

## ESTIMATOR KURVA *PRIESTLEY CHAO* MENGGUNAKAN FUNGSI KERNEL *TRIANGLE* UNTUK DATA RATA-RATA BULANAN BILANGAN *SUNSPOT* MATAHARI

Meilani Nurjaina\*, Bahridin Abapihi, Irma Yahya, Baharuddin, Ruslan, Makkulau  
Program Studi S1 Statistika FMIPA Universitas Halu Oleo

\*Email: [meilani.nurj15@gmail.com](mailto:meilani.nurj15@gmail.com)

### ABSTRACT

*This study aims to obtain a Priestley Chao curve estimator by using the Triangle kernel function in a nonparametric regression model for the monthly average solar number data from 1954 to 2008. Based on the results of the analysis and discussion, it can be concluded that the Priestley Chao curve estimator using the triangle kernel function produces a curve regression that almost follows the true curve pattern of the original data. The optimum bandwidth value obtained is 0.9 with a minimum GCV value of 1.0321 and a very small MSE value.*

**Keywords:** *Bandwidth, Sunspot Number, Priestley Chao Kernel Estimator, Kernel Triangle Function, Nonparametric Regression*

### ABSTRAK

Penelitian ini bertujuan untuk mendapatkan estimator kurva Priestley Chao menggunakan fungsi kernel Triangle pada model regresi nonparametrik untuk data rata-rata bulanan bilangan sunspot matahari dari tahun 1954 sampai 2008. Berdasarkan hasil analisis dan pembahasan dapat diketahui bahwa estimator kurva Priestley Chao dengan menggunakan fungsi kernel triangle menghasilkan kurva regresi yang hampir mengikuti pola kurva data asli yang sebenarnya. Nilai bandwidth optimum yang diperoleh adalah 0,9 dengan nilai GCV minimum sebesar 1,0321 serta nilai MSE yang sangat kecil.

**Kata Kunci:** *Bandwidth, Bilangan Sunspot, Estimator Kernel Priestley Chao, Fungsi Kernel Triangle, Regresi Nonparametrik*

### PENDAHULUAN

Terdapat tiga bentuk pendekatan regresi yaitu pendekatan parametrik, pendekatan semiparametrik, dan pendekatan nonparametrik. Pendekatan parametrik digunakan untuk mengestimasi model jika pola data diketahui. Pendekatan semiparametrik digunakan jika sebagian distribusi data dari sekumpulan variabel respon dan variabel prediktor diketahui dan untuk sebagian yang lain tidak diketahui. Sedangkan untuk mengestimasi model jika distribusi data tidak diketahui digunakan pendekatan nonparametrik.

Estimasi fungsi regresi nonparametrik dilakukan berdasarkan data pengamatan dengan menggunakan teknik (*smoothing*). Terdapat beberapa teknik *smoothing* dalam model regresi nonparametrik antara lain *histogram*, *estimator kernel*, *deret orthogonal*, *estimator spline*, *k-NN*, *deret fourier*, dan *wavelet*. Salah satu teknik yang akan digunakan untuk mengestimasi regresi nonparametrik dalam penelitian ini adalah pendekatan kernel. Pendekatan kernel memiliki bentuk yang lebih fleksibel. Pada regresi kernel terdapat tiga estimator, yaitu *Nadaraya Watson*, *Gasser Muller*, *Priestley Chao*.

Pendekatan kernel memiliki beberapa fungsi, diantaranya *uniform*, *triangle*, *epanechnikov*, *quartik*, *triweight*, *gaussian*, dan *cosinus*. Fungsi kernel yang akan digunakan dalam penelitian ini adalah fungsi kernel *triangle*. Hasil penelitian Nisa' (2016) telah membandingkan fungsi kernel *epanechnikov* dan kernel *triangle* menggunakan estimator *Nadaraya Watson* pada data rata-rata bulanan bilangan *sunspot* dimana fungsi kernel *triangle* memiliki nilai *Mean Square Error* (MSE) yang lebih kecil dibandingkan fungsi kernel lainnya. Hal yang terpenting dalam pendekatan kernel adalah pemilihan *bandwidth* atau parameter penghalus. *Bandwidth* berfungsi untuk mengatur kehalusan kurva yang akan diestimasi. Salah satu cara untuk memilih

*bandwidth* yang sesuai adalah dengan menggunakan kriteria *General Cross Validation* (GCV). Tujuan penelitian ini adalah untuk mendapatkan estimator kurva *Priestley Chao* menggunakan fungsi kernel *Triangle* pada model regresi nonparametrik untuk data rata-rata bulanan bilangan *sunspot* matahari dari tahun 1954 sampai 2008.

