

## Peramalan Jumlah Penumpang Kereta Api di Pulau Jawa Menggunakan Model BATS dan Model TBATS

Fitria Nurjanah\*, Marizsa Herlina

Prodi Statistika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Islam Bandung, Indonesia.

\*fitrianurjanah4321@gmail.com, Marizsa.herlina@unisba.ac.id

**Abstract.** Time series forecasting aims to find patterns in historical data series and extrapolate these patterns into the future. The success of a forecast depends on the use of forecasting methods, which must be adapted to the available data. Time series data that contains several seasonalities or has a dual-calendar effect is called data with complex seasonality, but not all forecasting methods can handle seasonal complexity. The methods that are able to handle seasonal complexity are the BATS (Box-Cox transform, ARMA errors, Trend, and Seasonal components) and TBATS (Trigonometric, Box-Cox transform, ARMA errors, Trend, and Seasonal components) models. Based on the results of the analysis carried out, it was found that the BATS model was better at predicting the number of train passengers on the island of Java compared to the TBATS model, as evidenced by the AIC value of 4318,364 and the RMSE of 1875,887 which were smaller than the AIC and RMSE values of the TBATS model. The results of forecasting using the BATS model showed that the highest number of train passengers was in December, namely 31,107,740 people and the lowest was in February 27,959,460 people.

**Keywords:** *BATS, Railways, Forecasting, TBATS.*

**Abstrak.** Peramalan deret waktu bertujuan untuk menemukan pola dalam deret data historis dan mengekstrapolasikan pola tersebut ke masa depan. Keberhasilan suatu peramalan bergantung pada penggunaan metode peramalan yang harus disesuaikan dengan data yang tersedia. Data deret waktu yang mengandung beberapa musiman atau terdapat efek *dual-calendar* disebut data dengan musiman yang kompleks, namun tidak semua metode peramalan dapat menangani kompleksitas musiman. Adapun metode yang mampu menangani kompleksitas musiman adalah model BATS (*Box-Cox transform, ARMA errors, Trend, and Seasonal components*) dan TBATS (*Trigonometric, Box-Cox transform, ARMA errors, Trend, and Seasonal components*). Berdasarkan hasil analisis yang dilakukan didapati hasil bahwa model BATS lebih baik dalam meramalkan jumlah penumpang kereta api di pulau Jawa dibandingkan dengan model TBATS dibuktikan dengan nilai AIC sebesar 4318.364 dan RMSE sebesar 1875.887 yang lebih kecil dibandingkan dengan nilai AIC dan RMSE dari model TBATS. Adapun hasil peramalan menggunakan model BATS didapat bahwa jumlah penumpang kereta api tertinggi adalah pada bulan Desember yaitu sebanyak 31.107.740 orang dan terendah pada bulan Februari 27.959.460 orang.

**Kata Kunci:** *BATS, Kereta Api, Peramalan, TBATS.*

## A. Pendahuluan

Menurut Hyndman [1] peramalan adalah memprediksi masa depan dengan akurat menggunakan semua informasi yang tersedia termasuk data historis dan pengetahuan tentang masa depan yang mungkin berdampak pada ramalan. Adapun tujuan dari dilakukannya peramalan adalah untuk kepentingan perencanaan sehingga dapat dilakukan tindakan yang tepat.

Peramalan tradisional sederhana untuk data dengan musiman tunggal (single seasonal) dikenal sebagai metode exponential smoothing Holt dan Winters. Pada dasarnya metode exponential smoothing Holt-Winters mampu menangani data deret waktu dengan satu pola musiman serta dua komponen lainnya (level dan trend). Metode-metode ini telah banyak digunakan untuk menghasilkan peramalan yang dapat diandalkan dengan deret waktu frekuensi rendah (bulanan atau triwulanan) [2]. Namun metode-metode tersebut tidak mampu mengatasi masalah kompleksitas musiman.

