

MODEL GABUNGAN (ANSAMBEL) SARIMA DAN JARINGAN SARAF TIRUAN UNTUK PERAMALAN BEBAN LISTRIK

Combination Model of SARIMA and Artificial Neural Network for Electrical Load Forecasting

Mega Silfiani^{1*}

¹Program Studi Statistika, JMTI, Institut Teknologi Kalimantan
Kampus ITK Karang Joang, Jalan Soekarno Hatta KM.15, Balikpapan, Kalimantan Timur

E-mail Corresponding Author: megasilfiani@lecturer.itk.ac.id

Abstrak: Penelitian ini bertujuan untuk menyelidiki efisiensi penggunaan *artificial neural network* (ANN) dalam model gabungan *seasonal autoregressive integrated moving average* (SARIMA) untuk meramalkan beban listrik. Model gabungan SARIMA terdiri dari beberapa model SARIMA. Selanjutnya, hasil ramalan dari beberapa SARIMA tersebut menjadi *input* dari ANN. Data penelitian ini mencakup beban listrik bulanan yang berkaitan dengan rumah tangga, bisnis, industri, dan masyarakat, mulai dari Januari 2016 hingga Desember 2020. Akurasi penelitian menggunakan *symmentrice* Hasil penelitian menunjukkan bahwa rata-rata SMAPE (%) model gabungan SARIMA dan ANN menunjukkan kinerja peramalan yang unggul dibandingkan dengan model individu terbaik. Potensi penelitian di masa depan dapat fokus pada eksplorasi beragam metodologi untuk menciptakan dan menggabungkan anggota dari model gabungan.

Kata Kunci: ANN, Beban Listrik, Model Gabungan, SARIMA, SMAPE

Abstract: This study aims to investigate the efficacy of employing artificial neural networks in conjunction with a seasonal autoregressive integrated moving average (SARIMA) ensemble for forecasting electrical load. The SARIMA ensemble comprises members generated by varying autoregressive orders or moving averages. Subsequently, these SARIMA ensemble members are integrated using artificial neural networks. The datasets encompass monthly electrical load data pertaining to households, businesses, industries, and the public, spanning from January 2016 to December 2020. The findings demonstrate that across various categories, SARIMA ensemble-based artificial neural networks demonstrated superior predictive performance compared to alternative models. Future research endeavors should focus on exploring diverse methodologies for both creating and amalgamating ensemble members.

Keywords: ANN, Combination Model, Electrical Load, SARIMA, SMAPE

1. PENDAHULUAN

Peramalan beban listrik memiliki peran penting dalam mengoptimalkan perencanaan sistem tenaga, operasi dan manajemen pemeliharaan [1]. Selain itu, peramalan beban listrik juga dapat membantu pemerintah dalam mengantisipasi kebutuhan beban listrik di masa depan, sehingga memberikan informasi untuk alokasi anggaran subsidi [2]. Akan tetapi peramalan beban listrik yang akurat menghadapi tantangan besar akibat beberapa faktor cuaca seperti suhu yang dapat mempengaruhi permintaan beban listrik [3]. Penelitian peramalan beban listrik telah dilakukan oleh [1]–[8].

Dataset beban listrik memiliki pola tren dan musiman [2]. Hal ini akibat dari permintaan beban listrik erat kaitannya dengan variasi musiman dan cuaca, di mana rumah tangga menunjukkan permintaan yang lebih tinggi pada musim dingin dibandingkan dengan musim panas [2]. Oleh karena itu, peramalan beban listrik seharusnya menggunakan metode yang dapat mengakomodasi adanya pola tren dan musiman. *Seasonal Autoregressive Integrated Moving Average* (SARIMA) merupakan salah satu pilihan tepat untuk meramalkan pola tren dan musiman yang terdapat dalam beban listrik [2][9]. SARIMA merupakan model *time series* yang dapat mengatasi deret waktu yang bersifat stasioner maupun tidak stasioner, menangkap pola baik non-musiman maupun musiman

[10]. Selain penggunaan SARIMA telah dilakukan di bidang peramalan beban listrik [1], [2], [9], [10], SARIMA juga diterapkan dalam berbagai bidang, termasuk peramalan permintaan pariwisata.

