FORECASTING THE NUMBER OF AIRPLANE PASSENGERS AT MINANGKABAU INTERNATIONAL AIRPORT USING THE P-SPLINE METHOD

(Studi Kasus: Jumlah Penumpang Pesawat Terbang di Bandara Internasioanl Minangkabau periode Januari 2016 – April 2021)

Muhamad Fajri (18337015), Muhammad Tibri Syofyan (18337021), Jimmi Darma Putra (18337052)
fjrmhd06@gmail.com, tibri.work@gmail.com, jimmidarma862@gmail.com
Mahasiswa Jurusan Statistika FMIPA Universitas Negeri Padang

ABSTRAK

Berdasarkan sensus penduduk pada bulan September 2020, telah terjadi peningkatan jumlah penduduk Indonesia sekitar 32,56 juta jiwa dengan pertumbuhan penduduk sebesar 1,25% per-tahun. Kementrian Perhubungan melakukan pencatatan tren peningkatan kebutuhan masyarakat terhadap jasa transportasi udara untuk mendukung efektivitas dan kelancaran berbagai aktivitas masyarakat. Secara total, jumlah penumpang angkutan udara domestik Januari – Maret 2021 sebanyak 69 juta jiwa, turun 58,68% dibanding periode yang sama tahun lalu sebanyak 16,7 juta jiwa. Adapun jumlah penumpang angkutan udara ke luar negeri/ internasional pada Maret 2021 sebanyak 40,7 ribu jiwa, naik 24,46% dibanding Februari 2021. Penelitian ini akan memodelkan jumlah penumpang pesawat terbang di Bandara Internasional Minangkabau dari Januari 2016 hingga April 2021 menggunakan regresi nonparametrik Spline dengan metode *Penalized Spline*. Nilai R^2 (R Square) sebesar 97,83% dapat dikatan model sangat baik untuk melakukan peramalan. Nilai MAPE diperoleh sebesar 1,14% kurang dari 10% maka dapat disimpulkan bahwa model memiliki kemampuan peramalan yang sangat baik.

Kata Kunci : Forecasting, *Penalized Spline*, Penumpang Pesawat Terbang, Spline

PENDAHULUAN

Berdasarkan sensus penduduk pada bulan September 2020, telah terjadi peningkatan jumlah penduduk Indonesia sekitar 32,56 juta jiwa dari data sensus penduduk sebelumnya dengan pertumbuhan penduduk sebesar 1,25% per-tahun. Kementrian Perhubungan mencatat tren peningkatan kebutuhan masyarakat terhadap jasa transportasi udara untuk mendukung efektivitas dan kelancaran berbagai aktivitas masyarakat. Menteri Perhubungan menyatakan, transportasi memiliki peranan strategis, bahkan menjadi salah satu kebutuhan utama masyarakat, terutama jasa transportasi udara. Untuk itu, Kementrian Perhubungan terus berupaya meingkatkan dan mengembangkan sejumlah infrastruktur bandara guna mengimbangi meningkatnya kebutuhan masyarakat.

Secara total, jumlah penumpang angkutan udara domestik Januari – Maret 2021 sebanyak 69 juta jiwa, turun 58,68% dibanding periode yang sama tahun lalu sebanyak 16,7 juta jiwa. Jumlah penumpang terbesar tercatat di Soekarno Hatta mencapai 1,7 juta jiwa, atau 23,13% dari keseluruhan penumpang domestik. Adapun jumlah penumpang angkutan udara ke luar negeri/ internasional pada Maret 2021 sebanyak 40,7 ribu jiwa, naik 24,46% dibanding Februari 2021. Peningkatan terbesar terjadi di Bandara Ngurah Rai, yakni mencapai 100%.

Oleh karena itu, salah satu metode untuk menganalisis data adalah analisis regresi. Dalam analisis regresi, jika pola hubungan antara variabel respon dengan variabel penjelas tidak diketahui bentuk fungsinya atau polanya maka pendekatan regresi nonparametrik digunakan. Salah satu pendekatan nonparametrik yang populer adalah Spline.

