W_2^q[a,b]$ and $b = (b_0, b_1,..., b_p)' \in R^{p+1}$ is the unknown parameter. This estimation of regression curve $\hat{m}$ dan $\hat{b}$ is obtained by minimizing *Penalized Least Square* (PLS) :**

$$n^{-1}\sum_{i=1}^{n}(y_i - x_i'b - m(t_i))^2 + g\int_a^b (m^{(q)}(t))^2\,dt,$$ **for each $m \in W_2^q[a,b]$, precision parameter. Some methods such as RKHS (*Reproducing Kernel Hilbert Space)* and Gateaux are frequently**

**used to solve the optimization of PLS. Those methods have advantages but need mathematical knowledge and special statistical interpretation that trouble common statistics users. In this article we use the approximation of family of truncated polynomial splines to estimate $\hat{m}$ and $\hat{b}$. This approach has easy and simple statistical interpretation and does not need extensive mathematical knowledge. The model of truncated polynomial of semiparametric relation between Billet production in steel factory with production process variable and Billet raw material with linear parametric pattern, and with period of Billet processing with non-parametric pattern.**

Keywords: Spline, spline family, semiparametric regression, non-parametric regression.

## 1. PENDAHULUAN

Dalam analisis regresi pola hubungan antara dua variabel atau lebih, tidak selalu berpola parametrik seperti linear, kuadratik, kubik dan yang lainnya. Terdapat banyak kasus dimana pola hubungan antara variabel berpola non-parametrik. Bahkan dalam beberapa kasus yang lain, sering hubungan antara variabel berpola semiparametrik (Eubank, 1988). Dalam regresi parametrik bentuk kurva regresi diasumsikan diketahui. Untuk dapat menggunakan pendekatan regresi parametrik, diperlukan pengetahuan masa lalu tentang karakteristik data yang dapat diselidiki. Berbeda dengan pendekatan parametrik, dalam regresi nonparametrik bentuk kurva regresi diasumsikan tidak diketahui. Kurva regresi hanya diasumsikan mulus *(smooth)* dalam arti termuat di dalam suatu ruang fungsi tertentu. Data diharapkan mencari sendiri bentuk estimasinya, tanpa dipengaruhi oleh faktor subyektifitas dari perancang penelitian (Eubank, 1988). Dengan demikian, pendekatan regresi nonparametrik memiliki fleksibelitas yang tinggi.

Regresi semiparametrik merupakan gabungan antara regresi parametrik dan regresi nonparametrik. Model regresi semiparametrik mulai berkembang pesat sejak Wahba pada tahun 1990 mempublikasikan suatu model hubungan semiparametrik dari penggunaan listrik pada negara bagian di Amerika Serikat, yang berpola parametrik linear dengan pendapatan dan berpola nonparametrik dengan temperatur. Saat ini sudah banyak para peneliti yang memfokuskan penelitiannya dalam regresi semiparametrik, seperti He dan Shi (1996), Shi dan Li (1994), Subanar dan Budiantara (1999; 2001). Misalkan diberikan data berpasangan ($x_i$, $t_i$, $y_i$) dan hubungan antara $x_i$, $t_i$ dan $y_i$ diasumsikan mengikuti model regresi semiparametrik:

$$y_i = x_i' b + m(t_i) + e_i,\ t_i \in [a,b],\ i = 1,2,\ldots,n \qquad (1)$$

Variabel respon $y_i$ berhubungan parametrik dengan variabel prediktor $x = (x_1, x_2,\ldots,x_p)'$ dan berhubungan nonparametrik dengan variabel prediktor $t_i$. Kurva regresi $m$ diasumsikan termuat di dalam ruang *Sobolev* $W_2^q[a,b]$, dengan :

$$W_2^q[a,b] = \left\{ h; \int_a^b \left(m^{(q)}(t)\right)^2 dt < \infty \right\}$$

Parameter $b = (b_0, b_1, \ldots, b_p)' \in R^{p+1}$ tidak diketahui. Tujuan utama dalam analisis regresi semiparametrik adalah mendapatkan estumasi untuk kurva regresi $\hat{m}$ dan $\hat{b}$. Estimasi ini diperoleh dari meminimumkan *Penalized Least Square* (PLS) :

$$n^{-1}\sum_{i=1}^{n} \left(y_i - x_i' b - m(t_i)\right)^2 + g\int_a^b \left(m^{(q)}(t)\right)^2,$$

