Application of Window Sliding ARIMA in Wind Speed and Solar Irradiance Forecasting

Sarita Sheoran
Department of Mathematics, Birla Institute of Technology and Science Pilani
Pilani, India (333031)
sheoransarita93@gmail.com

Sumanta Pasari
Department of Mathematics, Birla Institute of Technology and Science Pilani
Pilani, India (333031)
86.sumanta@gmail.com

Sakshi Shukla
Department of Mathematics, Birla Institute of Technology and Science Pilani
Pilani, India (333031)
200396.sakshi@gmail.com

***Abstract*—For better management and integration of renewable energy to the existing grid system, its accurate prediction is an inevitable requirement. For such a purpose, in literature, the statistical ARIMA model is often suggested to analyze wind speed and solar irradiance values. The present study explores window sliding ARIMA (WSARIMA) for energy prediction and reports its performance with respect to the conventional ARIMA method. The wind speed and solar irradiance data (2000–2014) from two test sites, namely Dhanora (Madhya Pradesh) and Nowlaipalle (Telangana) are used for the demonstration. It is observed that both datasets for both variables (wind speed and global horizontal irradiance, GHI) exhibit weak stationarity. The parameters for the ARIMA method are obtained through grid-search technique. Then, the proposed WSARIMA approach is applied to both datasets and results are noted. Based on the RMSE values, the WSARIMA method is found to be superior for both wind speed and GHI prediction. The involvement of sliding windows essentially incorporates seasonal fluctuations more productively in both data variables–wind speed and GHI. Therefore, the present study strongly recommends the WSARIMA model for energy prediction.**

Keywords*—*GHI; Wind Speed; Forecasting; ARIMA Model; WSARIMA.

1. Introduction

Reliable forecasting of renewable energy helps in planning and estimating the energy output on a short term to a long term basis. Among different renewable energy resources, the contribution of wind and solar energy is remarkable. Wind speed and GHI predictions are useful for several practical purposes, such as estimation of energy outputs of plants, marketing of renewable energy and maintenance planning of wind farms and solar plants. Daily forecasting can help decide the best months for solar and wind energy productions, whereas monthly weather forecasting can be used for long term planning of power plants.

The renewable energy forecasting techniques described in the literature are physical, statistical and artificial intelligence methods. The most popular statistical time series based forecasting approaches are ARIMA and its several variants, such as SARIMA, f-ARIMA, ARX and WSARIMA [1-2]. Using ARIMA(2,1,2), Pasari and Shah [3] analyzed daily and monthly wind speed data. Sheoran et al. [4] have studied the efficacy of ARIMA and its competitive models in monthly, daily and hourly wind speed forecasting at a study site in Madhya Pradesh. The seasonal ARIMA model was used to predict solar irradiation at 90 different stations from five different climatic zones in India [5]. The ARMA based models have been effectively used for GHI and wind speed prediction at distinct time scales [6-8]. The main characteristic of the wind speed and GHI datasets for the current study is yearly seasonal pattern. In wind data analysis, the WSARIMA method has recently been introduced, though it has been used in several other fields including solar irradiance where the data shows seasonal variation. The inclusion of adequate window lengths effectively predicted the daily equity returns in [9]. Alberg and Last in [10] introduced the sliding windows based models for forecasting the hourly load in smart meters. Yu et al. [11] showed the prediction capability of WSARIMA in traffic analysis. Reikard [2], and Reikard and Hansen [12] suggested the WSARIMA method in GHI prediction. Sheoran et al. [13] studied the performance of the WSARIMA along with MLP and LSTM models at four test sites in India. Recently, Sheoran et al. [14] noted remarkable improvement in forecasting results by WSARIMA when compared to the traditional ARIMA model for daily and weekly prediction of wind speed data.

In this study, we consider ARIMA and WSARIMA techniques for GHI and wind speed analysis. For illustration purposes, we consider two study locations, one from Madhya Pradesh and the other from Telangana state of India. We introduce the concept of wind speed and GHI forecasting in the current section, whereas Section 2 explains the dataset. The description of the WSARIMA method and its implementation are discussed in Section 3. Section 4 presents the results of the implemented models and observations based on the obtained results followed by conclusions in Section 5.