Regresi spline merupakan salah satu metode regresi nonparametrik yang dapat digunakan untuk memodelkan data. Regresi *spline* adalah suatu pendekatan ke arah pencocokan data dengan tetap memperhitungkan kemulusan kurva. *Spline* merupakan potongan polinomial tersegmen yang digabungkan oleh titik-titik knot yang dapat menjelaskan karakteristik dari data (Eubank, 1999).

Dalam regresi *spline*, pemilihan banyak dan letak knot merupakan isu yang penting. Untuk menentukan knot yang optimal, perlu dilakukan perhitungan sebanyak kombinasi banyaknya knot dari banyaknya data. Kemudian dipilih model optimal berdasarkan kriteria tertentu, misalnya nilai GCV minimum. Hal ini membutuhkan waktu yang lama dan jika dilakukan menggunakan software memerlukan memori yang besar. Karena itu diperlukan alternatif untuk mengatasi masalah ini, yaitu regresi *penalized spline* dimana knot terletak di titik-titik kuantil dari nilai *unique* (tunggal) variabel prediktor (Ruppert *et al*., 2003).

Estimator regresi *penalized spline* diperoleh dengan meminimumkan fungsi *Penalized Least Square* (PLS) yang terdiri dari jumlah kuadrat residual dan penalti kekasaran yang membuat model menjadi lebih mulus berdasarkan nilai parameter penghalus λ (Ruppert *et al*., 2003). Parameter penghalus λ yang optimal diperoleh dengan kriteria *Generalized Cross-Validation* (GCV) minimum.

Penelitian ini akan memodelkan jumlah penumpang pesawat terbang di Bandara Internasional Minangkabau dari Januari 2016 hingga April 2021.

TINJAUAN PUSTAKA

1. Regresi Nonparametrik

Dalam pendekatan model regresi nonparametrik, kurva regresi tidak diketahui atau tidak terdapat informasi masa lalu yang lengkap tentang bentuk pola datanya. Data diharapkan mencari sendiri bentuk estimasinya sehingga memiliki fleksibilitas yang tinggi. Kurva regresi hanya diasumsikan termuat dalam suatu ruang fungsi yang berdimensi tak hinggadan merupakan fungsi mulus (*smooth*). Estimasi fungsi $z(x_i)$ dilakukan berdasarkan data pengamatan dengan menggunakan teknik *smoothing* tertentu. Ada beberapa teknik *smoothing* yang dapat digurnakan antara lain estimator histogram, kernel, deret orthog- onal, spline, k-NN, deret fourier, dan wavelet.

2. Regresi Spline

Menurut Eubank (1999), regresi *spline* adalah suatu pendekatan ke arah pencocokan data dengan tetap memperhitungkan kemulusan kurva. *Spline* merupakan potongan polinomial tersegmen yang digabungkan oleh titik-titik knot yang dapat menjelaskan karakteristik dari data. Knot merupakan titik perpaduan bersama yang memperlihatkan terjadinya perubahan perilaku dari fungsi *spline* pada interval-interval yang berbeda.

Secara umum fungsi *spline* dengan derajat p dan m knot dinyatakan sebagai

$$f(x) = \beta_0 + \beta_1 x + \cdots + \beta_p x^p + \sum_{j=1}^{m} \beta_{p+j}(x - k_j)_+^p$$

Dengan fungsi *truncated* adalah sebagai berikut :

$$(x - k_j)_+^p = \begin{cases} (x - k_j)^p & , x \geq k_j \\ 0 & , x < k_j \end{cases}$$

Dan $k_1, k_2, \ldots, k_m$ merupakan titik knot.

