Jurnal Matematika Vol. 2 No. 1, Juni 2012. ISSN : 1693-1394

ESTIMATOR KERNEL DALAM MODEL REGRESI NONPARAMETRIK

I Komang Gede Sukarsa
e-mail: sukarsakomang@yahoo.com

I Gusti Ayu Made Srinadi
e-mail: srinadiigustiayumade@yahoo.co.id

Jurusan Matematika FMIPA, Universitas Udayana
Kampus Bukit Jimbaran, Badung, Bali

Abstract: Analisis regresi nonparametrik merupakan metode pendugaan kurva regresi yang digunakan jika tidak ada informasi sebelumnya te,ntang benttrk kurva regresi atau tidak terikat pada asumsi bentuk fungsi tertentu. Estimasi fungsi regresi nonparametrik dilakukan berdasarkan daA pengamatan dengan menggunakan teknik pemulusan (*smoothing*). Penelitian ini bertujuan untuk memperlihatkan pendekatan estimator kernel dalam regresi nonparametik padadata sekunder, yaitu data motorcycle. Hasil penelitian ini menunjukkan batrwa penggunaan fungsi kernel yang berbda yaitu fungsi kernel Triangle dan kernel Gaussian dengan bandwidth optimal menghasilkan estimasi kurva regresi yang hanrpir saura, sehingga dapat dituojukkan bahwa pemilihan bandwidth lebih penting dibandingkan dengan pernilihan fungsi kernel.
Keywords: Regresi Nonparametrik Estimator Kernel, *Bandwidth.*

1. Pendahuluan

Analisis regresi merupakan metode analisis data yang menggambarkan hubungan antara variabel respon dengan satu atau beberapa variabel prediktor [5]. Misalkan X adalah variabel prediktor dan Y adalah vaiabel respon untuk n pengamatan berpasangan $\{(x_i, y_i)\}_{i=1}^n$, maka hubungan linear antara variabel prediktor dan variabel respon tersebut dapat dinyatakan sebagai berikut:

$$y_i = m(x_i) + \varepsilon_i, \ i = 1, 2, \cdots, n \tag{1}$$

dengan ε_i adalah sisaan yang diasumsikan independen dingan mean nol dan variansi σ^2, serta $m(x_i)$ adalah fungsi regresi atau kurva regresi [2].

Pendekatan yang digunakan untuk mengestimasi fungsi regresi ada dua jenis, yaitu pendekatan parametrik dan nonparametrik. Dalam pendekatan parametrik, bentuk hubungan antara variabel respon dengan variabel prediktor diketahui atau diperkirakan dari bentuk kurva regresi, misalnya diasumsikan membentuk pola linear, kuadratik, eksponensial, dan polinomial. Dalam regresi parametrik yaitu dalam model regresi linear, harus memenuhi asumsi yang ketat yaitu sisaan berdistribusi normal dan memiliki variansi yang konstan. Untuk mengatasi penyimpangan asumsi dalam model regresi linear dapat dilakukan transformasi terhadap data sehingga diperoleh model regresi yang sesuai bagi data yang telah ditransformasi. Transformasi dipilih melalui teknik *trial* dan *error* sehingga penggunaan transformasi yang tepat akan membawa pada metode pendugaan yang relatif mudah, tetapi kesalahan penggunaan transformasi bisa juga membawa pada metode pendugaan dengan model yang lebih rumit [7]. Pendekatan kedua yaitu pendekatan nonparametik. Estimasi fungsi regresi nonparametrik dilakukan berdasarkan data pengamatan dengan menggunakan teknik (*smoothing*) [2]. Terdapat beberapa teknik smoothing dalam model regresi nonparametrik antara lain histogram, estimator kernel, deret orthogonal, estimator spline, k-NN, deret fourier, dan wavelet.

Ada beberapa jenis fungsi kernel, antara lain kemel *uniform*, *Triangle*, Epanechnikov, Gaussian, kuartik dan cosinus [3]. Dalam regresi kemel pemilihan parameter pemulus (*bandwidth*) jauh lebih penting dibandingkan dengan memilih fungsi kernel. Sehingga yang menjadi masalah dalam regresi kernel adalah pemilihan *bandwidth*, bukan pada pemilihan fungsi kernel. Fungsi kernel yang umum digunakan adalah kernel Gaussian dan kernel Epanechnicov [4]. Kernel *Triangle* sering digunakan karena lebih mudah dan cepat dalam perhitungan [6].