Salah satu innovations state space model untuk melakukan peramalan pada data deret waktu yang diindikasikan memiliki kompleksitas musiman adalah model BATS (Box-Cox transform, ARMA errors, Trend, and Seasonal components) dan model TBATS (Trigonometric, Box-Cox transform, ARMA errors, Trend, and Seasonal components).

Pada bulan-bulan tertentu, seperti saat libur sekolah seperti Natal, Idul Fitri, Imlek dan liburan lainnya, jumlah penumpang transportasi darat, laut dan udara biasanya meningkat. Khususnya pulau Jawa yang merupakan pulau dengan proporsi penduduk tertinggi di Indonesia dan merupakan pusat kegiatan ekonomi masyarakat membuat pulau Jawa menjadi wilayah yang paling merepresentasikan kegiatan transportasi masyarakat Indonesia secara umum. Kemungkinan adanya pola musiman bahkan musiman ganda serta adanya efek dual calendar pada data jumlah penumpang transportasi umum membuat peneliti merasa tertarik untuk meramalkan jumlah penumpang terkhusus jumlah penumpang kereta api di pulau Jawa dengan menggunakan model BATS serta model TBATS.

Meramalkan jumlah penumpang khususnya penumpang kereta api menjadi penting untuk mengetahui perkembangan atau pertumbuhan jumlah penumpang dari waktu ke waktu guna pemerintah atau pihak penyedia layanan transportasi untuk melakukan perencanaan baik dari segi kelayakan fasilitas maupun jumlah fasilitas agar selaras dengan jumlah penumpang atau pengguna layanan transportasi.

Berdasarkan latar belakang yang telah diuraikan, maka perumusan masalah dalam penelitian ini sebagai berikut: “Bagaimana perbandingan model BATS dengan model TBATS dalam meramalkan jumlah penumpang kereta api di pulau Jawa?” dan “Bagaimana meramalkan jumlah penumpang kereta api di pulau Jawa menggunakan model terbaik?”. Selanjutnya, tujuan dalam penelitian ini diuraikan dalam pokok-pokok sbb.

1. Untuk mengetahui perbandingan model BATS dengan model TBATS dalam meramalkan jumlah penumpang kereta api di pulau Jawa.
2. Untuk mengetahui hasil peramalan jumlah penumpang kereta api di pulau Jawa menggunakan model terbaik.

## B. Metodologi Penelitian

Data dalam penelitian ini adalah data sekunder yang didapat dari website bps.go.id. Data deret waktu yang digunakan dalam penelitian ini adalah data jumlah penumpang kereta api per bulan di pulau Jawa. Periode data yang digunakan adalah dari Januari 2006 sampai dengan Juni 2023 dengan banyaknya data adalah sebanyak 210 data.

### Model BATS

De Livera [3] memperkenalkan model BATS (*Box-Cox transform, ARMA errors, Trend, and Seasonal components*) yang merupakan *innovations state space model* yang menjadi alternatif dari metode *exponential smoothing* tradisional dimana model BATS bisa mengakomodasi beberapa musim dan dimodifikasi dengan penambahan transformasi Box-Cox untuk menangani non-linearitas serta penyesuaian residual ARMA untuk memperhitungkan setiap autokorelasi dalam residu dengan persamaan seperti berikut:

$$y_t^{(\omega)} = \begin{cases} \frac{y_t^{(\omega)} - 1}{\omega}; & \omega \neq 0 \\ \log y_t & \omega = 0 \end{cases} \quad (1)$$

$$y_t^{(\omega)} = \ell_{t-1} + \phi b_{t-1} + \sum_{i=1}^T s_{t-1}^{(i)} + d_t \quad (2)$$

$$l_t = l_{t-1} + \phi b_{t-1} + \alpha d_t \quad (3)$$

$$b_t = (1 - \phi)b + \phi b_{t-1} + \beta d_t \quad (4)$$

$$s_t^{(i)} = s_{t-m_i}^{(i)} + \gamma_i d_t \quad (5)$$

$$d_t = \sum_{i=1}^p \varphi_i d_{t-i} + \sum_{i=1}^q \theta_i \varepsilon_{t-i} + \varepsilon_t \quad (6)$$

Persamaan (1) adalah transformasi Box-Cox, persamaan (2) adalah pola musiman, persamaan (3) dan (4) adalah indeks *trend* jangka panjang dan pendek, persamaan (5) adalah indeks musiman dan persamaan (6) adalah residu yang dimodelkan ARMA.