Sehingga model regresi *spline* adalah

$$y_i = f(x_i) + \varepsilon_i$$

$$y_i = \beta_0 + \beta_1 x + \cdots + \beta_p x^p + \sum_{j=1}^{m} \beta_{p+j}(x - k_j)_+^p + \varepsilon_i$$

3. Regresi *Penalized Spline*

Menurut Ruppert et al. (2003),estimator regresi *penalized spline* diperoleh dengan meminimumkan fungsi *Penalized Least Square (PLS) sebagai berikut :*

$$\sum_{i=1}^{n} (y_i - f(x_i))^2 + \lambda^{2p} \sum_{j=1}^{m} \beta_{p+j}^2 \qquad , \lambda \geq 0$$

Suku pertama pada persamaan (3) adalah jumlah kuadrat residual dan suku keduanya adalah penalti kekasaran yang membuat model menjadi lebih halus berdasarkan nilai parameter penghalus λ (Ruppert *et al*., 2003). Dalam bentuk matriks persamaan (2) dinyatakan sebagai

$$\mathbf{L} = (\boldsymbol{Y} - \boldsymbol{X\beta})^T(\boldsymbol{Y} - \boldsymbol{X\beta}) + \lambda^{2p}\boldsymbol{\beta}^T\boldsymbol{D\beta} \qquad (3)$$

dimana λ = parameter penghalus

$$\mathbf{X} = \begin{bmatrix} 1 & x_1 & \cdots & x_1^p & (x_1 - k_1)_+^p & \cdots & (x_1 - k_m)_+^p \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_n & \cdots & x_n^p & (x_n - k_1)_+^p & \cdots & (x_n - k_m)_+^p \end{bmatrix}$$

$$\boldsymbol{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_p \\ \beta_{p+1} \\ \vdots \\ \beta_{p+m} \end{bmatrix}$$

$$\mathbf{D} = \begin{bmatrix} 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \\ \vdots & \cdots & \vdots & \vdots & \vdots & \cdots & \vdots \\ 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \\ 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \\ 0 & \cdots & 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & 0 & 0 & \cdots & 1 \end{bmatrix} = \begin{bmatrix} \mathbf{0}_{(p+1)\mathrm{x}(p+1)} & \mathbf{0}_{(p+1)\mathrm{x}m} \\ \mathbf{0}_{m\mathrm{x}(p+1)} & \mathbf{I}_{m\mathrm{x}m} \end{bmatrix} = \mathrm{diag}(\mathbf{0}_{p+1}, \mathbf{1}_m).$$

Pendugaan parameter dilakukan dengan meminimumkan persamaan (3) dengan prinsip $\frac{\partial L}{\partial \beta} = 0$ sehingga diperoleh estimasi parameter β sebagai berikut :

$$\hat{\boldsymbol{\beta}} = (\boldsymbol{X}^T\boldsymbol{X} + \lambda^{2p}\boldsymbol{D})^{-1}\boldsymbol{X}^T\boldsymbol{Y}$$

Sedangkan estimasi parameter untuk y adalah:

$$\hat{\boldsymbol{Y}} = \boldsymbol{X}\hat{\boldsymbol{\beta}}$$
$$\hat{\boldsymbol{Y}} = \boldsymbol{X}(\boldsymbol{X}^T\boldsymbol{X} + \lambda^{2p}\boldsymbol{D})^{-1}\boldsymbol{X}^T\boldsymbol{Y}$$
$$\hat{\boldsymbol{Y}} = \boldsymbol{S}_\lambda\boldsymbol{Y}$$

Dengan $S_\lambda = X(X^TX + \lambda^{2p}D)^{-1}X^T$ dinamakan matriks hat yang bersifat simetris dan definit positif.

4. Pemilihan Parameter Penghalus Optimal

Parameter penghalus λ yang optimal diperoleh dengan menghitung nilai *Generalized Cross-Validation* (GCV) minimum. Menurut Green dan Silverman (1994), fungsi GCV dapat dinyatakan sebagai berikut:

$$\text{GCV} = n^{-1}\frac{\text{RSS}(\lambda)}{[1\text{-}n^{-1}tr(\boldsymbol{S}_\lambda)]^2}$$

Dengan $\text{RSS}(\lambda) = \sum_{i=1}^{n}(y_i - \hat{y}_i)^2$

$\boldsymbol{S}_\lambda = \boldsymbol{X}(\boldsymbol{X}^T\boldsymbol{X} + \lambda^{2p}\boldsymbol{D})^{-1}\boldsymbol{X}^T$