Tujuan yang ingin dicapai dalam penelitian ini adalah untuk mengestimasi model regresi nonparametrik menggunakan estimator kernel *Triangle* dan kernel Gaussian.

2. Tinjauan Teori

2.1. Regresi Nonparametrik

Dalun regresi nonparametrik bentuk kurva regresi tidak diketahui, data diharapkan mencari sendiri bentuk estimasinya sehingga memiliki fleksibelitas yang tinggi. Kurva regresi hanya diasumsikan termuat dalam suatu ruang fungsi yang berdimensi tak hingga dan merupakan fungsi mulus (*smooth*). Estimasi frrngsi $z(x_i)$ dilakukan berdasarkan data pengamatan dengan menggunakan teknik smoothing tertentu. Ada beberapa teknik *smoothing* yang dapat digurnakan anttra lain estimator histogram, kernel, deret orthogonal, spline, k-NN, deret fourier, dan wavelet [2].

2.2. Estimator Densitas Kernel

Estimator kernel merupakan pengembangan dari estimator histogram. Estimator kernel diperkenalkan oleh Rosenblatt (1956) dan Parzen (1962) sehingga disebut estimator densitas kernel Rosenblatt-Parzen [3].

Secara umum kernel K dengan *bandwidth* h [8] didefinisikan sebagai:

$$K_h(x) = \frac{1}{h} K\left(\frac{x}{h}\right), \text{ untuk } -\infty < x < \infty,\ h > 0 \tag{2}$$

serta memenuhi:

(i) $K(x) \geq 0$, untuk semua x

(ii) $\int_{-\infty}^{\infty} K(x)dx = 1$

(iii) $\int_{-\infty}^{\infty} x^2 K(x)dx = \sigma^2 > 0$

(iv) $\int_{-\infty}^{\infty} xK(x)dx = 0$

maka estimator densitas kernel untuk fungsi densitas $f(x)$ adalah:

$$\hat{f}_h = \frac{1}{n}\sum_{i=1}^{n} K_h(x - x_i) = \frac{1}{nh}\sum_{i=1}^{n} K\left(\frac{x - x_i}{h}\right) \tag{3}$$

dari persamaan (3) terlihat bahwa $\hat{f}_h(x)$ tergantung pada fungsi kernel K dan parameter h. Bentuk bobot kernel ditentukan oleh fungsi kernel K, sedangkan ukuran bobotnya ditentukan oleh parameter pemulus h yang disebut *bandwidth*. Peran *bandwidth* seperti lebar interval pada histogram.

Beberapa jenis fungsi kernel [3] antara lain:

1. Kernel Uniform : $K(x) = \dfrac{1}{2}I(|x| \leq 1)$
2. Kernel Triangle : $K(x) = (1 - |x|)I(|x| \leq 1)$
3. Kernel Epanechnikov : $K(x) = \frac{3}{4}(1 - x^2)I(|x| \leq 1)$
4. Kernel Kuartik : $K(x) = \frac{15}{16}(1 - x^2)^2 I(|x| \leq 1)$
5. Kernel Triweight : $K(x) = \frac{35}{32}(1 - x^2)^3 I(|x| \leq 1)$
6. Kernel Cosinus : $K(x) = \frac{\pi}{4}\cos\left(\frac{\pi}{2}x\right) I(|x| \leq 1)$
7. Kernel Gaussian : $K(x) = \frac{1}{\sqrt{2\pi}}\left(\frac{1}{2}(-x^2)\right) \quad -\infty < x < \infty$

dengan I adalah indikator

2.3. Regresi Kernel

Regresi kernel adalah teknik statistika nonparametrik untuk mengestimasi fungsi regresi $m(x)$ pada model regresi nonparametrik $y_1 = m(x_i) + \varepsilon_i$. Nadaraya dan Watson pada tahun 1964 mendefinisikan estimator regresi kernel sehingga disebut estimator Nadaraya-Watson [3]

$$\hat{m}(x) = \frac{\frac{1}{n}\sum_{i=1}^{n} K_h(x - x_i)y_i}{\frac{1}{n}\sum_{i=1}^{n} K_h(x - x_i)} \tag{4}$$

$$\hat{m}(x) = \sum_{i=1}^{n} w_{hi}(x)y_i \tag{5}$$

dengan

$$w_{hi}(x) = \frac{\frac{1}{h}K\left(\frac{x-x_i}{h}\right)}{\frac{1}{h}\sum_{i=1}^{n} K\left(\frac{x - x_i}{h}\right)} = \frac{K\left(\frac{x-x_i}{h}\right)}{\sum_{i=1}^{n}\left(\frac{x - x_i}{h}\right)}$$