2. DATA

Both data variables, wind speed and GHI, are obtained from the National Solar Radiation Database (NSRDB). The NSRDB dataset comprises hourly values of meteorological values. For this work, the dataset is compiled from two locations, one from Dhanora (21.65ºN,75.25ºE) in Madhya Pradesh and the other from Nowlaipalle (17.35ºN, 78.05ºE) in Telangana state of India. The datasets are recorded for 14 years starting from January 1, 2000 to December 31, 2014. More details of the dataset are provided in Table 1

978-1-6654-8910-2/22/$31.00 ©2022 IEEE

Table 1: Information about the wind speed and GHI data

Wind Speed						
Location	Time Scale	Count	Minimum	Maximum	Mean	Standard Deviation
Dhanora (Madhya Pradesh)	Daily	5475	0.375	6.433	2.639	1.104
	Weekly	783	0.800	5.820	2.639	0.984
Nowlaipalle (Telangana)	Daily	5475	0.395	7.000	2.746	1.131
	Weekly	783	0.967	6.273	2.745	1.016
GHI						
Dhanora (Madhya Pradesh)	Daily	5475	79.059	460.655	331.460	78.980
	Weekly	783	130.226	456.035	331.477	68.414
Nowlaipalle (Telangana)	Daily	5475	68.562	460.000	341.902	69.271
	Weekly	783	188.196	445.910	341.926	50.625

3. Methodology

The two implemented models and steps of implementation are described below.

The ARIMA (p,d,q) model is expressed as:

$$\phi(B)(1 - B)^d Y_t = \theta(B)Z_t$$

$$\phi(B)Y_t = Y_t - \phi_1 Y_{t-1} - \phi_2 Y_{t-2} - \ldots - \phi_{t-p} Y_{t-p}$$

$$\theta(B)Z_t = Y_t + \theta_1 Z_{t-1} + \theta_2 Z_{t-2} + \ldots + \theta_{t-q} Z_{t-q}$$

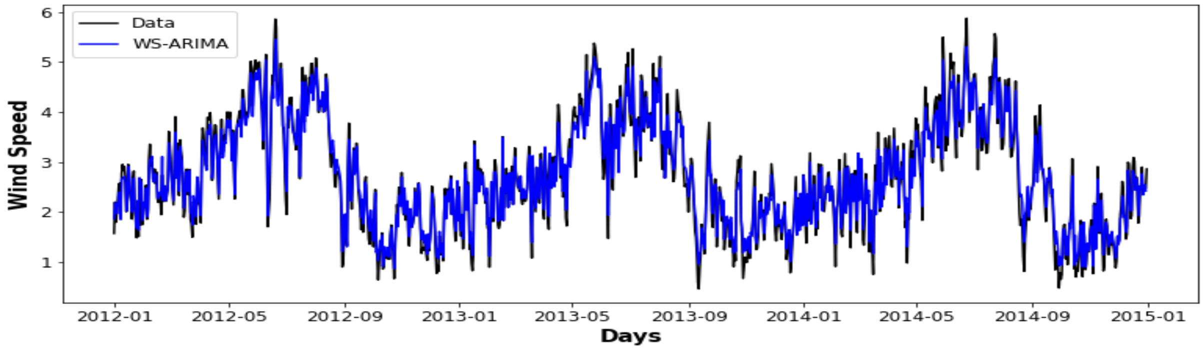
Here, $\{Y_t\}$ is the predicted variable and Z_t is the noise term; B denotes the backshift operator. The considered lags of the auto-regressive process (p), the order of differencing (d), and the number of lags in moving average (q) represent a specific order of ARIMA model.
The ARIMA model often lacks in dealing with non-linearity and high order seasonal variations. The seasonal ARIMA (s-ARIMA) model is often recommended for modeling seasonal data. However, as the SARIMA (p,d,q) $(P,D,Q)_s$ is quite complex due to its seven parameters, it is inefficient for a large dataset. The sliding window is another way of reducing the complexity of such models due to its dynamic nature using the fixed window size [2, 15]. In the sliding window technique, using data properties, first the size of the window is found. The detailed description of the WSARIMA method is explained in [14]. The optimal lengths for the sliding windows is 365 for the daily forecasting and 52 for the weekly forecasting, guided by the yearly seasonal pattern in the datasets.
The first step of methodology is data preprocessing, where re-sampling is carried out for daily and weekly time horizons. Then, the dataset is divided into training (first 11 years) and testing (3 years) sets. Augmented Dickey Fuller method is applied to test whether the dataset has stationarity property. Then, the ARIMA and WSARIMA are implemented on the test data. For ARIMA, a grid search technique is adopted to obtain the minimum AIC and RMSE values. Here the “statsmodels” and “pmdarima” python libraries are used. The performance of the WSARIMA model is realized through RMSE measures.