Demmler-Reinsch ortogonalization merupakan suatu metode yang dapat mempermudah dalam memilih G optimal untuk setiap parameter penghalus λ. Algoritma *Demmler-Reinsch ortogonalization* adalah sebagai berikut:

a. Menghitung dekomposisi cholesky X^TX untuk mendapatkan matriks R sedemikian sehingga $X^TX = R^TR$.
b. Menghitung dekomposisi nilai singular matriks $(R^{-1})^TDR^{-1}$ untuk mendapatkan $U\,diag(s)U^T$ sesedemikian sehingga $(R^{-1})^TDR^{-1} = U\,diag(s)U^T$.
c. Menghitung matriks $A = XR^{-1}U$ dan vektor $b = A^TY$.
d. Menghitung $\hat{Y} = A(\frac{b}{1+\lambda^{2p}s})$

$$\text{dimana } \mathbf{1} = \begin{bmatrix}1\\1\\\vdots\\1\end{bmatrix}_{(p+m+1)x1} \quad \text{dan } \mathbf{s} = \begin{bmatrix}s_1\\s_2\\\vdots\\s_{p+m+1}\end{bmatrix}_{(p+m+1)x1}$$

$$\text{dengan tr}(\boldsymbol{S}_\lambda) = \sum_{i=1}^{p+m+1}(1 + \lambda^{2p}s_i)^{-1}$$

5. Pemilihan Knot Optimal

Banyak knot (m) merupakan banyaknya titik dimana terjadi perubahan perilaku fungsi pada interval yang berlainan.Dalam regresi penalized spline,knot terletak pada titik kuantil dari nilai *unique* (tunggal) variabel prediktor $\{x_i\}_{i=1}^{n}$.

Algoritma yang digunakan untuk menentukan banyak knot optimal dalam regresi *penalized spline* adalah algoritma *full-search* .Dalam algoritma *full-search* ,dihitung nilai GCV untuk banyak knot m=1,2,3,… sampai banyak knot tertentu yang dicobakan,dengan ketentuan $m < (n_{unique} - p - 1)$,dimana n_{unique} adalah banyaknya nilai *unique* dari variabel prediktor $\{x_i\}_{i=1}^{n}$ dan p adalah derajat polinomial.Kemudian dipilih banyak knot yang memiliki nilai GCV terkecil (Ruppert et al.,2003).

6. Regresi Nonparametrik untuk Data Runtun Waktu

Menurut Hardle (1990), sifat statistik dari regresi penghalus secara umum dianalisis dalam kerangka struktur observasi yang berdistribusi identik dan independen (iid). Asumsi {(Xi,Yi), i =1,2,...,n} adalah sampel yang independen digunakan untuk penyederhanaan permasalahan. Namun data yang diperoleh dalam kehidupan sehari-hari seringkali tidak memenuhi asumsi bahwa observasi (X1,Y1), (X2,Y2), … (Xn,Yn) independen. Oleh karena itu perlu disusun suatu pemodelan data yang asumsi independensi datanya tidak dipenuhi. Terdapat tiga konsep dasar matematika yang mendasari pemodelan ini, yaitu:

a. Model (S): Suatu barisan stasioner {(Xi,Yi), i =1,2,...,n} adalah hasil observasi dan akan diestimasi f(x) = E(Y|X =x)
b. Model (T): Suatu runtun waktu {Zi, i =2,3,...n} adalah hasil observasi dan digunakan untuk memprediksi Zn+1 dengan f(x) = E(Zn+1| Zn = x).
c. Model (C): Error observasi {ein} dalam model regresi dengan rancangan tetap $Y_i = f\left(\frac{i}{n}\right) + e$, membentuk barisan variabel random yang berkorelasi.