*untuk setiap* $m \in W_2^q[a,b]$. (2)

Jika diperhatikan secara detail prosedur estimasi dalam regresi nonparametrik dan semiparametrik, seperti Kernel (Hardle, 1990), Deret Ortogonal (Eubank, 1988), Fourier dan Wavelets (Antoniadis, 1994), maka terlihat dengan jelas bahwa pendekatan-pendekatan tersebut diperoleh dengan memodelkan (menghampiri) kurva regresi yang tidak diketahui dengan suatu fungsi tertentu (fungsi keluarga) dari masing-masing pendekatan. Selanjutnya dilakukan optimasi *Least Square* (LS) untuk memperoleh estimasinya.

Memperhatikan prosedur estimasi yang dilakukan oleh peneliti-peneliti di atas, dalam tulisan ini, diberikan suatu pendekatan keluarga spline, yaitu spline polinominal *truncated*.Dengan pendekatan ini diharapkan diperoleh perhitungan matematik dan interprestasi Statistik yang relatif mudah dan sederhana. Proses inferensi didasarkan pada polinominal dan optimasi yang digunakan adlah LS (tidak optimasi PLS). Selanjutnya, pada bagian 2 tulisan ini, disajikan suatu optimasi untuk keluarga spline polinominal *truncated*. Sedangkan pada bagian 3, diberikan aplikasi model spline polinominal *truncated* pada data produksi Billet dari suatu perusahaan besi baja, yang berpola semiparametrik.

## 2. OPTIMASI KELUARGA SPLINE POLINOMINAL *TRUNCATED*

Sebelum menyajikan pendekatan keluarga spline Polinominal *truncated* dengan menggunakan optimasi LS, terlebih dahulu diberikan model umum untuk regresi spline. Diberikan model regresi semiparametrik (umum) :

$$y_i = L_i m + x_i' b + e_i, \; i = 1,2,..., n \qquad (3)$$

$L_i$ merupakan fungsional linear terbatas pada suatu ruang *Hilbert H*. Sesatan random $e_i$ berdistribusi independen dengan mean nol dan variasi $s^2$. Kurva regresi $m \in H$ dan parameter $b = (b_0, b_1,..., b_p)' \in R^{P+1}$ tidak diketahui. Karena $H$ ruang *Hilbert* maka $H$ dapat didekomposisi menjadi *direct sum* dari dua ruang, yaitu ruang $N$ dan $M$ (Kreyzsig, 1978) : $H = N \oplus M$, dengan ruang *null* dan $M = N^{\perp}$.

Untuk setiap $m \in H$ dapat dinyatakan menjadi $m = u + v,\; u \in N$ *dan* $v \in M$. Berdasarkan teorema *representasi Riersz* (Kreyzsig, 1978), terdapat dengan tunggal *representer* $g_i \in H$, sehingga :

$$L_i m = < g_i, m >, m \in H, i = 1,2,..., n$$

Jika ditulis $z_i = y_i - x_i' b$ maka model regresi (3) dapat ditulis menjadi :

$$z_i = < g_i, m > + e_i,\; i = 1,2,..., n$$

Estimasi $m$ dieroleh berdasarkan optimasi (Budiantara, 2000) :

$$\underset{m\in H}{Min}\left\{ n^{-1}\sum_{i=1}^{n}(z_i - < g_i, m>)^2 + g\|pm\|^2 \right\} = \underset{m\in H}{Min}\left\{ n^{-1}Q_1(m) + gQ_2(m) \right\} \qquad (4)$$

dengan

$$Q_1(m) = \sum_{i=1}^{n}(z_i - < g, m >)^2,\; Q_2(m) = \|Pm\|^2$$ dan $P$ proyeksi ortogonal $H$ onto $M$.

Penyelesaian optimal (4) adalah suatu fungsi yang merupakan anggota ruang $M$. Dengan sedikit penjabaran didapat :

$$Q_1(m) = \sum_{i=1}^{n}(z_i - < g_i, m >)^2$$

$$= \sum_{i=1}^{n}(z_i - < g_i, v >)^2$$

$$Q_2(m) = \|Pm\|^2$$

$$= \|Pu\|^2 + \|Pv\|^2 \geq \|Pv\|^2$$

Akibatnya untuk setiap $m \in H$ berlaku :

$$n^{-1}\sum_{i=1}^{n}(z_i - < g_i, m >)^2 + g\|Pm\|^2 \geq n^{-1}\sum_{i=1}^{n}(z_i - < g_i, v >)^2,\; v \in M$$

Ini berarti fungsi anggota $H$ yang meminimumkan optimasi (4) merupakan anggota ruang M.