2.4. Pemilihan Bandwidth Optimal

Bandwidth (h) adalah parameter pemulus (*smoothing*) yang berfungsi untuk mengontrol kemulusan dari kurva yang diestimasi. Bandwidth yang terlalu kecil akan menghasilkan kurva yang *under-smoothing* yaitu sangat kasar dan sangat fluktuatif, dan sebaliknya bandwidth yang terlalu lebar akan menghasilkan kurva yang *over-smoothing* yaitu sangat mulus, tetapi tidak sesuai dengan pola data (Hardle, 1994). Oleh karena itu perlu dipilih *bandwidth* yang optimal. Salah satu metode untuk mendapatkan h optimal adalah dengan menggunakan kriteria *Generalized Cross Validation* (GCV), yang didefinisikan sebagai berikut:

$$GCV = \frac{MSE}{\left(\frac{1}{n}tr(I - H(h))\right)^2} \tag{6}$$

dengan $MSE = \dfrac{1}{n}\sum_{i=1}^{n}(y - m_h(x_i))^2$. kebaikan suatu estimator dapat dilihat dari tingkat kesalahannya. Semakin kecil tingkat kesalahannya semakin baik estimasinya. Menurut [1], kriteria untuk menentukan estimator terbaik dalam model regresi nonparametrik, antara lain:

1. *Mean Square Error*(MSE)
$$\text{MSE} = \frac{1}{n}\sum_{i=1}^{n} e_i^2 = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2$$
2. *Root Mean Square Error*(RMSE)
$$\text{RMSE} = \sqrt{\text{MSE}}$$
3. *Mean Absolute Deviation*(MAD)
$$\text{MAD} = \frac{\sum_{t=1}^{n} |e_t|}{n} = \frac{\sum_{t=1} |y_i - \hat{y}_i|}{n}$$

3. Metode Penelitian

3.1. Sumber Data

Data yang digunakan dalam penelitian ini adalah data sekunder yang diambil dari buku *Applied Nonparametric Regression*. Data ini merupakan hasil penelitian yang dilakukan oleh Schmidt, Mattern, dan Schuler pada tahun 1981 yaitu data simulasi tabrakan sepeda motor pada suatu *Post Mortem Human Test Object* (PTMO) [3].

3.2. Identifikasi Variabel

Identifikasi variabel dalam penelitian ini adalah variabel prediktor (X) yaitu waktu (dalam milidetik) setelah simulasi tabrakan dan variabel respon (Y) yaitu percepatan (dalam $g, 1g = 9,81\text{m/s}^2$) setelah tabrakan yang disimulasikan.

3.3. Metode Analisis Data

Dalam penelitian ini model regresi nonparametrik diestimasi menggunakan estimator kernel, dengan fungsi kernel *Triangle* dan kernel Gaussian, dengan *macro* program menggunakan *software* S-plus. Adapun langkah-langkah yang dilakukan adalatt sebagai berikut:

1. Mengestimasi kurva regresi nonparametrik dengan estimator kernel *Triangle*.
2. Mengestimasi kurva regresi nonparametrik dengan estimator kernel Gaussian.
3. Membandingkan hasil estimasi antara estimator kernel *Triangle* dan kernel Gaussian menggunakan bandwidth yang optimal pada data sekunder, dengan membandingkan plot estimasi kurva regresi bersama-sama dengan plot data serta melihat nilai MSE, RMSE, dan MAD.

4. Hasil dan Pembahasan

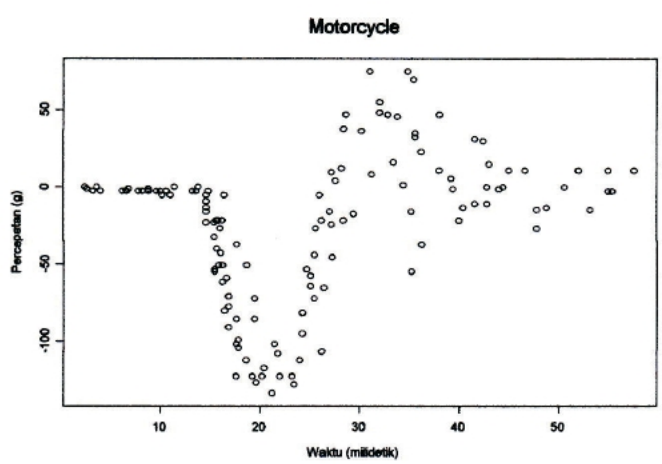
Gambaran umum data yang diolah menggunakan software S-Plus secara rinci dapat dilihat pada Tabel l.

