

# Perbandingan Peramalan Jumlah Kasus Kecelakaan Lalu Lintas Kota Balikpapan dengan Linear Trend Analysis dan Double Exponential Smoothing

Mega Silfiani<sup>1, a)</sup> dan Gebryani Rante Lembang<sup>2, b)</sup>

<sup>1,2</sup>Program Studi Statistika, Institut Teknologi Kalimantan, Jalan Soekarno Hatta KM.15, Balikpapan

<sup>a)</sup>Penulis korespondensi: megasilfiani@lecturer.itk.ac.id

<sup>b)</sup>16201001@student.itk.ac.id

**Abstrak.** Kecelakaan mengakibatkan berbagai kerugian sehingga penting untuk mengantisipasi kecelakaan secara akurat. Penelitian ini bertujuan untuk membandingkan dua metode untuk meramalkan jumlah kasus kecelakaan lalu lintas yang akan terjadi di Kota Balikpapan. Metode tersebut adalah *linear trend analysis* dan *double exponential smoothing*. Jumlah kasus kecelakaan lalu lintas yang terjadi di Kota Balikpapan dari Januari 2019 hingga Agustus 2022 merupakan dataset yang digunakan untuk penelitian ini. Hasil pemodelan menunjukkan bahwa *linear trend analysis* dengan RMSE 2,73 merupakan model yang paling akurat untuk memperkirakan jumlah kasus kecelakaan lalu lintas di Kota Balikpapan. Nilai RMSE untuk prediksi menggunakan metode *double exponential smoothing* adalah 2,86. Penelitian lanjutan yang dikembangkan dari penelitian ini dapat menggunakan pengaruh peraturan pembatasan sosial masyarakat akibat COVID 19 dan juga menerapkan metode *machine learning* yang cocok untuk sampel kecil.

**Kata Kunci :** *Double Exponential Smoothing, Linear Trend Analysis, Peramalan, dan RMSE.*

## PENDAHULUAN

Setiap tahun sekitar 1,3 juta penduduk dunia meninggal akibat kecelakaan lalu lintas. Selain itu, antara 20-50 juta penduduk menjadi korban dengan luka ringan sampai luka berat dan banyak korban yang mengalami kecacatan akibat kecelakaan lalu lintas tersebut [1]. Kecelakaan lalu lintas menyebabkan kerugian ekonomi yang cukup besar bagi individu, keluarga, dan negara secara keseluruhan. Kerugian ini timbul dari biaya pengobatan serta hilangnya produktivitas bagi mereka yang terbunuh atau cacat dan bagi anggota keluarga yang perlu mengambil cuti kerja atau sekolah untuk merawat korban yang terluka. Peramalan yang akurat terhadap jumlah kasus kecelakaan yang akan terjadi di masa yang akan datang menjadi sangat krusial. Dengan adanya peramalan jumlah kasus kecelakaan, pemerintah dan pihak kepolisian memiliki early warning (peringatan dini) sehingga jumlah kasus kecelakaan lalu lintas yang akan terjadi dapat dicegah dengan berbagai upaya.

Peramalan jumlah kasus kecelakaan lalu lintas dapat mengaplikasikan data deret waktu dimana hasil ramalan di masa yang akan datang didapatkan dari data historis data itu sendiri. Beberapa penelitian yang berfokus pada peramalan jumlah kasus kecelakaan lalu lintas antara lain Haryono, Widodo dan Abusini meramalkan jumlah kasus kecelakaan lalu lintas Kota Malang menggunakan metode Automatic Clustering-Fuzzy Time Series-Markov Chain [2] dan Maharsi, Mukid dan Wilandari yang meramalkan jumlah kasus kecelakaan lalu lintas Kota Semarang dengan metode AR, MA dan ARMA [3]. Pinata, Sukarsa dan Rusjyanthi juga melakukan penelitian dengan mengaplikasikan model XGBoost pada Python untuk meramalkan kecelakaan di Bali [4]. Beberapa metode dari peramalan jumlah kasus kecelakaan yang telah dilakukan membutuhkan sampel yang cukup seperti model AR, MA dan ARMA. Oleh karena itu, metode-metode di atas kurang sesuai untuk data dengan sampel yang relatif kecil sehingga dibutuhkan metode lain untuk meramalkan jumlah kasus kecelakaan lalu lintas.