Beberapa metode untuk menyelesaikan optimasi khusus (4) dalam regresi nonparametrik, telah diberikan oleh beberapa penulis seperti Craven dan Wahba (1979), Budiantara, dkk., (1997), dengan menggunakan RKHS. Metode yang sama juga telah diberikan untuk memperoleh estimator regresi semiparametrik spline parsial (Subanar dan Budiantara, 1999; wahba, 1990; Budiantara, 1999). Disamping itu Eubank (1988) memberikan metode *Gateaux* untuk memperoleh estimasi kurva regresi nonparametrik spline maupun estimator regresi semiparametrikspline parsial. Walaupun RKHS dan *Gateaux* mempunyai beberapa kelebihan, tetapi tidak dapat dihindari metode ini juga terdapat kelemahan-kelemahan seperti memerlukan pengetahuan matematik khusus (analisis Real dan analisis Fungsional) yang relatif tinggi. Karena alasan ini, He dan Shi (1996) dan Shi dan Li (1994) mencoba memberikan pendekatan keluarga B-Spline untuk memperoleh estimasi kurva regresi parsial. Tetapi hasil estimasi dengan B-Spline, juga mempunyai beberapa kelemahan, seperti kurang jelasnya interpretasi Statistik dan tidak menggambarkan secara visual perubahan perilaku kurva pada interval yang berbeda sebagai ciri khas dari pendekatan spline.

Berikut diberikan pendekatan keluarga spline lain, yaitu polinominal spline *truncated*. Dengan pendekatan ini diharapkan kelemahan - kelemahan di atas dapat diperbaiki. Diberikan suatu bentuk basis untuk ruang Spline (Schumaker, 1981) berbentuk :

$$\left\{1, t, \ldots, t^{q}, (t-K_1)^{q} I(t \geq K_1), \ldots, (t-K_r)^{q} I(t \geq K_r)\right\},$$

dengan $I$ merupakan fungsi indikator :

$$I_D(x) = \begin{cases} 1, x \in D \\ 0, x \in D \end{cases}$$

dan $K_1, K_2, \ldots, K_r$ titik-titik knots. Titik knots merupakan titik perpaduan bersama yang memperlihatkan terjadinya perubahan perilaku dari fungsi spline pada interval-interval yang berbeda. Untuk setiap fungsi $m$ dalam ruang ini dapat dinyatakan sebagai kombinasi linear

$$m(t) = q_0 + q_1 t + q_2 t^2 + \ldots + q_q t^q + f_1 (t-K_1)^q I(t \geq K_1) + \ldots + f_1 (t-K_r)^q I(t \geq K_r),$$

untuk $f_j, j = 0,1,\ldots, q$ dan $f_k, k = 1,2,\ldots, r$ konstanta bernilai real.

Model regresi (1) dapay dinyatakan menjadi :

$$z_i = \sum_{j=0}^{q} q_j t_i^j + \sum_{k=1}^{r} f_k (t_i - K_k)^q I(t_i \geq K_k) + e_i$$

untuk $z_i = y_i - x_i' b$ diperoleh model regresi :

$$y_i = x_i' b + \sum_{j=0}^{q} q_j t_i^j + \sum_{k=1}^{r} f_k (t_i - K_k)^q I(t_i \geq K_k) + e_i, i = 1, 2, \ldots n.$$

Dengan penyajian matriks, diperoleh model regresi :

$$y = B(x,t) q^* + e,$$

dengan $q^* = (b, q, f)'$ dan *B(x,t)* matriks berukuran *nx(p+q+r+2)* yang bergantung pada $x$ dan $t$. Estimasi $b = (b_0, b_1, \ldots, b_p)', q = (q_0, q_1, \ldots, q_q)'$ dan $f = (f_1, \ldots, f_r)'$ diperoleh dari menyelasaikan optimasi LS :

$$\underset{b \in R^{p+1}; q \in R^{q+1}; f \in R^r}{Min} (e'e) = \underset{b \in R^{p+1}; q \in R^{q+1}; f \in R^r}{Min} \left( (y - B(x,t) q^*)' (y - B(x,t) q^*) \right) \quad (5)$$