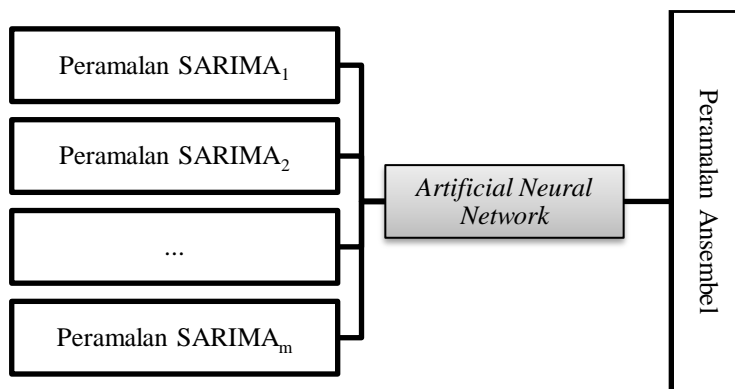
Model penggabungan beberapa metode peramalan telah menunjukkan efektivitasnya dalam meningkatkan akurasi peramalan dari model tunggal [2], [11]–[13]. Model tersebut sering disebut pendekatan ansembel [2], [13]. Penelitian peramalan beban listrik ini mengusulkan penggunaan model gabungan SARIMA dan *artificial neural network* (ANN). Hal ini didasarkan pada penelitian-penelitian dari hibrida SARIMA dan ANN yang dapat mengakomodasi pola linier dan nonlinier dalam data. Perbedaan mendasar pendekatan ansembel dengan hibrida SARIMA-ANN adalah pendekatan ansembel menggunakan hasil prediksi dari beberapa SARIMA sebagai *input* pada ANN, sementara pendekatan hibrida menggunakan residual dari SARIMA sebagai *input* pada ANN.

Berdasarkan uraian sebelumnya maka penelitian bertujuan untuk mengevaluasi akurasi peramalan dari penggabungan SARIMA dengan ANN di mana *symmetric mean absolute percentage error* (SMAPE) digunakan untuk mengevaluasi akurasi peramalan beban listrik. Dengan ide penggabungan model linier dan nonlinier diharapkan pendekatan ansembel SARIMA dan ANN memberikan kinerja yang lebih baik dibandingkan model tunggal dari SARIMA. Artikel ini terdiri dari beberapa sesi yang dapat dijelaskan sebagai berikut. Sesi 2 metodologi yang berisi pendekatan model gabungan, SARIMA, ANN, dan langkah penelitian. Sesi 3 berisi hasil dan pembahasan dari pemodelan SARIMA, pemodelan model gabungan SARIMA dan ANN, kemudian perbandingannya. Sesi terakhir adalah kesimpulan dan potensi penelitian lanjutan.

2. METODOLOGI

2.1. Pendekatan Model Gabungan (Ansembel)

Penggabungan beberapa nilai peramalan dari beberapa model dikenal sebagai metode ansambel [2], [9]. Beberapa penelitian membuktikan bahwa metode gabungan dapat menghasilkan nilai peramalan yang lebih baik daripada memilih model individual [2], [11], [12]. Namun, penyelidikan tentang kinerja teknik gabungan untuk *seasonal autoregressive integrated moving average* (SARIMA) dengan *artificial neural network* kurang mendapat perhatian. Oleh karena itu, makalah ini mengusulkan penyelidikan dalam ansambel SARIMA dengan *artificial neural network* (ANN). Prosedurnya adalah sebagai berikut: Ansambel SARIMA dibangun dengan menggabungkan beberapa nilai peramalan dari beberapa model SARIMA individu. Sementara itu, ANN digunakan untuk menggabungkan anggota ansambel dengan cara berapa model SARIMA individu digunakan sebagai *input* ANN dan targetnya adalah data aktual dari beban listrik. Prosedur pendekatan model gabungan SARIMA dan ANN dapat dijelaskan berdasarkan Gambar 1.