Double exponential smoothing (DES) merupakan metode peramalan yang sering diaplikasikan di berbagai bidang contohnya di bidang ekologi dengan peramalan emisi gas rumah kaca[5], di bidang ekonomi dengan peramalan jumlah pengangguran [6] dan peramalan jumlah produk yang terjual [7]. Metode tersebut memiliki kelebihan dalam mengakomodasi suatu deret waktu yang memiliki komponen tren. Selain itu, DES tidak memerlukan sampel besar untuk menghasilkan suatu ramalan. Model deret waktu lain yang juga dapat mengakomodasi komponen tren dan tidak memerlukan sampel besar untuk menghasilkan ramalan yang akurat adalah *linear trend analysis* (LTA). Penelitian yang menggunakan *linear trend analysis* sebagai metode peramalan adalah peramalan jumlah vaksin untuk balita pada Dinas Kesehatan Kabupaten Toba [8] dan peramalan jumlah kriminalitas di Kabupaten Sumenep[9].

Double exponential smoothing dan *linear trend analysis* yang memiliki karakteristik relatif sama karena mampu mengakomodasi komponen tren jarang dilakukan asesmen untuk melihat model mana yang memiliki akurasi yang tinggi dalam suatu kasus. Oleh karena, dalam penelitian ini bertujuan untuk membandingkan *linear trend analysis* dan double exponential smoothing untuk meramalkan jumlah kasus kecelakaan lalu lintas di Kota Balikpapan.

## METODOLOGI

Penelitian ini membandingkan dua metode peramalan untuk mendapatkan peramalan terbaik. Metode yang digunakan dalam penelitian ini adalah *linear trend analysis* dan *double exponential smoothing*. Sementara, ukuran akurasi yang aplikasikan untuk mengevaluasi kebaikan model menggunakan *root mean square error* (RMSE) dan *symmetric mean absolute percentage error* (SMAPE).

### *Linear Trend Analysis*

*Linear trend analysis* digunakan untuk memprediksi suatu deret waktu dengan mengasumsikan adanya hubungan linier suatu deret waktu dengan variabel tren yang merupakan urutan suatu data. Model *linear trend analysis* dapat ditulis dalam fungsi matematika sebagai Persamaan (1)[10].

$$\hat{Z}_t = \beta_0 + \beta_1 t \quad (1)$$

dimana  $\hat{Z}_t$  adalah hasil prediksi variabel Z pada periode t,  $\beta_0$  adalah konstanta,  $\beta_1$  adalah perubahan rata-rata dari satu periode ke periode berikutnya dan t adalah variabel satuan waktu.

### *Double Exponential Smoothing*

*Double exponential smoothing* merupakan salah satu metode perluasan *exponential smoothing* yang dapat mengakomodasi komponen tren. Misal  $\alpha$  merupakan parameter untuk level dan  $\gamma$  merupakan parameter untuk tren maka model *double exponential smoothing* dinotasikan sebagai Persamaan (2) [10].

$$\begin{aligned} L_t &= \alpha Z_t + (1 - \alpha)[L_{t-1} + T_{t-1}] \\ T_t &= \gamma[L_t - L_{t-1}] + (1 - \gamma)T_{t-1} \\ \hat{Z}_t &= L_{t-1} + T_{t-1} \end{aligned} \quad (2)$$

di mana  $\hat{Z}_t$  adalah hasil prediksi variabel Z pada periode t,  $L_t$  adalah level pada periode t, dan  $T_t$  adalah tren pada periode t.