Karena *B(x,t)* matriks *rank* penuh maka dengan sedikit penjabaran dan menggunakan derivatif parsial diperoleh estimasi parameter $q^*$ :

$$\hat{q}^* = (\hat{b}, \hat{q}, \hat{f})' = (B'(x,t) B(x,t))^{-1} B'(x,t) y.$$

Jadi estimasi kurva regresi semiparametrik, diberikan oleh :

$$\hat{E}(y | x,t) = B(x,t)(B'(x,t) B(x,t))^{-1} B'(x,t) y$$
$$= A(x,t) y \quad (6)$$

dengan $A(x,t) = B(x,t)(B'(x,t) B(x,t))^{-1} B'(x,t)$. Matriks *A(x,t)* mempunyai peran yang sangat penting dalam inferensi regresi semiparametrik spline. Berdasarkan uraian diatas, diperoleh pernyataan berikut :

(i). Optimasi yang digunakan untuk mengestimasi model keluarga polinomial *truncated* adalah optimasi LS (tidak PLS), sehingga secara matematik mudah dan sederhana.

(ii). Basis spline yang digunakan berupa polinominal *truncated* yang memuat titik-titik knots. Akibatnya secara visual dapat digambarkan secara jelas perubahan perilaku dari model keluarga spline pada interval-interval yang berbeda, sebagai ciri khas pendekatan spline.

(iii).Pendekatan dengan polinominal mempunyai sifat matematik dan Statistik yang baik.

(iv).Estimator keluarga spline polinominal *truncated* merupakan estimator linear dalam observasi. Sifat linear ini sangat membantu dalam pembangunan inferensi untuk model keluarga spline polinominal *truncated* (Budiantara, 2001).

(v). Estimator keluarga spline polinominal *truncated* merupakan estimator bias, sebab :

$$
\begin{aligned}
E\left(\hat{E}(y|x,t)\right) &= A(x,t)E(y) \\
&= A(x,t)[xb + m(t)] \\
&\neq m(t) + xb
\end{aligned}
$$

Walaupun estimator ini bias, tetapi tidak bias asimotik (Budiantara, 2001).

## 3. APLIKASI MODEL KELUARGA SPLINE POLINOMINAL *TRUNCATED*

Pada bagian ini disajikan aplikasi model semiparametrik dari keluarga spline polinominal *truncated*. Diberikan data produksi Billet (dalam kg) merupakan produk besi batangan dari suatu perusahaan besi baja *(y)*. Produksi Billet ini dipengaruhi oleh beberapa variabel independen, diantaranya HEAT ($x_h$) proses tertentu dalam produksi Billet, SCRAP ($x_s$) banyak bahan baku (dalam kg) untuk produksi Billet POT ( $x_p$ ) lama waktu (dalam detik) pengerjaan Billet dengan menggunakan listrik. Ingin diestimasi model pola hubungan antara produksi Billet dengan variabel independen HEAT, SCRAP dan POT. Diambil sampel random sebanyak 64 produksi Billet selama satu minggu. Hubungan ke-empat variabel diberikan oleh model :

$$y_i = f(x_{hi}, x_{si}, x_{pi}) + e_i,\ i = 1,2,\ldots,64.$$

Plot antara variabel $y_i$ dengan $x_{pi}$ disajikan dalam Gambar 1(a), sedangkan plot antara variabel $y_i$ dengan $x_{hi}$, dan variabel $y_i$ dengan $x_{si}$, masing-masing disajikan dalam Gambar 2(a) dan Gambar 3(a). Terlihat dari Gambar (2) dan Gambar 3(a) bahwa hubungan antara $y$ dengan $x_h$ cenderung linear. Demikian pula dengan hubungan antara $y$ dengan $x_s$ juga cenderung linear. Tetapi plot Gambar 1(a) memperlihatkan tidak adanya kecenderungan pola hubungan yang jelas antara $y$ dengan $x_p$, sehingga dalam hal ini sulit untuk dimodelkan secara parametrik. Berdasarkan plot ini, dicoba model hubungan antara variabel respon $y_i$ dengan $x_{hi}$, dan variabel $y_i$ dengan $x_{pi}$, $x_{hi}$, $x_{si}$ menggunakan model semiparametrik, dimana variabel $y_i$ dengan $x_{pi}$ berhubungan nonparametrik. Model regresi semiparametrik dinyatakan sebagai :