## METODE PENELITIAN

Data yang digunakan dalam penelitian ini adalah data sekunder hasil pengamatan *sunspot* dari *Solar Indices Data NOAA* (*National Oceanic and Atmospheric Administration*) yang dipublikasikan oleh NOAA, divisi *Space Weather & Solar Event* dan diambil dari alamat <http://www.ngdc.noaa.gov> atau [sidc.be/silso/datafiles](http://sidc.be/silso/datafiles). Data yang digunakan adalah hasil pengamatan dari tahun 1954 sampai tahun 2008 data rata-rata bulanan bilangan *sunspot*.

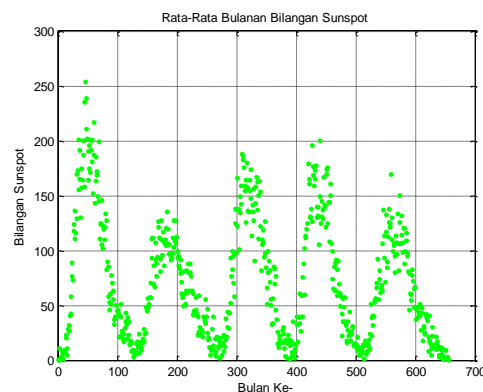
Variabel dalam penelitian ini terdiri dari variabel respon ( $Y$ ) dan variabel prediktor ( $X$ ). Variabel respon ( $Y$ ) yaitu data rata-rata bulanan bilangan *sunspot* dan variabel prediktor ( $X$ ) menunjukkan waktu pengamatan jumlah aktivitas matahari.

Adapun langkah-langkah yang akan dilakukan dalam penelitian ini adalah sebagai berikut :

1. Analisis statistika deskriptif.
2. Menentukan estimator *Priestley Chao* dengan menggunakan fungsi kernel *triangle*.
3. Menentukan *bandwidth* yang digunakan dalam estimator.
4. Memasukkan *bandwidth* ke dalam estimator kernel *triangle*.
5. Menghitung nilai GCV dari *bandwidth* yang telah digunakan dalam estimator kernel *triangle*.
6. Pemilihan *bandwidth* optimum berdasarkan nilai GCV minimum.
7. Memperoleh hasil estimasi terbaik untuk data rata-rata bulanan bilangan *sunspot* berdasarkan *bandwidth* optimum.
8. Kesimpulan.

## HASIL DAN PEMBAHASAN

Data yang digunakan dalam penelitian ini merupakan data rata-rata bulanan bilangan *sunspot* dari hasil penelitian sebelumnya yaitu fungsi kernel *triangle* menggunakan estimator Nadaraya Watson, pengamatan bulan Januari tahun 1954 hingga bulan Juli tahun 2008. Data rata-rata bulanan bilangan *sunspot* tersebut kemudian diplot terhadap waktu pengamatan yang dimulai dari bulan ke-1 sampai bulan ke-655.



**Gambar 1** Plot Data Rata-rata Bulanan Bilangan *Sunspot* Tahun 1954 sampai 2008

Gambar (1) merupakan plot data rata-rata bulanan bilangan *sunspot* yang menggambarkan hubungan antara waktu pengamatan (bulan ke-1 sampai 655) sebagai variabel bebas ( $X$ ) dan rata-rata bulanan bilangan *sunspot* sebagai variabel terikat ( $Y$ ). Plot data rata-rata bulanan bilangan *sunspot* pada Gambar (1) menunjukkan 5 siklus yaitu siklus ke-19 sampai siklus ke-23.