## Evaluasi Model

Evaluasi kinerja peramalan dari *linear trend analysis* dan *double exponential smoothing* menggunakan *root mean square error* (RMSE) sebagai Persamaan (3) [10].

$$RMSE = \sqrt{\frac{\sum_{t=1}^n (Z_t - \hat{Z}_t)^2}{n}} \quad (3)$$

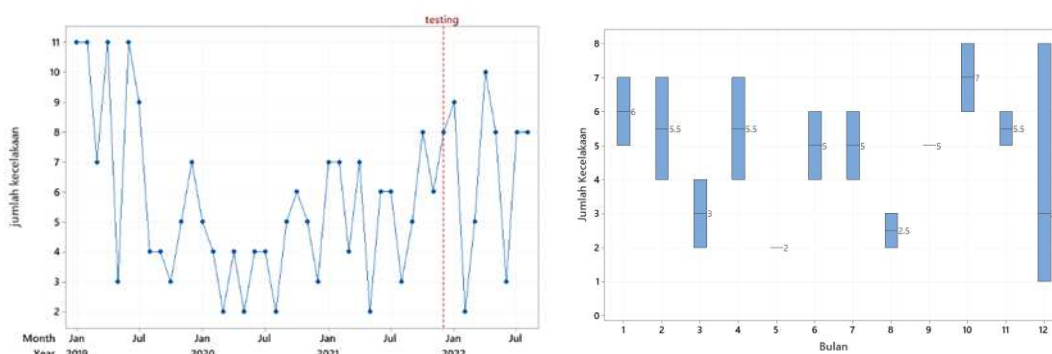
di mana  $\hat{Z}_t$  adalah prediksi variabel Z pada periode t dan  $Z_t$  adalah nilai aktual variabel Z pada periode t.

## Dataset

Dataset yang digunakan dalam penelitian ini merupakan jumlah kasus kecelakaan bulanan Kota Balikpapan yang direkam oleh Kepolisian dari Januari 2019 sampai dengan Agustus 2022. Dataset dibagi menjadi dua bagian, yaitu data training dan data testing. Data training merupakan data yang digunakan untuk pemodelan sementara data testing merupakan data yang digunakan untuk mengevaluasi kebaikan model. Data training dimulai dari Januari 2019 sampai dengan Desember 2021 dan sisanya digunakan untuk data testing sehingga dapat dikatakan proporsi data training sekitar 82% dan data testing sekitar 12%.

## HASIL DAN PEMBAHASAN

Jumlah kasus kecelakaan di Kota Balikpapan menunjukkan adanya fluktuasi dari tahun ke tahun. Sebelum adanya kasus pertama COVID 19 di Indonesia (Maret 2020), jumlah kasus kecelakaan di Kota Balikpapan cenderung tinggi dengan rata-rata sekitar tujuh kasus per bulan (Gambar 1a). Sementara pada puncak COVID 19 pertama yang terjadi di pertengahan tahun 2020, jumlah kasus kecelakaan bulanan Kota Balikpapan menurun hingga sekitar 2-4 kasus per bulan. Hal ini dikarenakan adanya pembatasan sosial untuk menekan kasus COVID 19. Selanjutnya, dengan semakin longgarnya pembatasan sosial dan peralihan ke *new normal*, jumlah kasus kecelakaan di Kota Balikpapan pun cenderung memiliki tren kenaikan yang cukup berfluktuatif tiap bulannya. Oleh karena itu, dapat dikatakan jumlah kasus kecelakaan di Kota Balikpapan menunjukkan adanya pola tren pada deret waktu. Selanjutnya, untuk melihat adanya komponen musiman pada jumlah kasus kecelakaan Kota Balikpapan, penulis memvisualisasikan Box Plot berdasarkan masing-masing bulan dari 2019 sampai dengan 2022. Gambar 1b menunjukkan median jumlah kasus kecelakaan untuk masing-masing bulan sekitar di angka 5 dan 6. Sementara itu, median terendah yaitu 2 yang terjadi di Mei dan median tertinggi yaitu 7 terjadi di Oktober. Berdasarkan Gambar 1a dan Gambar 1b maka dapat disimpulkan bahwa jumlah kasus kecelakaan Kota Balikpapan tidak menunjukkan adanya komponen musiman dan hanya memiliki komponen tren. Oleh karena itu, pemodelan menggunakan *linear trend analysis* dan *double exponential smoothing* yang mengakomodasi komponen tren dapat dipakai untuk pemodelan jumlah kasus kecelakaan di Kota Balikpapan.