Permasalahan model runtun waktu (T) dapat digambarkan ke dalam model (S) dengan mendefinisikan dalam runtun waktu {Zi, i =2,3,...n}, nilai lag Z_{i-1} sebagai Xi dan nilai Zi sebagai Yi. Selanjutnya masalah prediksi Z_{n+1} dari {Zi}, i =1,2,...,n dapat dianggap sebagai masalah penghalusan regresi untuk {(Xi,Yi), i =2,3,...n} = {(Zi-1,Zi), i =2,3,...n}.

7. Pemilihan Parameter *Smoothing* λ

Metode yang digunakan untuk memilih parameter *smoothing* λ adalah menggunakan *Generalized Cross Validation* (GCV). Parameter penghalus yang optimal akan diperoleh berdasarkan nilai GCV yang minimum. Fungsi GCV dapat dinyatakan sebagai berikut:

$$GCV = \frac{(N)^{-1}\sum_{i=1}^{m}\sum_{j=1}^{n}(y_{i;j} - \hat{y}_{i;j})^2}{\{1 - tr(\boldsymbol{A}_\lambda)/N\}^2}$$

dimana GCV adalah nilai GCV dengan N adalah banyaknya pengamatan sebanyak m subjek x n pengamatan, $y_{i;j}$ adalah data aktual subyek ke-i pada pengamatan ke-j dan adalah hasil estimasi subyek ke-i pada pengamatan ke-j dengan $tr(A_\lambda) = tr[X(X^TX + \lambda D)^{-1}X^T]$.

8. Pemilihan Metode Terbaik

Untuk mengetahui seberapa akurat peramalan yang dihasilkan dapat digunakan nilai koefisien determinasi (R^2). Koefisien determinasi adalah koefisien yang mengukur seberapa jauh kemampuan sebuah model dalam menerangkan variasi variabel dependen. Nilai R^2 dapat dinyatakan sebagai berikut :

$$R^2 = \frac{\sum_{i=1}^{m}\sum_{j=1}^{n}(\hat{Y}_{i;j} - \bar{Y})^2}{\sum_{i=1}^{m}\sum_{j=1}^{n}(Y_{i;j} - \bar{Y})^2}$$

Dimana R^2 adalah koefisien determinasi $Y_{i;j}$, adalah data aktual subjek ke-i dan pengamatan ke-j, $Y_{i;j}$ adalah data prediksi subjek ke-i dan pengamatan ke-j dan $\bar{Y}$ adalah rata-rata data aktual. Salah satu cara untuk mengetahui ketepatan peramalan sebuah model dapat menggunakan nilai *Mean Absolute Percentage Error* (MAPE). Rumus MAPE dituliskan sebagai berikut :

$$MAPE = \frac{\sum_{i=1}^{m}\sum_{j=1}^{n}\left|\left(\frac{y_{i;j} - \hat{y}_{i;j}}{y_{i;j}}\right)\right| x100\%}{N}$$

dengan $y_{i;j}$ adalah data aktual subjek ke-i pada pengamatan ke-j,$\hat{y}_{i;j}$ adalah hasil estimasi subjek ke-i pada pengamatan ke-j dan N adalah banyaknya pengamatan (m subjek x n pengamatan). Semakin kecil nilai MAPE semakin akurat peramalan sebuah model. Untuk MAPE < 10% maka kemampuan peramalan sangat baik, 10% ≤ MAPE < 20% artinya bahwa kemampuan peramalan baik, 20% ≤ MAPE < 50% artinya bahwa kemampuan peramalan cukup dan bila MAPE ≥ 50% maka artinya bahwa kemampuan peramalan buruk .

METODE PENELITIAN

Data yang kami gunakan dalam penelitian ini adalah data sekunder yang diambil di situs Badan Pusat Statistika Sumatera Barat. Jenis data ini adalah data *time series* Jumlah penumpang pesawat terbang di Bandara Internasional Minangkabau dari Januari 2016 sampai April 2021.