Gambar 1. Arsitektur Model Gabungan (Ansembel) dari SARIMA dan ANN

2.2. Seasonal Autoregressive Integrated Moving Average (SARIMA)

SARIMA adalah perluasan dari model ARIMA yang dapat mengakomodasi pola tren dan musiman dalam data deret waktu [2], [14]. SARIMA memiliki orde $(p, d, q) (P, D, Q)^s$ di mana $p, d,$ dan q mewakili komponen non-musiman, $P, D,$ dan Q mewakili komponen musiman, dan s mewakili frekuensi pola musiman. Persamaan SARIMA dapat dinotasikan sebagai (1) [2], [15], [16].

$$Z_t = \frac{\theta_q(B)\Theta_Q(B^s)N_t}{\Phi_p(B^s)\phi_p(B)(1-B)^d(1-B^s)^D} \tag{1}$$

di mana $\Phi_p(B^s) = 1 - \Phi_1 B^s - \Phi_2 B^{2s} - \dots - \Phi_p B^{ps}$ adalah komponen musiman *autoregressive*, $\phi_p(B) = 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p$ adalah komponen *autoregressive nonmusiman*, $(1-B)^d$ adalah *nonmusiman differencing*, $(1-B^s)^D$ adalah musiman *differencing*, $\theta_q(B) = 1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q$ adalah *nonmusiman moving average*, Z_t adalah beban listrik aktual pada periode t , $\Theta_Q(B^s) = 1 - \Theta_1 B^s - \Theta_2 B^{2s} - \dots - \Theta_Q B^{Qs}$ adalah komponen *moving average musiman*.

Pada pemodelan SARIMA akan dilakukan penambahan pemodelan pencilan jika asumsi residual berdistribusi normal tidak memenuhi. Hal ini dilakukan karena adanya dugaan pencilan menyebabkan ketidaknormalan dalam residual sehingga memerlukan penanganan tambahan. Penelitian ini menggunakan dua jenis pencilan dalam pemodelannya, yaitu *Additive Outlier (AO)* dan *Level Shifts (LS)*. Model SARIMA yang menggabungkan pencilan yang dikembangkan oleh program SAS dan berikut persamaannya umumnya [2]:

$$Z_t = \sum_{j=1}^k \varpi_j v_j(B) I_j^{(T_j)} + \frac{\theta_q(B)\Theta_Q(B^s)}{\Phi_p(B^s)\phi_p(B)(1-B)^d(1-B^s)^D} N_t \tag{2}$$

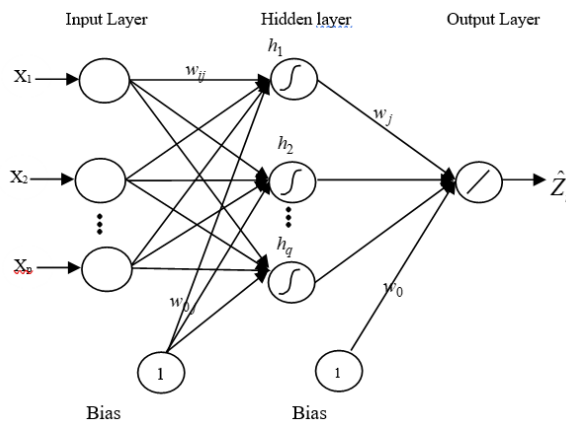
di mana $I_j^{(T_j)}$ adalah pencilan pada periode T_j , dan $v_j(B) = 1$ untuk AO, $v_j(B) = \frac{1}{(1-B)}$ untuk LS. Jika model SARIMA memenuhi seluruh asumsi maka dapat dilanjutkan ke tahap peramalan.