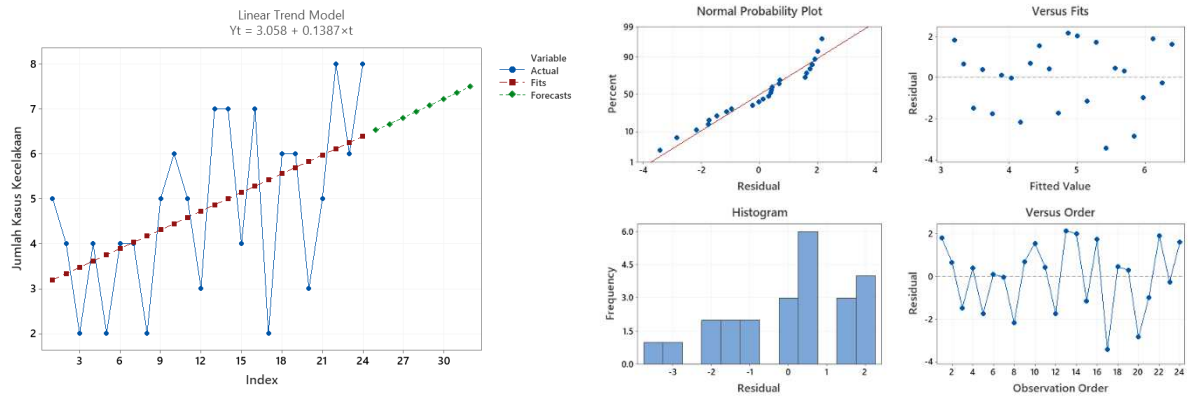


**Gambar 1.** (a) Time Series Plot dan (b) Box Plot Jumlah Kasus Kecelakaan Kota Balikpapan

Pemodelan *linear trend analysis* (LTA) menggunakan prosedur yang sama seperti regresi sederhana. Pada pemodelan LTA, variabel respon yang digunakan adalah jumlah kasus kecelakaan Kota Balikpapan dan variabel prediktor yang digunakan adalah  $t$  atau variabel yang berisi urutan data. Dari pemodelan LTA yang telah dilakukan seperti pada Gambar 2a didapatkan model seperti pada Persamaan (4):

$$\hat{Z}_t = 3,058 + 0,1387t \quad (4)$$

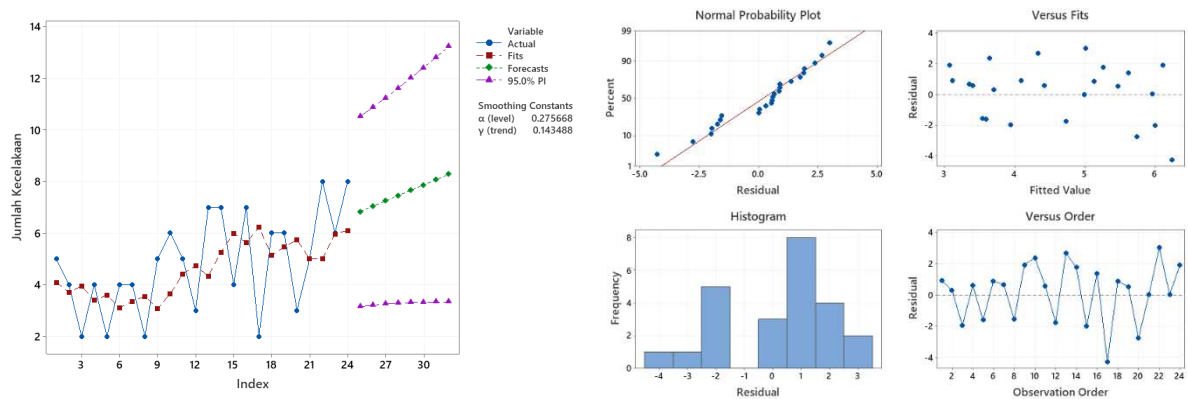
Setelah didapatkan model *linear trend analysis* (LTA) maka dapat dilakukan evaluasi pada residual. Pada Gambar 2b *Normal Probability Plot* menunjukkan bahwa residual sudah memenuhi asumsi distribusi normal karena titik-titik residual mengikuti garis distribusi normal. Selain itu, Gambar 2b juga menunjukkan bahwa residual juga sudah memenuhi asumsi identik dan independen karena residual sudah homogen dan tidak memiliki autokorelasi. Selanjutnya, dapat dilakukan peramalan delapan bulan ke depan yang kemudian dihitung nilai RMSE dari peramalan menggunakan model LTA.