Tahapan analisis yang digunakan dalam penelitian ini adalah :

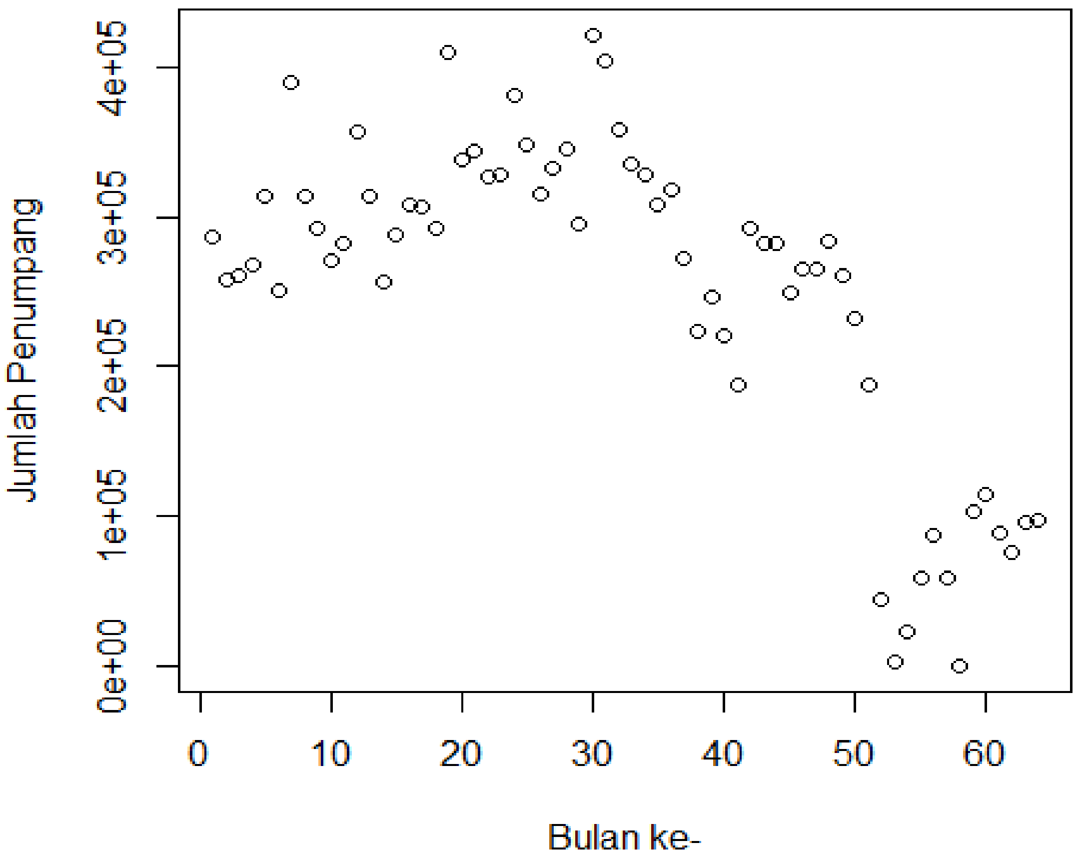
1. Membuat *scatter plot* antara jumlah penumpang (y) dengan tahun (x).
2. Menentukan Jumlah Knot yang digunakan
3. Menghitung nilai GCV untuk jumlah knot yang diperoleh

4. Memilih jumlah knot optimum berdasarkan nilai GCV minimun
5. Menentukan lokasi knot
6. Menghitung parameter pemulus yang optimal berdasarkan jumlah knot optimal menggunakan kriteria GCV
7. Mendapatkan estimator spline terpenalti menggunakna metode *penalized least square* (PLS)
8. Mendapatkan estimasi model nonparametrik
9. Menghitung Ketepatan Peramalan Model Regresi Spline Terbaik dengan Metode *Penalized Spline*
10. Menarik kesimpulan.

ANALISIS DAN PEMBAHASAN

1. *Scatterplot*

Dibawah ini akan ditampilkan *scatterplot* untuk memperlihatkan hubungan antara jumlah penumpang (y) dengan tahun (x) sebagaimana yang diperlihatkan pada Gambar 1.