2.3. Artificial Neural Network (ANN)

ANN yang digunakan dalam penelitian ini adalah *feed-forward neural network (FFNN)*. FFNN biasa dikenal sebagai *multilayer perceptron (MLP)* [17]. FFNN memiliki kelebihan dibandingkan arsitektur yang lain karena cukup sederhana. Representasi matematis dari hubungan *input-output* dalam FFNN adalah sebagai berikut:

$$\hat{Z}_t = w_0 + \sum_{j=1}^q w_j h_j \text{ dan } h_j = \frac{1}{1 + \exp\left(-\left(w_{0j} + \sum_{i=1}^p w_{ij} X_i\right)\right)} \tag{3}$$

di mana $w_{ij} (i=1, 2, \dots, p; j=1, 2, \dots, q)$ dan $w_j (j=1, 2, \dots, q)$ adalah parameter dari FFNN yang biasa disebut sebagai pembobot, p adalah banyaknya neuron pada *input layer* dan q adalah banyaknya neuron pada *hidden layer*. Penelitian ini menggunakan FFNN dengan satu *hidden layer* dengan fungsi aktivasi logistik sigmoid, dan fungsi aktivasi linier pada *output layer*. Selanjutnya, algoritma *backpropagation* digunakan untuk memperkirakan bobot dan bias. Gambar 2 menampilkan arsitektur FFNN.



Gambar 2. Arsitektur Feed-Forward Neural Network (FFNN)

2.4. Langkah Penelitian

Langkah penelitian dalam paper ini adalah sebagai berikut:

- i. Data dari masing-masing kategori beban listrik dibagi menjadi dua bagian, data training dan data testing. Data training digunakan untuk pemodelan dan data testing digunakan untuk evaluasi peramalan. Data training terdiri dari Januari 2016 sampai dengan Desember 2019. Sementara data testing menggunakan Januari 2020 sampai dengan Desember 2020
- ii. Mengaplikasikan data training ke model SARIMA. Prosedur dalam pemodelan SARIMA adalah sebagai berikut.
 1. Identifikasi model: Langkah ini bertujuan untuk mengidentifikasi stasioneritas mean dan varians. Uji stasionaritas varians mengimplementasikan Transformasi Box-Cox, sedangkan uji stasionaritas rata-rata menggunakan ilustrasi plot ACF. Jika data tidak stasioner dalam hal varians, transformasi Box-Cox harus dilakukan; Jika gagal memenuhi asumsi stasioner dalam hal rata-rata, data harus di-differencing. Setelah data telah stasioner dalam hal varians dan mean, pola grafik fungsi autokorelasi (ACF) dan fungsi autokorelasi parsial (PACF) dapat digunakan untuk memperkirakan model SARIMA (p, d, q) (P, D, Q).
 2. Estimasi Model: Penelitian ini menggunakan *conditional least square* untuk mengestimasi parameter dari SARIMA. Setelah mendapatkan estimasi parameter, statistik uji-t digunakan untuk melakukan uji signifikansi parameter SARIMA.
 3. Pemeriksaan diagnosa: Pemeriksaan diagnostik dilakukan untuk mengevaluasi asumsi *white noise* dan distribusi normal dalam residual SARIMA. Evaluasi residual *white noise* menggunakan uji Ljung Box sementara distribusi normal dilakukan menggunakan uji distribusi Kolmogorov Smirnov.
 4. Peramalan
- iii. Membuat model prakiraan gabungan dengan memodelkan ANN menggunakan prediksi data pelatihan di masing-masing SARIMA pada langkah 2 dan langkah 3 sebagai *input* kemudian meramalkan pengujian data
- iv. Setelah mendapatkan peramalan data pengujian untuk setiap langkah 2, langkah 3 dan langkah 4, akurasi hasil perkiraan masing-masing model dapat dievaluasi. Evaluasi ini menggunakan SMAPE seperti yang ditunjukkan pada (4).

$$SMAPE = \frac{100\%}{n} \sum_{t=1}^n \frac{|Z_t - \hat{Z}_t|}{\left| \frac{Z_t + \hat{Z}_t}{2} \right|} \quad (4)$$

dimana Z_t adalah beban listrik aktual pada periode t and \hat{Z}_t adalah prediksi beban listrik pada periode t.