### Model TBATS

Model TBATS (*Trigonometric, Box-Cox transform, ARMA errors, Trend, and Seasonal components*) adalah kebaruan dari model BATS (*Box-Cox transform, ARMA errors, Trend, and Seasonal components*) dimana dalam model TBATS terdapat penambahan kata *Trigonometric* yang mengisyaratkan adanya penambahan model *trigonometric seasonal* [4]. Penambahan model *trigonometric seasonal* bertujuan untuk memungkinkan model secara keseluruhan dapat mengakomodasi musim non-bilangan bulat, dan efek *dual-calendar* yang sebelumnya tidak dapat dilakukan model BATS. Model TBATS menggunakan persamaan-persamaan yang sama dengan model BATS, hanya pada persamaan (5) akan digantikan oleh persamaan berikut:

$$s_t^{(i)} = \sum_{j=1}^{k_i} s_{j,t}^{(i)} \quad (7)$$

$$s_{j,t}^{(i)} = -s_{j,t-1}^{(i)} \cos \lambda_j^{(i)} + s_{j,t-1}^{*(i)} \sin \lambda_j^{(i)} + \gamma_1^{(i)} d_t \quad (8)$$

$$s_{j,t}^{*(i)} = -s_{j,t-1}^{(i)} \sin \lambda_j^{(i)} + s_{j,t-1}^{*(i)} \cos \lambda_j^{(i)} + \gamma_2^{(i)} d_t \quad (9)$$

Persamaan (7) adalah indeks musiman, persamaan (8) adalah stokastik komponen musiman ke- $i$  dan persamaan (9) adalah pertumbuhan stokastik untuk komponen musiman ke- $i$ .

### Pemilihan Model Terbaik

Adapun untuk menentukan model terbaik untuk penelitian ini adalah menggunakan *Akaike's Information Criterion* (AIC) dan *Root Mean Square Error* (RMSE) untuk melihat akurasi hasil peramalan dengan persamaan seperti berikut:

$$AIC = \mathcal{L}^*(\hat{\boldsymbol{\theta}}, \hat{\boldsymbol{x}}_0) + 2K \quad (10)$$

$$RMSE_h = \sqrt{\frac{1}{n} \sum_{t=1}^n (y_t - \hat{y}_t)^2} \quad (11)$$

Dengan kriteria semakin rendah tingkat *error* pada hasil peramalan, maka semakin layak hasil peramalan tersebut digunakan dan model yang digunakan untuk peramalan semakin baik.

### C. Hasil Penelitian dan Pembahasan

Adapun dalam menentukan panjang musim ( $m$ ) peneliti menentukan  $m = 12$ . Karena data yang dipakai merupakan data bulanan, serta satu musiman terjadi dalam 1 tahun (12 bulan). Namun tidak ditemukan adanya efek *dual-calendar* karena pada bulan dimana terdapat hari libur nasional atau hari raya keagamaan tidak terjadi kenaikan atau penurunan jumlah kereta api yang signifikan. Sehingga didapat hasil sebagai berikut:

**Tabel 1.** Koefisien Parameter Model BATS

	Parameter
$\omega$	1
$\alpha$	1.191731
$\beta$	(tidak ada)
$\phi$	(tidak ada)
$p$	(tidak ada)
$q$	(tidak ada)
$\gamma$	-0.02652024

Sumber: Data Penelitian yang Sudah Diolah, 2023.

Berdasarkan koefisien parameter yang diperoleh pada Tabel 4.1, dapat dilihat bahwa  $\omega = 1$  yang berarti tidak dilakukan transformasi Box-Cox pada data sehingga notasi model yang terbentuk adalah BATS (1, {0,0}, -, {12}).

