residuals follow a distribution the data points should align closely with a reference line. The alignment of points in Figure 8, along the line indicates that model 2's residuals adhere closely to a distribution suggesting that the model is well tuned, and its predictions are trustworthy.

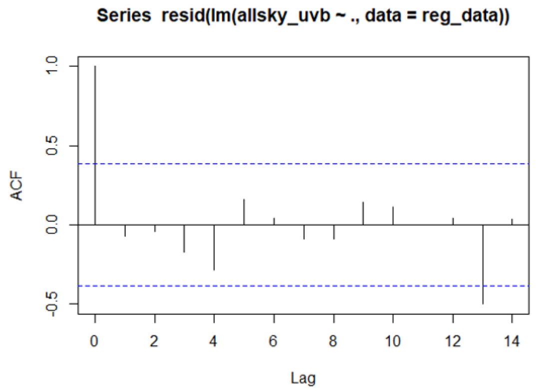

Figure 9: Series residual plot of model 2.

Figure 9 is the Autocorrelation Function (ACF) plot, for the residuals from model 2. It illustrates how the series correlates with itself at time lags. In a scenario where the model fits well, we would anticipate the autocorrelations at all lag points to fall within the confidence bounds (indicated by the dashed lines). This indicates no autocorrelation. It also shows that the model's residuals are random. The randomness signifies that the model has effectively captured all information from the data.

Our LASSO regression model's predictive accuracy is quantified through several statistical measures. The RMSE of 0.0386 highlights that our predictions deviate from the actual values by this small margin, indicating a tightly fitted model. The MSE, at 0.00149, reaffirms this, showing a minimal average of squared errors.

Impressively, the model accounts for approximately 81.3% of the variability in the target variable, as suggested by an R-squared of 0.813. Such a high R-squared value reflects the model's robustness in capturing the underlying data patterns.

To round up, these metrics underscore a strong predictive capability, suggesting the model is well-tuned to the nuances of our data.

4 CONCLUSIONS

The aim of the study was achieved through a systematic approach involving LASSO regression to forecast solar irradiation in Tomsk, which is a method particularly well-suited to handle the challenges posed by climatic variability. The study created two distinct models to interpret the complex relationships between meteorological factors and solar power output. Model 1 included specific humidity at 2 meters as an independent variable, while Model 2 excluded it, thus allowing for the analysis of other significant meteorological factors.

The effectiveness of these models was demonstrated by robust statistical metrics: Model 1 showed an R-squared value of 0.843, indicating that it could explain over 84% of the variability in the dependent variable, while Model 2 had an R-squared value of 0.813, accounting for approximately 81.3% of the variability in the target variable. These high R-squared values signify that both models have strong predictive capabilities and can reliably capture the underlying data patterns.

Furthermore, the use of the NASA POWER database provided a comprehensive set of meteorological variables, ensuring that the models had a solid data foundation to work from. The choice of Tomsk, with its distinctive climate, offered a unique case for examining solar irradiation patterns, contributing to the literature on renewable energy forecasting and the operational practices in the field.

In essence, the study succeeded in contributing to predictive modeling by developing interpretable models that effectively address climatic variability and demonstrate strong predictive performance, thus advancing the field of solar power output forecasting.

To further enhance the research presented, one could explore the integration of real-time data feeds to improve model responsiveness, use Explainable Artificial Intelligence (XAI) tools to broaden the interpretability spectrum, and test the scalability of the proposed framework across different regions and forms of renewable energy.

REFERENCES

[1] "Transforming our world: the 2030 Agenda for Sustainable Development," UN Doc. A/RES/70/1, Sept. 25, 2015.

[2] S. Krumdieck, Transition Engineering: Building a Sustainable Future. Taylor & Francis, 2020.

[3] J. Stephenson, B. Barton, G. Carrington, A. Doering, R. Ford, D. Hopkins et al., "The energy cultures framework: exploring the role of norms, practices and material culture in shaping energy behaviour in New Zealand," Energy Research & Social Science, vol. 7, pp. 117-123, 2015, [Online]. Available: https://doi.org/10.1016/j.erss.2015.03.005.

[4] B. Sovacool, M. Burke, L. Baker, C. Kotikalapudi, and H. Wlokas, "New frontiers and conceptual frameworks for energy justice," Energy Policy, vol. 105, pp. 677-691, 2017, [Online]. Available: https://doi.org/10.1016/j.enpol.2017.03.005.

[5] D. McCollum, W. Zhou, C. Bertram, H. Boer, V. Bosetti, S. Busch et al., "Energy investment needs for fulfilling the Paris agreement and achieving the sustainable development goals," Nature Energy, vol. 3, no. 7, pp. 589-599, 2018, [Online]. Available: https://doi.org/10.1038/s41560-018-0179-z.

[6] N. Tang, S. Mao, W. Yu, and R. Nelms, "Solar power generation forecasting with a lasso-based approach," IEEE Internet of Things Journal, vol. 5, no. 2, pp. 1090-1099, 2018, [Online]. Available: https://doi.org/10.1109/jiot.2018.2812155.

[7] S. Pasari and A. Shah, "Time series auto-regressive integrated moving average model for renewable energy forecasting," Sustainable Production, Life Cycle Engineering and Management, pp. 71-77, 2020, [Online]. Available: https://doi.org/10.1007/978-3-030-44248-4_7.

[8] C. Voyant, G. Notton, S. Kalogirou, M. Nivet, C. Paoli, F. Motte et al., "Machine learning methods for solar radiation forecasting: a review," Renewable Energy, vol. 105, pp. 569-582, 2017, [Online]. Available: https://doi.org/10.1016/j.renene.2016.12.095.

[9] F. Niu and Z. O'Neill, "Recurrent neural network based deep learning for solar radiation prediction," Building Simulation Conference Proceedings, 2017, [Online]. Available: https://doi.org/10.26868/25222708.2017.507.

[10] M. Sengupta, Y. Xie, A. Lopez, A. Habte, G. Maclaurin, and J. Shelby, "The national solar radiation data base (nsrdb)," Renewable and Sustainable Energy Reviews, vol. 89, pp. 51-60, 2018, [Online]. Available: https://doi.org/10.1016/j.rser.2018.03.003.

[11] V. Sharma and S. Chandel, "Performance and degradation analysis for long term reliability of solar photovoltaic systems: a review," Renewable and Sustainable Energy Reviews, vol. 27, pp. 753-767, 2013, [Online]. Available: https://doi.org/10.1016/j.rser.2013.07.046.

[12] S. Sobri, S. Koohi-Kamali, and N. Rahim, "Solar photovoltaic generation forecasting methods: a review," Energy Conversion and Management, vol. 156, pp. 459-497, 2018, [Online]. Available: https://doi.org/10.1016/j.enconman.2017.11.019.

[13] R. Ahmed, V. Sreeram, Y. Mishra, and M. Arif, "A review and evaluation of the state-of-the-art in PV solar power forecasting: techniques and optimization," Renewable and Sustainable Energy Reviews, vol. 124, p. 109792, 2020, [Online]. Available: https://doi.org/10.1016/j.rser.2020.109792.

[14] C. Yen, H. Hsieh, K. Su, M. Yu, and J. Leu, "Solar power prediction via support vector machine and random forest," E3S Web of Conferences, vol. 69, p. 01004, 2018, [Online]. Available: https://doi.org/10.1051/e3sconf/20186901004.

[15] Y. Kim and J. Kim, "Gradient lasso for feature selection," Twenty-First International Conference on Machine Learning - ICML '04, 2004, [Online]. Available: https://doi.org/10.1145/1015330.1015364.