Tabel 1. Statistika Deskriptif dan Motorcycle

Variabel	N	Min	Maks	Mean	Q_2	Standar deviasi
Waktu (x)	133	2,4	57,6	25,18	23,4	13,132
Percepatan	133	$-134,0$	25,0	$-25,55$	$-13,3$	48,322

Banyak data pengamatan adalah 133, dengan waktu minimum sebesar 2,4 milidetik waktu maksimum sebesar 57,6 milidetik dan percepatan minimum sebesar $-134,0$ g, percepatan maksimum 25,0 g. Rata-rata waktu sebesar 25,18 milidetik, dan percepatan sebesar $-25,55$ g, dengan nilai tengah (median) waktu sebesar 23,4 milidetik dan percepatan sebesar $-13,3$ g, serta standar deviasi waktu sebesar 13,132 milidetik dan percepatan 48,332 g.

Bentuk hubungan antara variabel prediktor (waktu) dengan variabel respon (percepatan) dilihat dari plot antara kedua variabel tersebut (Gambar 1).



Gambar 1. Diagram Pencar Data *Motorcycle*

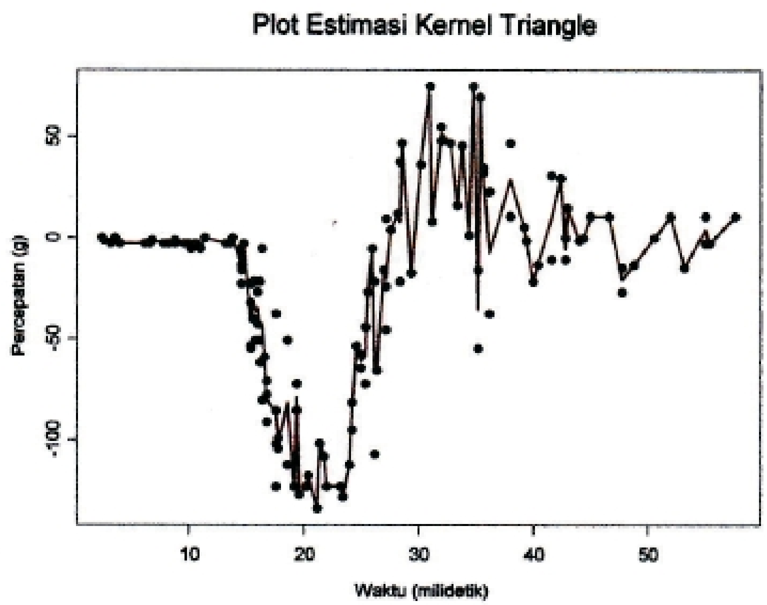
Gambar 1 menunjukkan bentuk kurva yang menggambarkan hubungan antara waktu (milidetik) dengan percepatan (g), yang sangat sulit diestimasi bila digunakan pendekatan regresi parametrik, karena kurva tidak membentuk pola linear, kuadratik, eksponensial, atau kubik. Kurva regresi akan diestimasi menggunakan pendekatan regresi nonparametrik dengan estimator kernel.

4.1. Estimasi Data *Motorcycle* dengan Estimator Kernel

Langkah-langkah yang dilakukan dalam mengestimasi kurva regresi nonparametrik menggunakan estimator kernel adalah menentukan fungsi bobot atau fungsi kernel dan ukuran bobot yaitu nilai bandwidth (h) yang optimal. Sebelum menentukan nilai bandwidth (A), terlebih dahulu dipilih fungsi kernel yang akan digunakan. Dalam penelitian ini digunakan fungsi kernel *Triangle* dan kernel Gaussian.