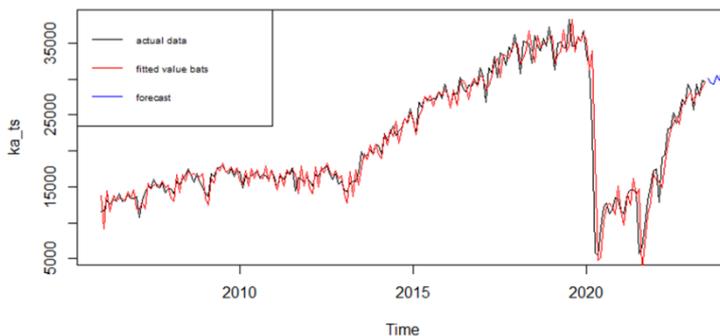
**Tabel 2.** Koefisien Parameter Model TBATS

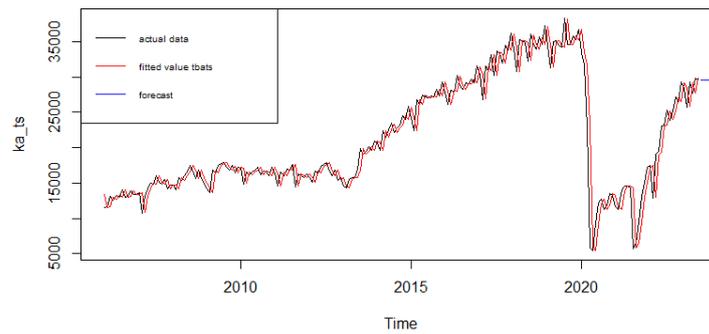
	Parameter
$\omega$	1
$\alpha$	0.9723592
$\beta$	(tidak ada)
$\phi$	(tidak ada)
$\varphi$	(tidak ada)
$\theta$	(tidak ada)
$\gamma_1$	(tidak ada)
$\gamma_2$	(tidak ada)

Sumber: Data Penelitian yang Sudah Diolah, 2023.

Pemodelan TBATS yang dilakukan oleh *software* R tidak terbentuk, melainkan terbentuk model BATS kembali. Peneliti menduga hal tersebut terjadi karena *software* R tidak bisa mendeteksi atau tidak bisa mengaplikasikan *trigonometric seasonal* pada data. Sehingga notasi model yang terbentuk adalah BATS (1, {0,0}, -, -). Karena yang terbentuk adalah model BATS maka selanjutnya hanya akan digunakan model BATS untuk melakukan peramalan.

### Pemilihan Model Terbaik

**Gambar 1.** Perbandingan Data Aktual dengan Model BATS(1, {0,0}, -, {12})



**Gambar 2.** Perbandingan Data Aktual dengan Model BATS(1, {0,0}, -, -)

Perbandingan secara visual saja tidak cukup karena bersifat subjektif, maka akan dihitung nilai AIC dan RMSE untuk menentukan model terbaik. Model terbaik dilihat berdasarkan pada residual yang dihasilkan kedua model dengan menggunakan kriteria model AIC dan RMSE. Kriteria kebaikan dari model-model yang terbentuk ditunjukkan pada Tabel 3 berikut.

**Tabel 3.** AIC dan RMSE Model

Kriteria	BATS(1, {0,0}, -, {12})	BATS(1, {0,0}, -, -)
AIC	4318.364	4368.616
RMSE	1875.887	2249.332

Sumber: Data Penelitian yang Sudah Diolah, 2023.

Dari Tabel 3 dapat disimpulkan bahwa model BATS(1, {0,0}, -, {12}) lebih baik digunakan untuk data jumlah penumpang kereta api di pulau Jawa karena memiliki nilai AIC dan RMSE yang lebih kecil dari model BATS (1, {0,0}, -, -). Sehingga untuk melakukan peramalan akan menggunakan model BATS(1, {0,0}, -, {12}).