3. HASIL DAN PEMBAHASAN

Dataset yang digunakan dalam penelitian ini adalah data beban listrik bulanan di Kota Jember dari beberapa kategori pengguna seperti rumah tangga, bisnis, industri dan publik.

Untuk membangun SARIMA individu, para peneliti menerapkan pendekatan ARIMA Box Jenkins. Dalam identifikasi model, stasioner data deret waktu diidentifikasi, baik mean maupun varians, dan musiman dataset deret waktu. Dalam estimasi model, urutan SARIMA (p, d, q) (P, D, Q) diperkirakan menggunakan penaksir kuadrat terkecil bersyarat dan urutan awal SARIMA didasarkan pada pola dan plot ACF dan PACF. Akhirnya, dalam pemeriksaan diagnostik, *white noise* dan distribusi normalitas residu diperiksa dengan penerapan uji Ljung-Box dan uji Kolmogorov-Smirnov. Uji Ljung-Box digunakan untuk mengevaluasi *white noise* residual, sedangkan uji Kolmogorov-Smirnov digunakan untuk mengevaluasi distribusi normalitas residual yang digunakan setiap tes $\alpha=5\%$.

Tabel 1. Model SARIMA pada Peramalan Beban Listrik

Kategori	Model SARIMA	SMAPE (%)
Rumah Tangga	SARIMA (0,1,0) (1,0,0) ¹²	3,530
Bisnis	SARIMA (0,1,0) (1,0,0) ¹²	16,142
Industri	SARIMAX (0,1,1) (1,0,0) ¹²	7,639
Publik	SARIMAX (0,1,1) (0,0,1) ¹²	2,918

Setelah mendapatkan model SARIMA, dimungkinkan untuk menggabungkan nilai peramalan dengan menerapkan ANN. Nilai prediksi beberapa SARIMA untuk data pelatihan akan menjadi *input* model ANN sementara nilai aktual dalam data pelatihan akan menjadi target ANN. Sebelum membuat model ANN, *input* dan target akan menjadi standardisasi. Penelitian ini menerapkan *backproagation* untuk melatih *neural network*. Arsitektur *neural network* adalah satu *hidden layer* dan jumlah neuron di *hidden layer* digunakan oleh 1-20 neuron. Jumlah optimal neuron di *hidden layer* dievaluasi oleh Mean Square Error (MSE) dari data pelatihan. Selain itu, kami mengadaptasi fungsi aktivasi logistik sigmoid untuk neuron di *hidden layer* dan fungsi aktivasi linier untuk neuron di *output layer*. Model gabungan SARIMA melalui ANN untuk setiap kategori mengikuti Tabel 2.

Tabel 2. Model Gabungan SARIMA dan ANN pada Peramalan Beban Listrik

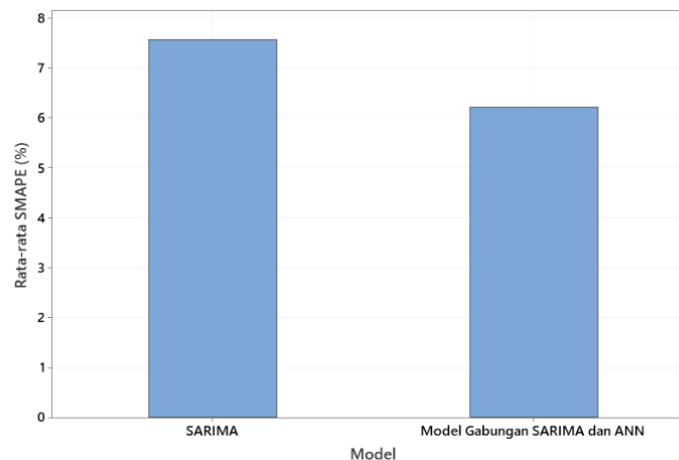
Kategori	Banyaknya Anggota Ansembel	SMAPE (%)
Rumah Tangga	6	5,827
Bisnis	6	8,361
Industri	6	8,187
Publik	6	2,411