$$y_i = b_1 x_{hi} + b_2 x_{si} + m(x_{pi}) + e_i,\ i = 1,2,\ldots,64.$$

Kurva $m$ didekati dengan model keluarga spline polinominal *truncated* kuadratik dengan dua titik knots, yaitu $K_1$ dan $K_2$ ($K_1$ = $K_2$) :

$$m(x_p) = \sum_{j=1}^{2} q_j x_p^j + f(x_p - K_1)^2 I(x_p \geq K_1) + f_2(x_p - K_2)^2 I(x_p \geq K_2)$$

Untuk mendapatkan spline optimal perlu dipilih letak titik-titik knots $K_1$ dan $K_2$ yang optimal dengan menggunakan *generalized Cross Validation* (GCV) (Budiantara,2000). Fungsi GCV diberikan oleh :

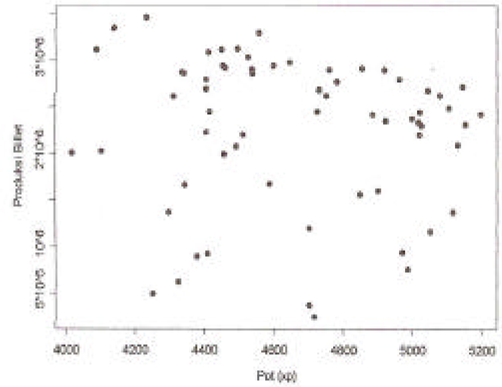
$$G(K_1,K_2) = \frac{n^{-1} y'[I - A(x_h, x_s, x_p\{K_1,K_2\})]'[I - A(x_h, x_s, x_p\{K_1,K_2\})]y}{\left(n^{-q}\, trace[I - A(x_h, x_s, x_p\{K_1,K_2\})]\right)^2}$$

Model spline optimal berkaitan dengan nilai fungsi $G$ yang terkecil. Tabel 1 menyajikan ringkasan nilai GCV untuk beberapa titik-titik knots. Terlihat dari Tabel 1, nilai GCV terkecil adalah $G_{opt}(K_1,K_2)$ = 2511300803, bersesuaian dengan titik-titik knots :