### Peramalan

Peramalan dilakukan dengan menggunakan model BATS (1, {0,0}, -, {12}) sesuai dengan hasil perbandingan model yang menyatakan bahwa model BATS lebih baik digunakan untuk data jumlah penumpang kereta api di pulau Jawa, serta digunakan selang kepercayaan sebesar 95%. Berikut hasil peramalan jumlah penumpang kereta api di pulau Jawa untuk tiga puluh bulan kedepan:

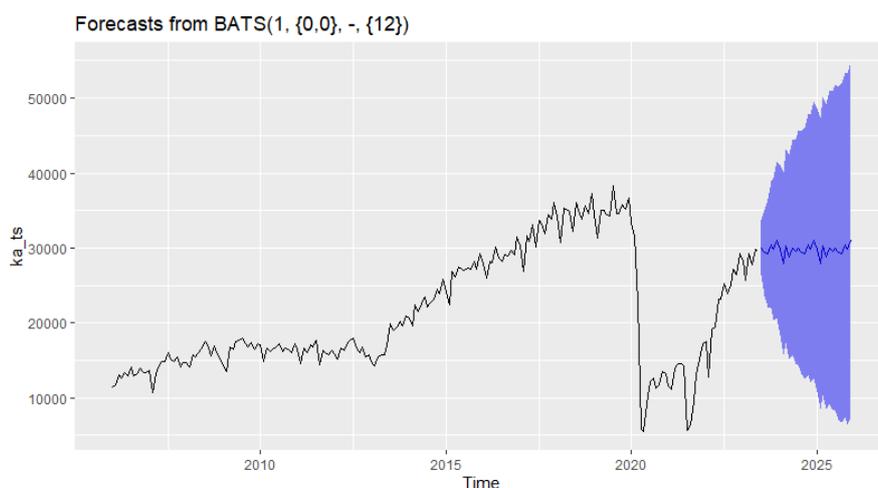
**Tabel 4.** Hasil Ramalan Jumlah Penumpang Kereta Api di Pulau Jawa

Bulan - Tahun	Nilai Peramalan	Batas Bawah Peramalan	Batas Atas Peramalan
Jul-23	30042.1	26365.4	33718.8
Aug-23	29377.7	23657.8	35097.5
Sep-23	29308.2	22103	36513.4
Oct-23	30450.6	22017.8	38883.5
Nov-23	29859	20355.8	39362.3
Dec-23	31107.7	20643	41572.5
Jan-24	29776.6	18431.6	41121.6
Feb-24	27959.5	15797.8	40121.2
Mar-24	30278.4	17351.5	43205.3

Bulan - Tahun	Nilai Peramalan	Batas Bawah Peramalan	Batas Atas Peramalan
Apr-24	28812.4	15163.1	42461.7
May-24	30042.4	15707.1	44377.8
Jun-24	29579.1	14589.1	44569.1
Jul-24	30042.1	14451.9	45632.3
Aug-24	29377.7	13183.4	45571.9
Sep-24	29308.2	12531.7	46084.7
Oct-24	30450.6	13111.4	47789.9
Nov-24	29859	11974.7	47743.3
Dec-24	31107.7	12694.5	49521
Jan-25	29776.6	10849.2	48703.9
Feb-25	27959.5	8531.56	47387.4
Mar-25	30278.4	10362.5	50194.3
Apr-25	28812.4	8420.25	49204.6
May-25	30042.4	9184.82	50900
Jun-25	29579.1	8266.24	50891.9
Jul-25	30042.1	8302.94	51781.3
Aug-25	29377.7	7201.34	51554
Sep-25	29308.2	6703.18	51913.3
Oct-25	30450.6	7424.87	53476.4
Nov-25	29859	6420.06	53298
Dec-25	31107.7	7262.76	54952.7

Sumber: Data Penelitian yang Sudah Diolah, 2023.