SMAPE digunakan untuk memilih model terbaik untuk beban listrik. Tabel 3 menunjukkan bahwa kinerja SARIMA dan model gabungan SARIMA dan ANN sama. Hal ini dapat dilihat dari model terbaik dari masing-masing kategori. Model SARIMA individu memberikan kinerja peramalan yang lebih baik pada peramalan beban listrik kategori rumah tangga dan industri. Sementara model gabungan SARIMA dengan ANN memberikan kinerja peramalan beban listrik yang lebih baik pada kategori bisnis dan publik. Walaupun demikian, jika kita menghitung rata-rata SMAPE (%) dari keseluruhan kategori berdasarkan pemodelannya memberikan hasil yang berbeda.

Tabel 3. Perbandingan SARIMA dan Model Gabungan SARIMA dan ANN

Kategori	Model	SMAPE (%)
Rumah Tangga	SARIMA	3,530
	Gabungan SARIMA-ANN	5,827
Bisnis	SARIMA	16,142
	Gabungan SARIMA-ANN	8,367
Industri	SARIMA	7,639
	Gabungan SARIMA-ANN	8,187
Publik	SARIMA	2,918
	Gabungan SARIMA-ANN	2,411

Gambar 3 menunjukkan adanya perbedaan rata-rata SMAPE (%) dari keseluruhan kategori beban listrik. Rata-rata SMAPE (%) yang dihasilkan model gabungan SARIMA dan ANN memiliki nilai yang lebih kecil dari rata-rata SMAPE (%) model SARIMA individu. Dengan kata lain, model gabungan SARIMA dan ANN secara keseluruhan memberikan kinerja peramalan yang lebih baik dari SARIMA individu.



Gambar 3. Perbandingan Rata-Rata SMAPE (%) SARIMA dengan Model Gabungan SARIMA dan ANN

4. KESIMPULAN

Studi ini menginvestigasi penggunaan efisien ANN dalam model gabungan SARIMA untuk melakukan prediksi beban listrik. Model gabungan SARIMA terdiri dari beberapa model SARIMA yang berbeda. Hasil prediksi dari berbagai SARIMA ini kemudian dijadikan *input* untuk ANN. Temuan dari penelitian ini menunjukkan bahwa di berbagai kategori, model gabungan SARIMA dan ANN memiliki kinerja prediksi dari model individu. Sebagai arah penelitian mendatang, disarankan untuk mengeksplorasi berbagai metodologi dalam menciptakan dan menggabungkan anggota dari model gabungan.