**Analisis Data**

Fungsi kernel *triangle* disubstitusikan sehingga memperoleh estimator *Priestley Chao* sebagai berikut :

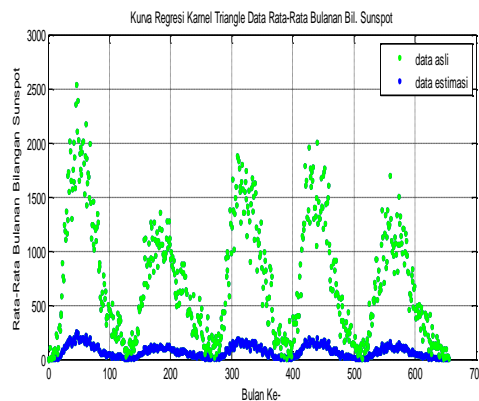
$$\hat{m}(x_i) = \frac{1}{h} \sum_{i=1}^n (x_i - x_{i-1}) \left( \left( 1 - \left| \frac{x - x_i}{h} \right| \right) I \right) Y_i \tag{4.1}$$

sehingga estimator regresi *Priestley Chao* :

$$\hat{Y}_i = \frac{1}{h} \sum_{i=1}^n (x_i - x_{i-1}) \left( \left( 1 - \left| \frac{x - x_i}{h} \right| \right) I \right) Y_i \tag{4.2}$$

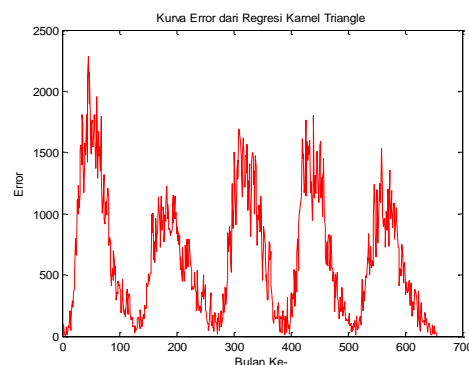
Hasil substitusi akan digunakan untuk memperoleh hasil estimasi data rata-rata bulanan bilangan *sunspot* yang telah diplotkan pada Gambar (1). Hal yang paling penting dalam memperoleh hasil estimasi kurva regresi pada pendekatan kernel adalah memilih *bandwidth* yang sesuai. Pemilihan *bandwidth* yang terlalu besar mengakibatkan kurva hasil estimasi menjadi sangat halus. Sebaliknya pemilihan *bandwidth* yang terlalu kecil mengakibatkan kurva hasil estimasi menjadi kasar.

Kurva regresi rata-rata bulanan bilangan *sunspot* matahari terhadap waktu (bulan ke-) dari model diatas dapat dilihat pada Gambar (2).



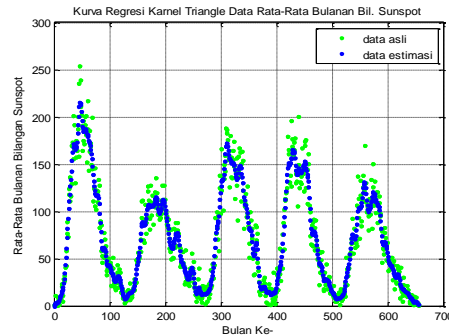
**Gambar 2** Kurva Regresi Kernel *Triangle* Rata-rata Bulanan Bilangan *Sunspot* terhadap waktu (bulan ke-) dengan nilai *bandwidth* 0,1

Pada Gambar (2) terlihat bahwa kurva yang dihasilkan sangat kasar. Kurva regresi (titik biru) tidak mengikuti pola kurva asli (titik hijau), kurva regresi berada jauh dibawah kurva asli. Hal ini berarti nilai *bandwidth* yang dipilih terlalu kecil untuk digunakan pada data rata-rata bulanan bilangan *sunspot*. Kurva error untuk kurva regresi kernel *triangle* dengan nilai *bandwidth* sebesar 0,1 dapat dilihat pada Gambar (3). Selisih nilai hasil estimasi dan data asli yang sangat besar menyebabkan nilai error yang sangat besar pula.