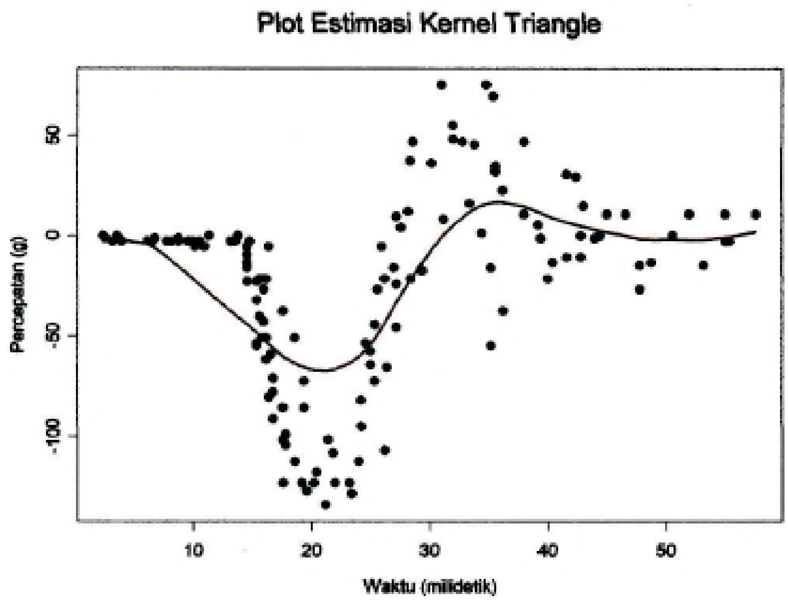
4.2. Estimasi Data *Motorcycle* dengan Estimator Kernel *Triangle*

Pemilihan bandwidth (h) merupakan langkah terpenting dalam kernel *smoothing*, apabila nilai h yang dipilih terlalu kecil akan diperoleh kurva regresi yang sangat kasar (*under-smoothing*), sebaliknya apabila nilai h terlalu besar akan menghasilkan kurva yang sangat mulus (*over-smoothing*).



Gambar 2. Plot Estimasi Kernel *Triangle* dengan *Bandwidth* = 0,1

Nilai bandwidth yang terlalu kecil, misalkan $h = 0,1$ menghasilkan kurva regresi yang sangat kasar, seperti terlihat pada Gambar 2, sedangkan nilai bandwidth yang terlalu besar, misalkan $h = 10$ menghasilkan kurva regresi yang sangat mulus dan tidak sesuai dengan pola data, seperti terlihat pada Gambar 3.



Gambar 3. Plot Estimasi Kernel *Triangle* dengan *Bandwidth* = 10

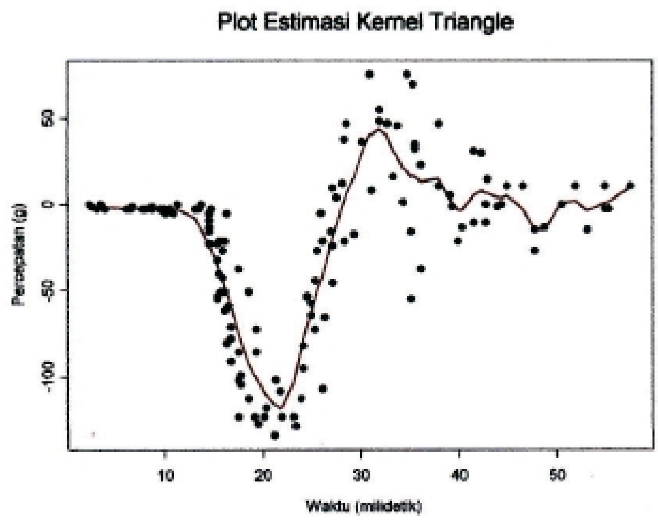
Untuk memperoleh estimasi kurva regresi yang optimal, yaitu kurva yang mulus dan sesuai dengan pola data, perlu dipilih nilai bandwidth (h) yang optimal. Nilai bandwidth (h) berdasarkan kriteria GCV minimum dengan *macro* program *software* S-Plus pada selang kenaikan nilai h yang cukup kecil, misallran diambil kenaikan nilai h sebesar 0,005 sehingga diperoleh nilai bandwidth (h) dan GCV yang ditunjukkan pada Tabel 2.