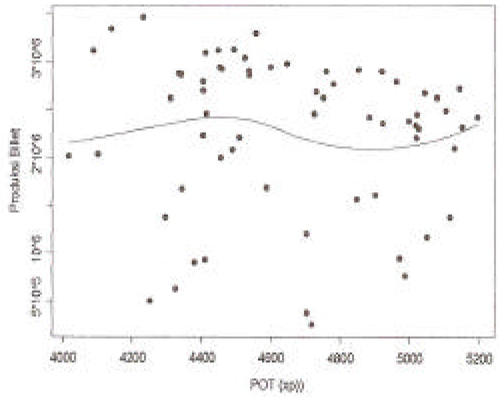
$K_1$ = 4369 dan $K_2$ = 4620

**Tabel 1. Nilai GCV model keluarga spline polinominal *truncated* dari berbagai titik-titik knots**

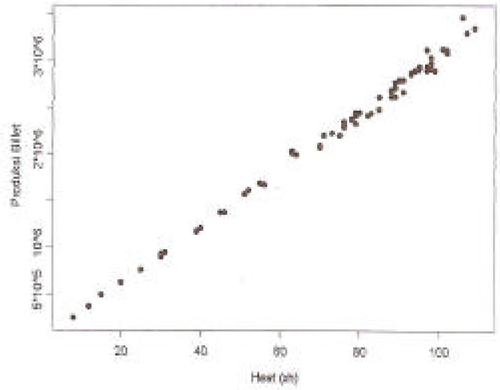
| Titik-titik knots | Nilai GCV | Titik-titik knots | Nilai GCV |
|---|---|---|---|
| $K_1$ = 4367, $K_2$ = 4624 | 2511439320 | **$K_1$ =4369, $K_2$ =4620*** | **2511300803*** |
| $K_1$ = 4368, $K_2$ = 4624 | 2511895253 | $K_1$ =4369, $K_2$ =4621 | 2511563382 |
| $K_1$ = 4369, $K_2$ = 4624 | 2512354376 | $K_1$ =4369, $K_2$ =4622 | 2511826418 |
| $K_1$ = 4370, $K_2$ = 4624 | 2512816671 | $K_1$ =4369, $K_2$ =4623 | 2512090044 |
| $K_1$ = 4371, $K_2$ = 4624 | 2513282288 | $K_1$ =4369, $K_2$ =4624 | 2512354376 |
| $K_1$ = 4372, $K_2$ = 4624 | 2513751160 | $K_1$ =4369, $K_2$ =4625 | 2512619303 |
| $K_1$ = 4373, $K_2$ = 4624 | 2514223441 | $K_1$ =4369, $K_2$ =4626 | 2512884913 |
| $K_1$ = 4374, $K_2$ = 4624 | 2514699006 | $K_1$ =4369, $K_2$ =4627 | 2513151300 |
| $K_1$ = 4375, $K_2$ = 4624 | 2515178021 | $K_1$ =4369, $K_2$ =4628 | 2513418292 |
| $K_1$ = 4376, $K_2$ = 4624 | 2515660580 | $K_1$ =4369, $K_2$ =4629 | 2513686151 |
| $K_1$ = 4377, $K_2$ = 4624 | 2516146642 | $K_1$ =4369, $K_2$ =4630 | 2513954770 |



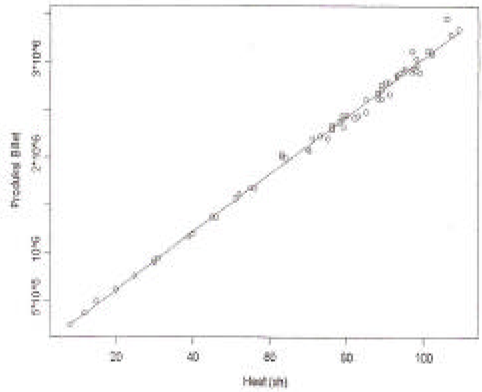

Gambar 1($a$). Plot antara $y$ dengan $x_p$.



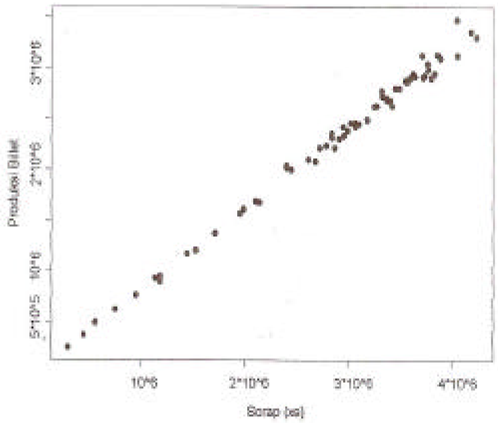

Gambar 1($b$). Model keluarga spline polinomial *truncated* dengan dua knots : 4369 dan 4620.



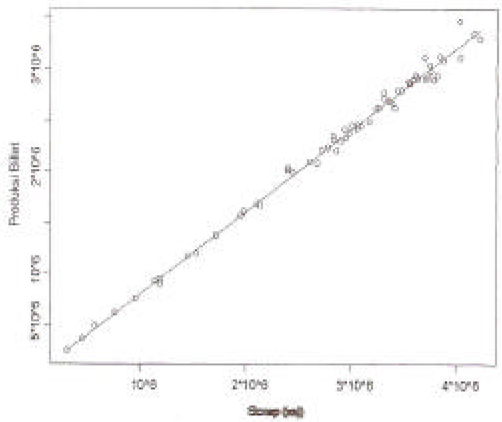

Gambar 2($a$). Plot antara $y$ dengan $x_h$.




Gambar 2($b$). Estimasi model regresi parametrik linear antara $y$ dengan $x_h$.




Gambar 3($a$). Plot antara $y$ dengan $x_s$.




Gambar 3($b$). Estimasi model regresi parametrik linear antara $y$ dengan $x_s$.