**Gambar 2.** (a) *Linear Trend Analysis* dan (b) *Residual Plot Linear Trend Analysis*

Prosedur dalam pemodelan *double exponential smoothing* (DES) yang paling krusial adalah pemilihan parameter  $\alpha$  dan  $\gamma$  yang masing-masing untuk pembobot level dan pembobot tren. Dalam penelitian ini, parameter  $\alpha$  dan  $\gamma$  didapatkan dari memodelkan data ke model ARIMA (0, 2, 2) untuk meminimumkan SSE kemudian komponen tren dan level kemudian diinisiasi dengan *backcasting*. Pemodelan DES yang digunakan dalam penelitian ini dinotasikan dalam Persamaan (5).

$$\begin{aligned} L_t &= 0,2757Z_t + (1-0,2757)[L_{t-1} + T_{t-1}] \\ T_t &= 0,1435[L_t - L_{t-1}] + (1-0,1435)T_{t-1} \\ \hat{Z}_t &= L_{t-1} + T_{t-1} \end{aligned} \quad (5)$$



**Gambar 3.** (a) *Double Exponential Smoothing* dan (b) *Residual Plot Double Exponential Smoothing*

Setelah didapatkan model *double exponential smoothing* (DES) maka dapat dilakukan evaluasi pada residual. Pada Gambar 3b *Normal Probability Plot* menunjukkan bahwa residual sudah memenuhi asumsi distribusi normal karena titik-titik residual mengikuti garis distribusi normal. Selain itu, Gambar 3b juga menunjukkan bahwa residual juga sudah memenuhi asumsi identik dan independen karena residual sudah homogen dan tidak memiliki autokorelasi. Selanjutnya, dapat dilakukan peramalan delapan bulan ke depan yang kemudian dihitung nilai RMSE dari peramalan menggunakan model DES.

Perbandingan kebaikan model *linear trend analysis* (LTS) dan *double exponential smoothing* (DES) berdasarkan *root mean square error* (RMSE) dapat dirangkum seperti pada Tabel 1.

Berdasarkan Tabel 1, model memiliki nilai RMSE terkecil adalah *linear trend analysis* (LTA) dengan nilai RMSE sebesar 2.73. Oleh karena itu, model yang memiliki kinerja terbaik dalam meramalkan jumlah kasus kecelakaan Kota Balikpapan dalam penelitian ini adalah *linear trend analysis* (LTA). Hasil peramalan yang dihasilkan model LTA memiliki pola tren deterministik yang linier (Gambar 2a). Hal ini terjadi karena setiap bertambahnya periode

peramalan maka nilai prediksi dari jumlah kasus kecelakaan lalu lintas di Kota Balikpapan akan meningkat sebesar 0,1387.

**Tabel 1.** Rangkuman Model *Linear Trend Analysis* dan *Double Exponential Smoothing*

Model	RMSE
<i>Linier Trend Analysis</i>	2.73
<i>Double Exponential Smoothing</i>	2.86

Penelitian ini memiliki batasan pada jumlah data training sebanyak 36 pengamatan sementara untuk data testing sebanyak delapan pengamatan. Selain itu, batasan juga dilakukan pada pemilihan model untuk *trend analysis* yang hanya berupa tren linier dan pemilihan parameter  $\alpha$  dan  $\gamma$  pada *double exponential smoothing* yang menggunakan optimasi SSE dari ARIMA (0, 2, 2). Penelitian lanjutan mungkin dapat ditambahkan efek peraturan pembatasan sosial masyarakat akibat COVID 19. Serta, dapat juga diinvestigasi beberapa metode dari *machine learning* yang sesuai untuk peramalan dengan sampel kecil.