Tabel 2. Nilai Bandwidth dan GCV dengan Kernel *Triangle*

Bandwidth	GCV	*Bandwidth*	GCV
2,295	655,6136	2,375	655,4753
2,300	655,5928	2,380	655,4785
2,305	655,5738	2,385	655,4830
2,310	655,5565	2,390	655,4886
2,315	655,5411	2,395	655,4953
2,320	655,5273	2,400	655,5031
2,325	655,5152	2,405	655,5222
2,330	655,5047	2,410	655,5428
2,335	655,4957	2,415	655,5649
2,340	655,4883	2,420	655,5883
2,345	655,4823	2,425	655,6131
2,350	655,4778	2,430	655,6392
2,355	655,4746	2,435	655,6666
2,360	655,4729	2,440	655,6952
2,365	655,4724	2,445	655,7250
2,370	655,4732	2,450	655,7560

Tabel 2 memperlihatkan GCV minimum bernilai 655,4724 yaitu pada nilat *bandwidth* (h) sebesar 2,365, maka nilai *bandwidth* (h) optimal untuk fungsi kernel *Triangle* adalah 2,365.

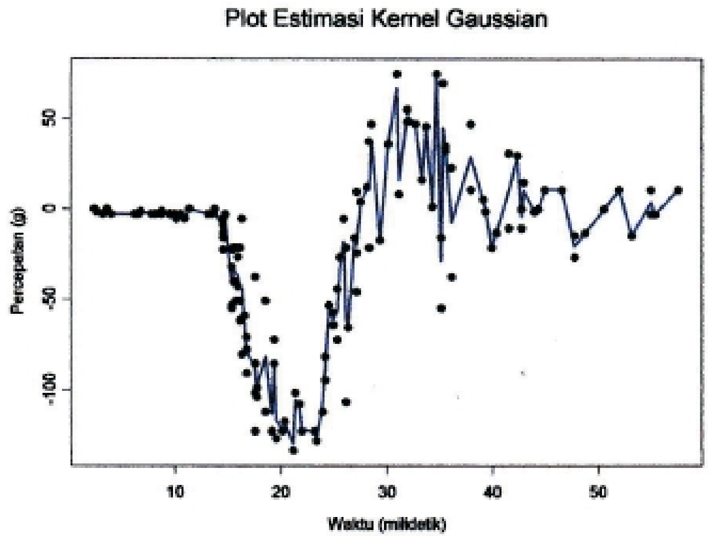
Setelah diperoleh nilai bandwidth yang optimal berdasarkan kriteria GCV minimum, kemudian dilakukan estimasi model regresi nonparametrik dengan estimator kernel *Triangle* pada bandwidth yang optimal, yaitu menghitung nilai $\hat{m}(x)$ dengan macro progam *software* S-plus, sehingga diperoleh nilai dugaan $\hat{m}(x)$ untuk kernel *Triangle* dan estimasi kurva regresi yang ditunjukkan pada Gambar 4.



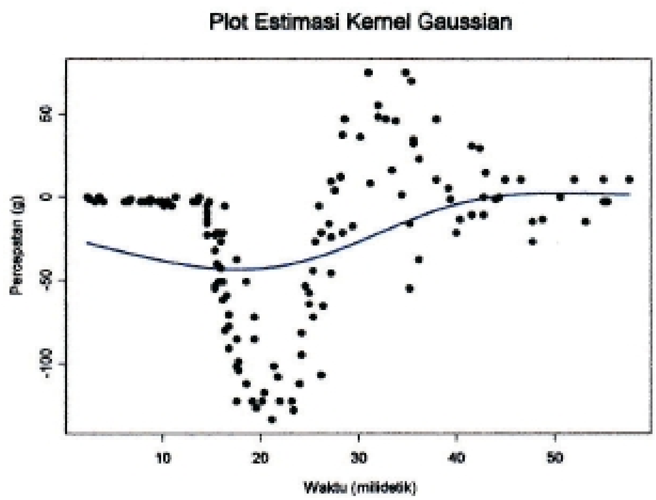
Gambar 4. Plot Estimasi Kernel *Triangle* dengan *Bandwidth* Optimal = 2,365

4.3. Estimasi Data *Motorcycle* dengan Estimator Kernel Geussian

Nilat *bandwidth* yang terlalu kecil, misalkan $h = 0,1$ menghasilkan kurva regresi yang sangat kasar, seperti terlihat pada Gambar 5. Sebaliknya nilai *bandwidth* yang terlalu besar, misalkan $h = 10$ menghasilkan kurva regresi yang sangat mulus dan tidak sesuai dengan pola data seperti yang ditunjukkan pada Gambar 6.