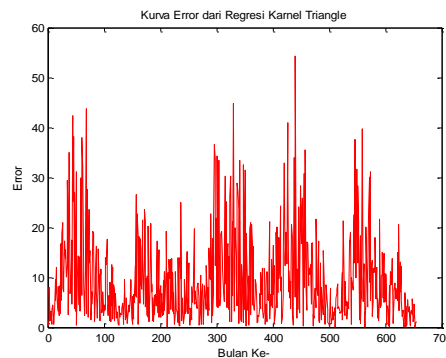


**Gambar 3** Kurva Error Regresi Kernel *Triangle* terhadap Rata-rata Bulanan Bilangan *Sunspot* dengan *bandwidth* 0,1

Pada Gambar (4) yang menampilkan kurva estimasi dengan nilai *bandwidth* 5 dapat dilihat bahwa kurva regresi (titik biru) mendekati kurva asli (titik hijau), namun kurva regresi tersebut terlalu mulus. Kurva regresi tidak melewati sebagian pola data asli sehingga dapat dikatakan bahwa kurva regresi ini belum baik. Kurva error yang ditampilkan pada Gambar (5) menunjukkan bahwa nilai error untuk hasil estimasi data dengan nilai *bandwidth* 5 sangat besar.



**Gambar 4** Kurva Regresi Kernel *Triangle* Rata-rata Bulanan Bilangan *Sunspot* terhadap waktu (bulan ke-) dengan *bandwidth* sebesar 5



**Gambar 5** Kurva Error Regresi Kernel *Triangle* terhadap Rata-rata Bulanan Bilangan *Sunspot* dengan *bandwidth* 5

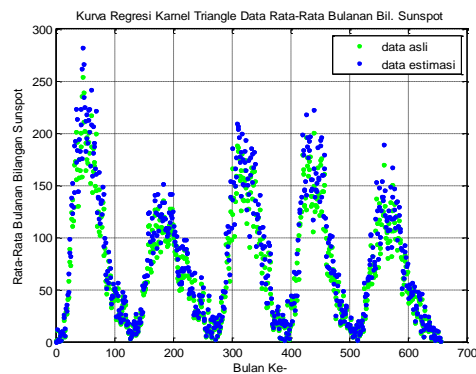
Untuk memilih *bandwidth* optimum dapat dilakukan dengan melihat nilai GCV yang minimum. Nilai GCV dari interval *bandwidth* 0,1 sampai 5 akan ditampilkan pada tabel berikut:

**Tabel 1** Nilai GCV dari interval *bandwidth* 0,1 sampai 5

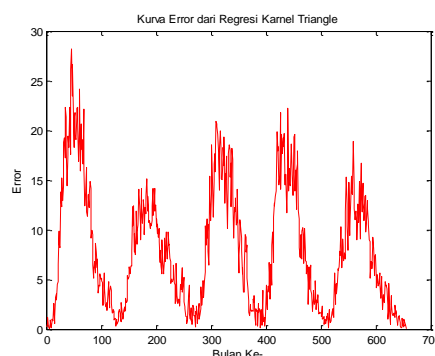
<i>Bandwidth</i>	GCV	<i>Bandwidth</i>	GCV
0,1	6,7718	2,6	106,0195
0,2	1,3376	2,7	107,1356
0,3	4,5517	2,8	108,0401
0,4	1,8810	2,9	109,5802
0,5	8,3603	3	112,3951
0,6	3,7156	3,1	116,4774
0,7	1,5355	3,2	121,0009
0,8	5,5751	3,3	125,1744
<b>0,9</b>	<b>1,0321</b>	3,4	128,6756
1	4,4410	3,5	131,4649
1,1	4,8287	3,6	133,6643
1,2	1,1044	3,7	135,4798
1,3	1,4331	3,8	137,1519

<i>Bandwidth</i>	<i>GCV</i>	<i>Bandwidth</i>	<i>GCV</i>
1,4	1,4735	3,9	138,9254
1,5	1,3309	4	141,0315
1,6	1,1094	4,1	143,3839
1,7	88,6681	4,2	145,8886
1,8	71,3436	4,3	148,2906
1,9	61,9227	4,4	150,4689
2	61,8464	4,5	152,3933
2,1	71,6309	4,6	154,0926
2,2	83,2881	4,7	155,6328
2,3	93,0696	4,8	157,1006
2,4	99,8957	4,9	158,5926
2,5	103,9783	5	160,2079