## KESIMPULAN

Peramalan jumlah kasus kecelakaan sangat krusial karena kecelakaan menimbulkan berbagai kerugian. Pada tahapan identifikasi data deret waktu, jumlah kecelakaan lalu lintas Kota Balikpapan mengikuti pola tren dan tidak musiman. Berdasarkan pola dalam deret waktu maka metode *linear trend analysis* dan *double exponential smoothing* sesuai untuk digunakan dalam meramalkan jumlah kasus kecelakaan Kota Balikpapan. Model terbaik untuk meramalkan jumlah kasus kecelakaan Kota Balikpapan adalah *linear trend analysis* dengan nilai RMSE sebesar 2.73. Hal ini mengindikasikan bahwa jumlah kecelakaan lalu lintas Kota Balikpapan ke depannya akan mengalami kenaikan secara linier sebesar 0,1387 dari satu periode ke periode berikutnya. Tentunya hal ini harus mendapatkan perhatian lebih dari pemerintah dan kepolisian agar angka jumlah kecelakaan lalu lintas ini menurun.

## REFERENSI

- [1] WHO. Global Status Report on Road Safety. France. (2018)
- [2] E. Haryono, A. Widodo, S. Abusini, Kajian model Automatic Clustering-Fuzzy Time Series-Markov Chain dalam memprediksi data historis jumlah kecelakaan lalu lintas di kota Malang. J. Sains Dasar **2** (1), 63 – 71, (2013)
- [3] I.R. Maharsi, M. A. Mukid, Y. Wilandari, Peramalan Jumlah Kecelakaan di Kota Semarang Tahun 2017 Menggunakan Metode Runtun Waktu (Studi Kasus : Data Jumlah Kecelakaan Lalu Lintas di Kota Semarang Periode Januari 2012 – Desember 2016). Jurnal Gaussian **6** (3), 355-364 (2017)
- [4] N. N. P. Pinata, I. M. Sukarsa, N. K. D. Rusjyanthi, Prediksi Kecelakaan Lalu Lintas di Bali dengan XGBoost pada Python. Jurnal Ilmiah Merpati, **8** (3), 188-196, (2020).
- [5] R. Alhindawi, Y.A. Nahleh, A. Kumar, N. Shiwakoti, Projection of Greenhouse Gas Emissions for the Road Transport Sector Based on Multivariate Regression and the Double Exponential Smoothing Model. Sustainability, **12**(21), 9152, (2020)
- [6] H. Syafwan I, M. Syafwan, E. Syafwan, A. F. Hadi, P. Putri, Forecasting Unemployment in North Sumatra Using Double Exponential Smoothing Method. Annual Conference on Science and Technology Research (ACOSTER) 2020. Journal of Physics: Conference Series 1783, 012008, (2021).
- [7] F. Sidqi, I. D. Sumitra. Forecasting Product Selling Using Single Exponential Smoothing and Double Exponential Smoothing Methods. IOP Conf. Series: Materials Science and Engineering 662, 032031, (2019).
- [8] F. S. Aritonang, I. M. Sarkis, A. Situmorang, Peramalan Penyediaan Jumlah Vaksin Untuk Balita Dengan Metode Trend Projection di Dinas Kesehatan Kabupaten Toba, METHOSISFO: Jurnal Ilmiah Sistem Informasi, 2(1), 39-45, (2022).
- [9] E. Rahmanita, L. Fauziah, W. Agustiono, Forecasting Crime: Application of Trend Projection Method in Sumenep District, International Joint Conference on Science and Technology, 78-83, (2019).
- [10] W. W. S. Wei, Time Series Analysis Univariate and Multivariate Methods 2<sup>nd</sup>. Pearson Education (2006).