Gambar 5. Plot Estimasi Kernel Gaussian dengan *Bandwidth* = 0,1



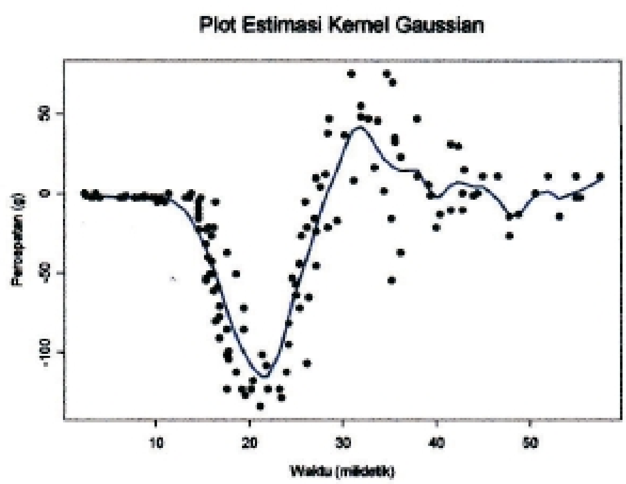
Gambar 6. Plot Estimasi Kernel Gaussian dengan *Bandwidth* = 10

Nilai *bandwidth* (h) berdasarkan kriteria GCV minimum dengan macro program *software* SPlus, untuk memperoleh nilai *bandwidth* (h) yang lebih akurat, selang kenaikan nilai h dibuat kecil, misalkan sebesar 0,005, sehingga diperoleh nrlar *bandwidth* (h) dan GCV seperti pada Tabel 3.

Tabel 3. Nilai *Bandwidth* dan GCV dengan Kernel Gaussian

Bandwidth	GCV	*Bandwidth*	GCV
1,040	650,1474	1,125	649,9816
1,045	650,0821	1,130	650,0300
1,050	650,0241	1,135	650,0843
1,055	649,9736	1,140	650,1446
1,060	649,9302	1,145	650,2107
1,065	649,8940	1,150	650,2826
1,070	649,8648	1,155	650,3603
1,075	649,8425	1,160	650,4436
1,080	649,8271	1,165	650,5326
1,085	649,8184	1,170	650,6271
1,090	649,8163	1,175	650,7271
1,095	649,8208	1,180	650,8326
1,100	649,8319	1,185	650,9435
1,105	649,8493	1,190	651,0598
1,110	649,8730	1,192	651,1814
1,115	649,9031	1,200	651,3083
1,120	649,9393	1,205	651,4404

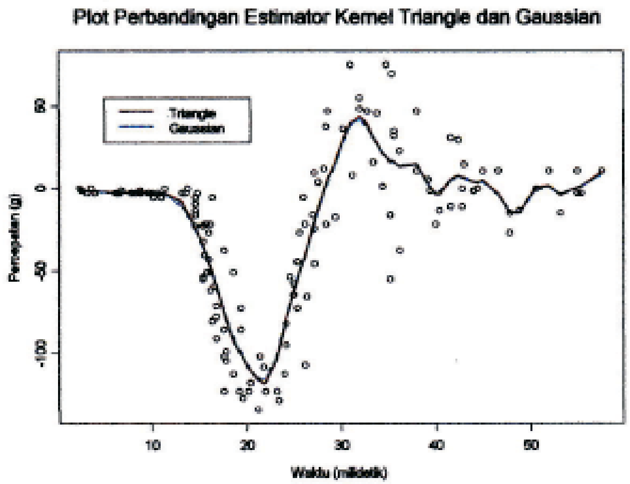
Tabel 3. menunjukkan GCV minimum bernilai 649.8163 pada nilai *bandwidth* (h) sebesar 1,090. Sehingga nilai *bandwidth* (h) optimal untuk fungsi kernel Gaussian adalah 1,090. Setelah diperoleh nilai *bandwidth* yang optimal berdasarkan kriteria GCV, kemudian dilakukan estimasi model regresi nonparametrik dengan estimator kernel Gaussian menggunakan *bandwidth* yang optimal, yaitu menghitung nilai $\hat{m}(x)$ dengan macro progam *software* S-plus, sehingga diperoleh nilai dugaan $\hat{m}(x)$ tmtuk kernel Gaussian dan estimasi kurva regresi yang ditunjukkan padra Gambar 7.



Gambar 7. Plot Estimasi Kernel Gaussian dengan *Bandwidth* = 1,090

4.4. Perbandingan Estimator Kernel *Triangle* dan Kernel Gaussian

Pendekatan estimator kernel *Triangle* dan Gaussian dalam mengestimasi kurva regresi terlihat dalam Gambar 8 berikut.