DAFTAR PUSTAKA

- [1] E. Chodakowska, J. Nazarko, and Ł. Nazarko, "Arima models in electrical load forecasting and their robustness to noise," *Energies*, vol. 14, no. 23, 2021, doi: 10.3390/en14237952.
- [2] M. Silfiani, H. Aprillia, and Y. Fitriani, "Comparing Various Combined Techniques at Seasonal Autoregressive Integrated Moving Average (SARIMA) for Electrical Load Forecasting," in *2023 International Seminar on Intelligent Technology and Its Applications (ISITIA)*, Jul. 2023, pp. 376–381. doi: 10.1109/ISITIA59021.2023.10221130.
- [3] K. Chapagain, S. Kittipiyakul, and P. Kulthanavit, "Short-term electricity demand forecasting: Impact analysis of temperature for Thailand," *Energies*, vol. 13, no. 10, pp. 1–29, 2020, doi: 10.3390/en13102498.
- [4] E. Yukseltan, A. Yucekaya, and A. H. Bilge, "Forecasting electricity demand for Turkey: Modeling periodic variations and demand segregation," *Appl. Energy*, vol. 193, pp. 287–296, 2017, doi: 10.1016/j.apenergy.2017.02.054.
- [5] M. Al Mamun and K. Nagasaka, "Artificial neural networks applied to long-term electricity demand forecasting," *Proc. - HIS'04 4th Int. Conf. Hybrid Intell. Syst.*, pp. 204–209, 2005, doi: 10.1109/ichis.2004.27.
- [6] M. H. M. R. S. Dilhani and C. Jeenanunta, "Daily electric load forecasting: Case of Thailand," *7th Int. Conf. Inf. Commun. Technol. Embed. Syst. 2016, IC-ICTES 2016*, vol. 2016, pp. 25–29, 2016, doi: 10.1109/ICTEmSys.2016.7467116.
- [7] Y. J. Chang, S. H. Wang, and H. Y. Sun, "Research of short-term load forecasting algorithm based on wavelet analysis and Radial Basis Function Neural Network," *PEITS 2009 - 2009 2nd Conf. Power Electron. Intell. Transp. Syst.*, vol. 1, pp. 81–84, 2009, doi: 10.1109/PEITS.2009.5406961.
- [8] J. Bedi and D. Toshniwal, "Deep learning framework to forecast electricity demand," *Appl. Energy*, vol. 238, no. January, pp. 1312–1326, 2019, doi: 10.1016/j.apenergy.2019.01.113.
- [9] M. Silfiani, F. N. Hayati, D. Nurlaily, and I. Fitria, "Household Electrical Load Forecasting: A Hybrid of Linear Models and Radial Basis Function Neural Network," 2021. doi: 10.1109/ICAMIMIA54022.2021.9807693.
- [10] A. Kumar Dubey, A. Kumar, V. García-Díaz, A. Kumar Sharma, and K. Kanhaiya, "Study and analysis of SARIMA and LSTM in forecasting time series data," *Sustain. Energy Technol. Assessments*, vol. 47, no. July, p. 101474, 2021, doi: 10.1016/j.seta.2021.101474.
- [11] U. Gunter, "Improving Hotel Room Demand Forecasts for Vienna across Hotel Classes and Forecast Horizons: Single Models and Combination Techniques Based on Encompassing Tests," *Forecasting*, vol. 3, no. 4, pp. 884–919, 2021, doi: 10.3390/forecast3040054.
- [12] U. Gunter, I. Önder, and E. Smeral, "Are Combined Tourism Forecasts Better at Minimizing Forecasting Errors?," *Forecasting*, vol. 2, no. 3, pp. 211–229, Jun. 2020, doi: 10.3390/forecast2030012.
- [13] C. Lu, J. Liang, W. Jiang, J. Teng, and C. Wu, "High-resolution probabilistic load forecasting: A learning ensemble approach," *J. Franklin Inst.*, vol. 360, no. 6, pp. 4272–4296, 2023, doi: 10.1016/j.jfranklin.2023.02.010.

- [14] A. Malki, E. S. Atlam, A. E. Hassanien, A. Ewis, G. Dagneu, and I. Gad, "SARIMA model-based forecasting required number of COVID-19 vaccines globally and empirical analysis of peoples' view towards the vaccines," *Alexandria Eng. J.*, vol. 61, no. 12, pp. 12091–12110, 2022, doi: 10.1016/j.aej.2022.05.051.
- [15] C. A. Mendoza, J. G. Brida, and N. Garrido, "The impact of earthquakes on Chile's international tourism demand," *J. Policy Res. Tour. Leis. Events*, vol. 4, no. 1, pp. 48–60, 2012, doi: 10.1080/19407963.2011.636923.
- [16] Y. W. Chang and M. Y. Liao, "A seasonal ARIMA model of tourism forecasting: The case of Taiwan," *Asia Pacific J. Tour. Res.*, vol. 15, no. 2, pp. 215–221, 2010, doi: 10.1080/10941661003630001.
- [17] N. An, W. Zhao, J. Wang, D. Shang, and E. Zhao, "Using multi-output feedforward neural network with empirical mode decomposition based signal filtering for electricity demand forecasting," *Energy*, vol. 49, no. 1, pp. 279–288, 2013, doi: 10.1016/j.energy.2012.10.035.