4. Results

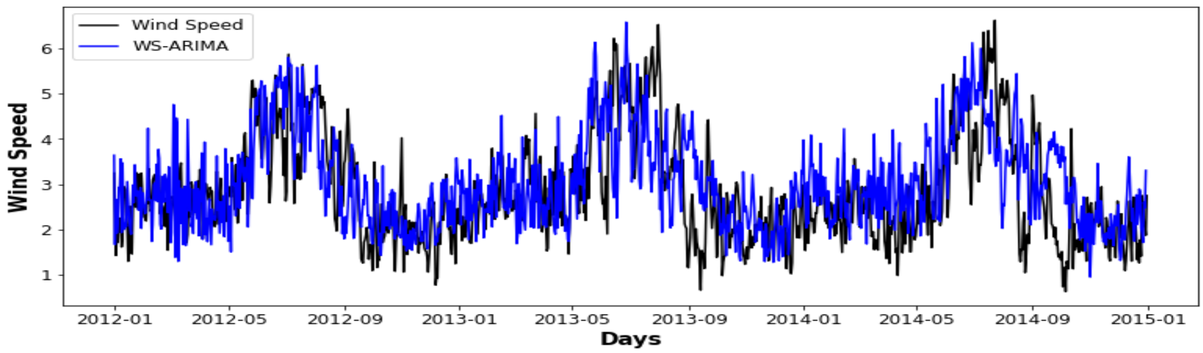
Using RMSE values, Table 2 and Table 3 display the performance of the WSARIMA method in daily and weekly prediction of wind speed and GHI data. Actual measurements and predicted values of daily and weekly time horizons for both variables across both locations are shown in Figure 1 and Figure 2, respectively. It is noted that the differencing parameter turns out to be zero for all cases, as expected in stationary time series data. Therefore, the ARIMA model here is nothing but an ARMA model. The RMSE values for daily and weekly wind data corresponding to WSARIMA method are significantly lower than that of ARIMA model, indicating a better performance of the WSARIMA method. Similarly, the RMSE values related to the WSARIMA method in GHI prediction across both locations and both time horizons are significantly lesser than the corresponding ARIMA model. Therefore, the WSARIMA model achieves better performance in wind speed and GHI prediction. Particularly, as noted in [], the RMSE values of ARIMA are almost twice that of WSARIMA in daily wind speed prediction, and they are almost four times in case of weekly prediction of wind speed data. In the present analysis, moving from daily to weekly wind speed prediction increases RMSE for the WSARIMA model, whereas in GHI prediction, the RMSE measures of the WSARIMA model decrease if one moves from daily to weekly time horizon.

5. Conclusion

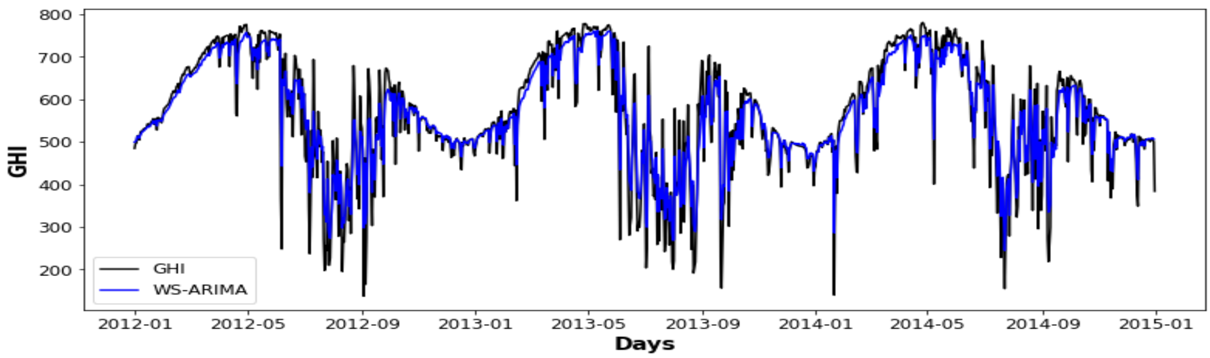
Based on the wind speed and GHI data from two locations in India, the present work highlights the usefulness of the WSARIMA technique in energy prediction. It is observed that the WSARIMA technique significantly produces lesser RMSE values than that of the ARIMA model. The involvement of sliding windows essentially incorporates seasonal fluctuations better in both data variables. Therefore, the present analysis strongly suggests to consider the WSARIMA model in energy prediction.