Gambar 1. Diagram pencar variabel jumlah penumpang setiap bulan

Dari *scatterplot* pada gambar 1 dapat kita lihat bahwa titik-titik tidak memiliki pola tertentu atau kurva regresi tertentu sehingga bisa didekati dengan regresi non-parametrik salah satunya dengan pendekatan regresi spline dengan metode *penalized spline.*

2. Pemilihan Banyak Knot dan Parameter Penghalus λ Optimal

Pemilihan banyak knot dan parameter penghalus λ optimal sangat penting pada regresi spline dengan metode *penalized spline*. Penentuan banyak knot dan parameter penghalus λ optimal ditentukan oleh nilai GCV minimum pada tiap ordenya. Diperoleh hasil operasi penentuan banyak knot dan parameter penghalus λ optimal menggunakan suatu R Software pada tabel 1.

Tabel 1. Banyak Knot Optimal, Parameter Penghalus λ Optimal dan GCV Minimun

Orde	Banyak Knot	λ Optimal	GCV
2	5	2,3	2.318.670.053

Berdasarkan tabel 1 dapat kita ketahui bahwa orde terbaik untuk membentuk model terbaik terletak pada orde 2 dengan jumlah knot sebanyak 5 knot, λ optimal bernilai 2,3, dan nilai GCV sebesar 2.318.670.053.

3. Model Regresi Spline Terbaik dengan Metode *Penalized Spline*

Nilai titik knot dapat dilihat pada Tabel 2 sebagai berikut:

Tabel 2. Nilai Titik Knot untuk Model Terbaik

Titik Knot ke-	Nilai Titik Knot
1	11.5
2	22
3	32.5
4	43
5	53.5

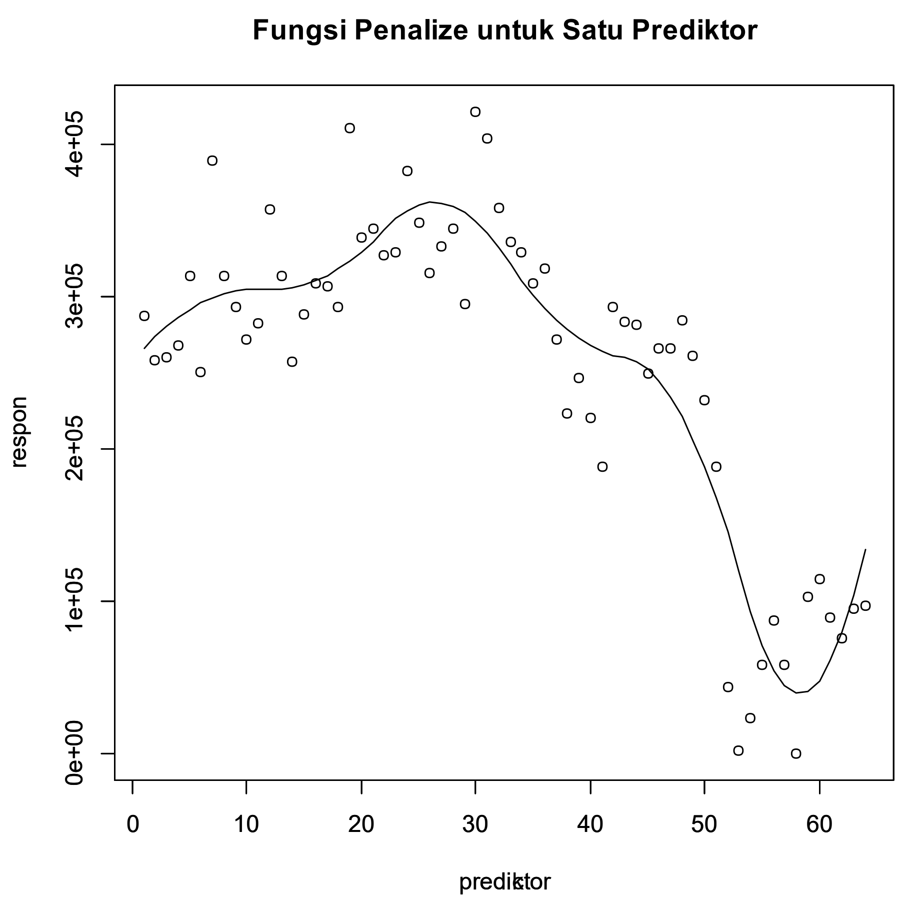
Selanjutnya untuk membentuk model diperlukan nilai parameter. Nilai Parameter dapat dilihat pada Tabel 3 sebagai berikut:

Tabel 3. Nilai Parameter untuk Model Terbaik

Parameter	Estimasi Parameter
$\beta_{1;0}$	257148.7013
$\beta_{1;1}$	8955.5125
$\beta_{1;2}$	-421.0776
$\beta_{1;3}$	847.1465
$\beta_{1;4}$	-1366.8163
$\beta_{1;5}$	1434.0454
$\beta_{1;7}$	-1771.2301
$\beta_{1;8}$	4204.2919

Dari tabel nilai knot dan nilai parameter yang sudah diperoleh maka dibuat model regresi spline terbaik dengan metode *penalized spline* sebagai berikut:

$$\hat{y} = \begin{cases} 257148{,}7013 + 8955{,}5125x - 421{,}0776x^2 & ; x < 11{,}5 \\ 369183{,}8259 - 10528{,}857x + 426{,}0695x^2 & ; 11{,}5 \leq x < 22 \\ -292355{,}2633 + 49611{,}0602x - 940.7468x^2 & ; 22 \leq x < 32{,}5 \\ 1222355{,}1905 - 43601{,}8908x + 493{,}2986x^2 & ; 32{,}5 \leq x < 43 \\ -2052649{,}2645 + 108723{,}8978x - 1277{,}9315x^2 & ; 43 \leq x < 53{,}5 \\ 9981085{,}2263 - 341135{,}3355x + 2926{,}3604x^2 & ; x \geq 53{,}5 \end{cases}$$



Gambar 2. Plot P-Spline kuadratik dengan 5 titik knot

4. Menghitung Ketepatan Peramalan Model Regresi Spline Terbaik dengan Metode *Penalized Spline*

   Untuk mengetahui seberapa akurat model yang dihasilkan dapat digunakan koefisien determinasi (R^2) dan dapat menggunakna nilai *Mean Absolute Percentage Error* (MAPE). Berdasarkan perhitungan manual R^2 didapatkan bahwa nilai R^2 sebesar 97,83%. Karena nilai R^2 mendekati nilai 100% maka dapat disimpulkan bahwa model sangat baik untuk data peramalan. Sedangkan untuk nilai MAPE berdasarkan perhitungan manual diperoleh sebesar 1,14%. Nilai MAPE tersebut kurang dari 10% maka dapat disimpulkan bahwa model memiliki kemampuan peramalan yang sangat baik.

KESIMPULAN

Model regresi nonparametrik spline dengan metode *penalized spline* pada data *time series* dalam kasus pemodelan jumlah penumpang pesawat di Bandara Internasional Minangkabau terletak pada orde 2 dengan 5 knot dan nilai lambda 2,3 dengan nilai GCV minimun adalah 2.318.670.053.

Nilai R^2 yang diperoleh adalah sebesar 97,83% yang memiliki arti bahwa model sangat baik untuk data peramalan. Sedangkan nilai MAPE yang diperoleh 1,14%. Nilai MAPE tersebut kurang dari 10% sehingga dapat dinyatakan bila model ini memiliki kemampuan peramalan yang sangat baik.

DAFTAR PUSTAKA

Eubank, R.L., 1999, Spline Smoothing and Nonparametric Regression Second Edition. New York: Marcel Dekker.

Green, P. J., and Silverman, B. W., 1994, Nonparametric Regression and Generalized Linear Models: A Roughness Penalty Approach. London: Chapman & Hall.

Hardle, W., 1990, Applied Nonparametric Regression. New York: Cambridge University.

Ruppert, D., Wand, M.P., and Carroll, R.J., 2003, Cambridge Series in Statistical and Probabilistic Mathematics: Semiparametric Regression. New York: Cambridge University.