Berikut disajikan plot hasil ramalan jumlah penumpang kereta api di pulau Jawa untuk tiga puluh periode atau tiga puluh bulan, yaitu tahun 2023, 2024 dan 2025.



**Gambar 3.** Hasil Peramalan Jumlah Penumpang Kereta Api di Pulau Jawa

Dapat terlihat dari Gambar 3 bahwa hasil ramalan (garis berwarna biru) jumlah penumpang kereta api di pulau Jawa tidak mengalami kenaikan dari tahun ke tahun. Satu pola

terus berulang untuk setiap tahun, seperti nilai peramalan yang ditunjukkan pada Tabel 4, nilai ramalan untuk bulan yang sama di setiap tahun tidak mengalami perubahan nilai. Misalnya nilai ramalan jumlah penumpang kereta api di pulau Jawa untuk bulan Juli tahun 2023, 2024 dan 2025 bernilai sama yaitu sebanyak 30.042.100 orang.

Adapun batas atas dan batas bawah peramalan dengan menggunakan selang kepercayaan sebesar 95% terlihat semakin besar dari waktu ke waktu (daerah yang diarsir terlihat semakin besar dari waktu ke waktu).

#### D. Kesimpulan

Berdasarkan pembahasan dalam penelitian ini, peneliti menyimpulkan beberapa hasil penelitian sebagai berikut:

1. Untuk data jumlah penumpang kereta api di pulau Jawa yang tidak terdapat efek dual-calendar pada data serta berfrekuensi bulanan, model BATS terbukti lebih baik digunakan untuk melakukan peramalan dibuktikan dengan nilai AIC dan RMSE yang lebih rendah (ditunjukkan pada Tabel 3).
2. Hasil ramalan jumlah penumpang kereta api menggunakan model terbaik berdasarkan nilai AIC dan RMSE terkecil yaitu model BATS, didapat jumlah penumpang kereta api tertinggi adalah pada bulan Desember 2023, 2024 dan 2025; jumlah penumpang kereta api terendah adalah pada bulan Februari 2024 dan 2025; kenaikan jumlah penumpang kereta api tertinggi adalah pada bulan Maret 2024 dan 2025; penurunan jumlah penumpang kereta api tertinggi adalah pada bulan Februari 2024 dan 2025 (ditunjukkan pada Tabel 4).

#### Acknowledge

Peneliti mengucapkan terima kasih kepada semua pihak yang telah membantu dalam penyelesaian penelitian ini.

#### Daftar Pustaka

- [1] Hyndman, R. J., & Athanasopoulos. (2018). *Forecasting : Principles and Practice*. Australia: Monash University.
- [2] Baek, M. (2010). *Forecasting Hourly Electricity Loads of South Korea: Innovations State Space Modeling Approach*. The Korean Journal of Economics, 17.
- [3] De Livera, A., & Hyndman, R. (2009). *Forecasting Time Series With Complex Seasonal Patterns Using Exponential Smoothing*.
- [4] De Livera, A. M., Hyndman, R. J., & Snyder, R. D. (2010). *Forecasting Time Series With Complex Seasonal Patterns using Exponential Smoothing*.
- [5] Philip Trisa Parera, & Suwanda. (2023). Model Random Coefficient Autoregressive Orde Pertama dan Penerapannya. *Jurnal Riset Statistika*, 155–162. <https://doi.org/10.29313/jrs.v3i2.3048>
- [6] Jasmine Wildani Arisa, & Nusar Hajarisman. (2023). Determinasi Derajat Kelangsungan Hidup Anak Menggunakan Multigroup Structural Equation Modeling. *Jurnal Riset Statistika*, 147–154. <https://doi.org/10.29313/jrs.v3i2.3047>
- [7] Sahwa Chanigia Viqri, Z., & Kurniati, E. (2023). Perbandingan Penerapan Metode Fuzzy Time Series Model Chen-Hsu dan Model Lee dalam Memprediksi Kurs Rupiah terhadap Dolar Amerika. 1(1), 19–26. <https://doi.org/10.29313/datamath.v1i1.12>