Gambar 8. Plot Perbandingan Estimasi Kernel *Triangle* dan Gaussian

Hasil estimasi kurva regresi antara fungsi kernel *Triangle* dan kernel Gaussian sangat berimpit dimana menghasilkan bentuk kurva regresi yang sangat mirip. Selanjutnya dilihat dari perbandingan nilai MSE, RMSE, dan MAD yang dihasilkan kedua fungsi kernel tersebut yang tercantum pada Tabel 4.

Tabel 4. Perbandingan Estimator Kernel *Triangle* dan Gaussian

Fungsi Kernel	***Bandwidth* (h) Optimal**	**MSE**	**RMSE**	**MAD**
Triangle	2,365	452,1965	21,26491	15,75821
Gaussian	1,090	469,5878	21,66997	16,20430

Tabel 4. menunjukkan nilai MSE, RMSE, dan MAD yang dihasilkan fungsi kernel *Triangle* dan kernel Gaussian dengan menggunakan bandwidth optimal. Secara statistik nilai MSE, RMSE, dan MAD yang dihasilkan kernel *Triangle* hampir mendekati nilai-nilai pada kernel Gaussian, sehingga dapat dikatakan nilai MSE, RMSE dan MAD yang dihasilkan kedua fungsi kernel tersebut hampir sama.

Berdasarkan plot hasil estimasi untuk fungsi kernel *Triangle* dan kernel Gaussian dengan menggunakan bandwidth optimal, sangat berimpit, serta perbandingan nilai MSE, RMSE, dan MAD yang menunjukkan hasil yang hampir sama sehingga dapat dikatakan bahwa penggunaan fungsi kernel yang berbeda dengan bandwidth yang optimal untuk masing-masing fungsi kernel tersebut akan menghasilkan estimasi kurva regresi yang sama. Hasil penelitian ini mendukung pendapat yang dikemukakan oleh Hastie dan Tibshirani [4], yang menyatakan bahwa dalam regresi kernel pemilihan parameter pemulus (*bandwidth*)jauh lebih penting dibandingkan dengan memilih fungsi kernel.

5. Kesimpulan

Berdasarkan hasil dan pembahasan dapat diambil simpulan bahwa untuk data *motorcycle* diperoleh *bandwidth* optimal untuk estimator kernel *Triangle* sebesar 2,365 dan kemel Gaussian sebesar 1,090. Dalam regresi kernel yang terpenting adalah pemilihan nilai bandwidth optimal, bukan pemilihan fungsi kernel, karena penggunaan fungsi kernel yang berbeda dengan nilai bandwidth optimal menghasilkan estimasi kurva regresi yang hampir sama Hal ini sesuai dengan pendapat yang dikemukakan oleh Hastie dan Tibshirani [4], yaitu dalam regresi kernel pemilihan parameter pemulus (*bandwidth*) jauh lebih penting dibandingkan dengan memilih fungsi kernel.

Daftar Pustaka

[1] Aydin, Dursun. 2007. *A Comparison of the Nonparametric Regression Models using Smoothing Spline and Kernel Regression.* World Academy of Science, Engineering and Technology, 36, 253-257, Turkey. http://www.waset.org/journals/waset/v36/v36-46.pdf. Diakses tanggal 9 Februari 2010.

[2] Eubank, R. 1998. *Spline Smoothing and Nonparametric Regression.* Marcel Dekker. New York.

[3] Hardle, W. 1994. *Applied Nonparametric Regression.* Cambridge University Press. New York.

[4] Hastie, T.J. and R.J. Tibshirani. 1990. *Generalized Additive Models.* Chapman and Hall. New York. London

[5] Hosmer, D.W. and S.Lemeshow. 2000. *Applied Logistic Regression*, 2^{nd}.John Wiley and Sons, Inc.New York.

[6] MathSoft. 1993. *S-Plus Guide to Statistical and Mathematical Analysis.* Version 3.2. A Division of Math Soft, lnc. Seattle, Washington.

[7] Neter,J., W. Wasserrman dan M. H. Kutner. 1997.*Model Linier Terapan Analisis Regresi Linier Sederhana.* Diterjemahkan oleh Bambang Sumantri. Jurusan Statistika FMIPA IPB. Bogor.

[8] Wand M.P. and M.C.Jones. 1995. *Kernel Smoothing.* Chapman and Hall. New York.