Berdasarkan Tabel (1) dapat dilihat bahwa nilai Pada interval 0,1 hingga 1,9 nilai GCV memiliki nilai yang tidak stabil, naik dan turun, sedangkan pada interval 2 sampai 10, nilai GCV konsisten mengalami kenaikan. Dari Tabel (1) dapat diketahui pula nilai GCV yang paling minimum adalah *bandwidth* bernilai 0,9 dengan nilai GCV sebesar 1,0321. Kurva regresi hasil estimasi dan kurva *error* dengan *bandwidth* optimum dapat dilihat pada Gambar (6) dan Gambar (7). Sedangkan pada penelitian sebelumnya diperoleh nilai GCV yang paling minimum adalah *bandwidth* bernilai 1 dengan nilai GCV sebesar 0.



**Gambar 6** Kurva regresi kernel *triangle* rata-rata bulanan bilangan *sunspot* terhadap waktu (bulan ke-) dengan nilai *bandwidth* 0,9



**Gambar 7** Kurva Error Regresi Kernel *Triangle* terhadap Rata-rata Bulanan Bilangan *Sunspot* dengan *Bandwidth* 0,9

Dapat dilihat dari Gambar (6) bahwa kurva regresi (titik berwarna biru) hampir mengikuti pola data rata-rata bulanan bilangan *sunspot* sebenarnya (titik berwarna hijau). Pada kurva *error* dapat dilihat bahwa nilai *error* yang dihasilkan sangat kecil

## KESIMPULAN

Berdasarkan hasil analisis data dan pembahasan pada bab sebelumnya, maka penulis dapat menarik kesimpulan bahwa estimator kurva *Priestley Chao* dengan menggunakan fungsi kernel *triangle* menghasilkan kurva regresi yang hampir mengikuti pola kurva data asli yang sebenarnya. Nilai *bandwidth* optimum yang diperoleh adalah 0,9 dengan nilai GCV minimum sebesar 1,0321 serta nilai MSE yang sangat kecil. Sedangkan pada penelitian sebelumnya Nisa' (2016) telah melakukan penelitian fungsi kernel *triangle* dengan menggunakan estimator *Nadaraya Watson* pada data rata-rata bulanan bilangan *sunspot* dengan memperoleh hasil Nilai *bandwidth* optimum yang diperoleh adalah 1 dengan nilai GCV minimum sebesar 0. Hal ini menunjukkan bahwa fungsi kernel *triangle* menggunakan estimator *Nadaraya Watson* lebih baik digunakan untuk meneliti data rata-rata bulanan bilangan *sunspot* matahari dari tahun 1954 sampai 2008 dibandingkan dengan fungsi kernel *triangle* menggunakan estimator *Priestley Chao*.

## DAFTAR PUSTAKA

- Ayuningtyas, Tri. 2018. *Regresi Nonparametrik Kernel Nadaraya-Watson dalam Data Time Series (Studi Kasus: Indeks Harga Saham Gabungan terhadap KURS, Inflasi, dan Tingkat Suku Bunga Periode Januari 2015-Maret 2018)*. Skripsi. Yogyakarta: Universitas Islam Indonesia.
- Marlin, A, F. 2016. *Penduga Kurva Regresi Nonparametrik Linear dan Nonlinear dengan metode Priestly-Chao, Nadaraya-Watson, dan Metode Fourier*. Skripsi. Bandar Lampung: Universitas Lampung.
- Nisa', Sofiyatin. 2016. *Estimator Kernel Epanechnikov dan Kernel triangle pada Data Rata-Rata Bulanan Bilangan Sunspot, NOAA*. Skripsi. Malang: Universitas Islam Negeri Maulana Malik Ibrahim.
- Puspitasari, Icha, Suparti, dan Yuciana Wulandari. 2012. Analisis Indeks Harga Saham Gabungan (IHSG) dengan Menggunakan Model Regresi Kernel. *Jurnal Gaussian*. Vol. 1, No. 1.
- Rifai, N, A, K. 2019. Pendekatan Regresi Nonparametrik dengan Fungsi Kernel untuk Indeks Harga Saham gabungan. *Jurnal Statistika*. Vol. 19, No.1.