**Tabel 2. Estimasi model semiparametrik keluarga spline polinominal *truncated*.**

| Parameter | Estimasi |
|---|---|
| $b$ | $\hat{b} = 30101,46$ |
| $b_2$ | $\hat{b}_2 = 0.013165$ |
| $q_1$ | $\hat{q}_1 = 306,1156$ |
| $q_2$ | $\hat{q}_2 = -0,06948$ |
| $f_1$ | $\hat{f}_1 = 0,8171149$ |
| $f_2$ | $\hat{f}_2 = -1,077396$ |

Tabel 2 menyajikan ringkasan eatimasi parameter model semiparametrik dari keluarga spline polinominal *truncated*. Model estimasi semiparametrik keluarga spline polinominal *truncated* diberikan oleh :

$$\hat{E}(y|x_h, x_s, x_p) = 30101{,}46x_h + 0{,}013165x_s + 306{,}1156x_p - 0{,}06948x_p^2 + $$
$$+0{,}8171149(x_p - 4369)^2 I(x_p \geq 4369) - 1.077396(x_p - 4620)^2 I(x_p \geq 4620)$$

Estimasi pola hubungan produksi Bullet secara parsial dengan variabel prediktor $x_p$ diberikan dalam Gambar 1*(b)*, berupa model keluarga spline polinominal *truncated* kuadratik dengan dua titik knots yaitu $K_1$ = 4369 dan $K_2$ = 4620. Estimasi pola hubungan produksi Billet secara parsial dengan variabel prediktor $x_h$ diberikan dalam Gambar 2*(b)*, berpola parametrik linear. Sedangkan, Estimasi pola hubungan produksi Billet secara parsial dengan variabel prediktor $x_s$ diberikan dalam Gambar 3*(b)*, yang juga berpola parametrik linear.

## Daftar Pustaka


Antoniadis, A., Gregorire, G. and Mackeagu, W., 1994, "Wavelets Methods for Curve Estimation", **Journal of the American Statistical Association**, 89, 1340-1353.

Budiantara, I.N, Subanar, and Soejoeti, Z., 1997, "Weighted Spline Estimator", **Bulletin of the International Statistical Insitute,** 51, 333-334.

Budiantara, I.N, 1999, "Estimator Spline Terbobot Dalam Regresi Semiparametrik", **Majalah Ilmu Pengetahuan dan Teknologi (IPTEKS),** 10, 103-109.

Budiantara, I.N, 2000a, "Metode U, GML, CV dan GCV Dalam Regresi Nonparametrik Spline", **Majalah Ilmiah Himpunan Matematika Indonesia (MIHMI),** 6, 285-290

Budiantara, I.N, 2000b, "Optimasi dan Proyeksi Dalam Regresi Nonparametrik Spline", **Majalah Berkala Matematika dan Ilmu Pengetahuan Alam** (BIMIPA)-Universitas Gadjah Mada, 10, 35-44.

Budiantara, I.N, 2001a, "Regresi Nonparametrik dan Semiparametrik Serta Perkembangannya", **Makalah Pembicara Utama pada Seminar Nasional Alumni Pasca Sarjana Matematika Universitas Gadjah Mada, Yogyakarta.**

Budiantara, I.N, 2001b. "Estimasi Parametrik dan Nonparametrik Untuk pendekatan Kurva Regresi", Makalah Pembicara Utama pada Seminar Nasional Statistika V, **Jurusan Statistika,** Fakultas Matematika dan Ilmu Pengetahuan Alam, Institut Teknologi Sepuluh Nopember (ITS), Surabaya.

Budiantara, I.N, 2004,"Spline : Historis, Motivasi, dan Perannya Dalam Regresi Nonparametrik", **Makalah Pembicara Utama pada Konferensi Nasional Matematika XII,** Jurusan Matematika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Udayana (UNUD), Denpasar.

Craven, P. and Wahba, G., 1979, "Smoothing Noise Data with Spline Functions", **Numerische Mathematics**, 31, 377-403.

Hardle, W., 1990, "*Applied Non-parametrik Regression*", Cambridge University Press, New York.

He, X. and Shi, 1996, "Brivariated tensor Product B-Spline in a Partly Linear Model Linear", **Journal of Multivariate Analysis,** 58, 162-181.

Kreyszig, E., 1978, ***"Introductory Functional Analysis with Application",*** John Wiley and Sons, New York.

Subanar and Budiantara, I.N., 1999, “Weighted Spline Estimator in a Partially Linear Models”, **Procedings of the SEAMS-GMU International Conference 1999 on Mathematics and Its Applications** , 61-70.

Wahba G., 1990, ***“Spline Models for Observasion Data”,*** SIAM Pensylvania.