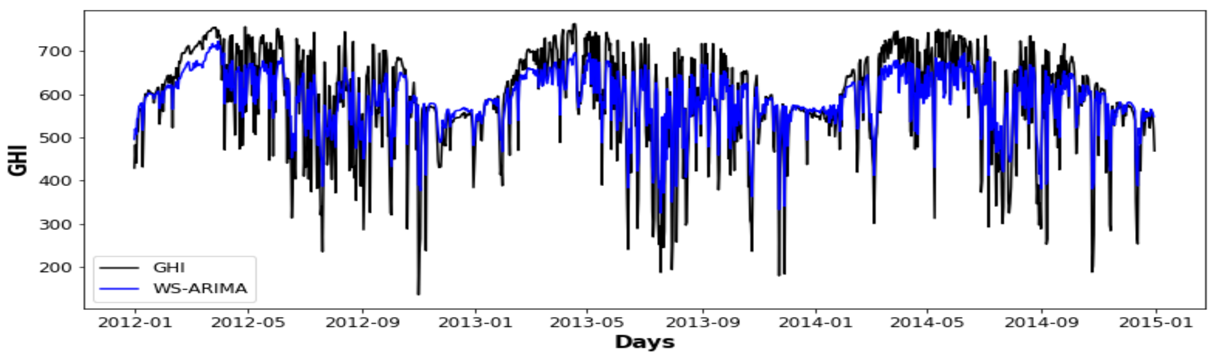
(A)



(B)

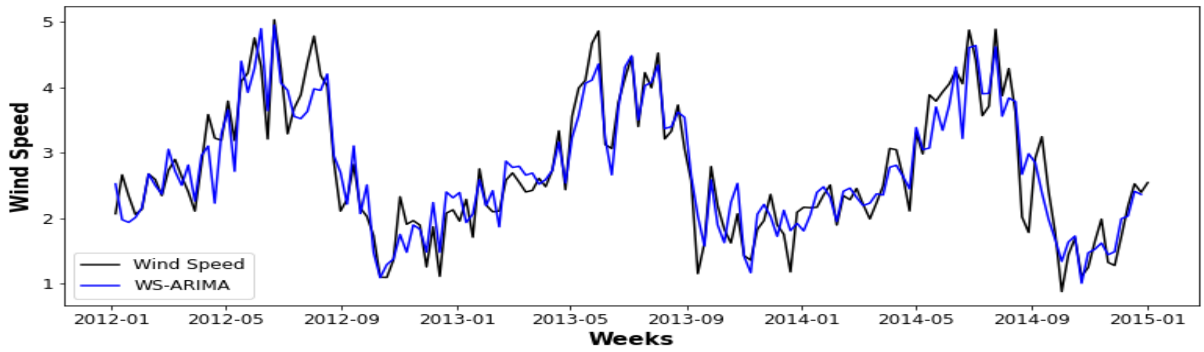


(C)

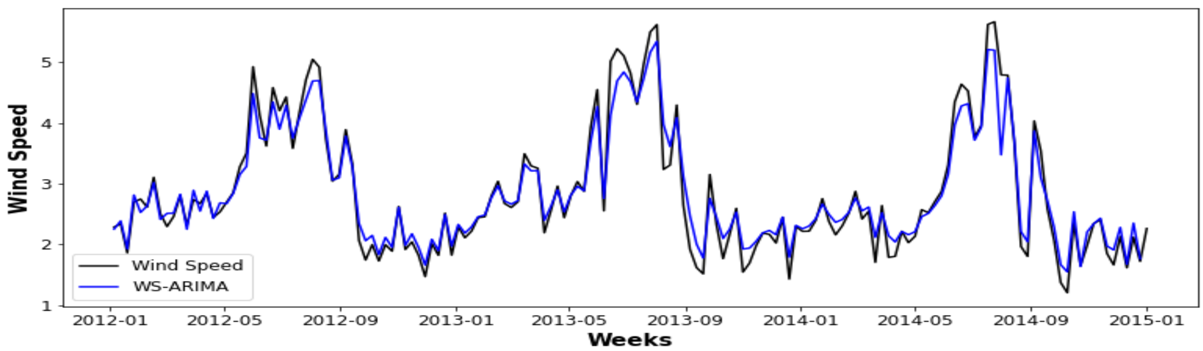


(D)

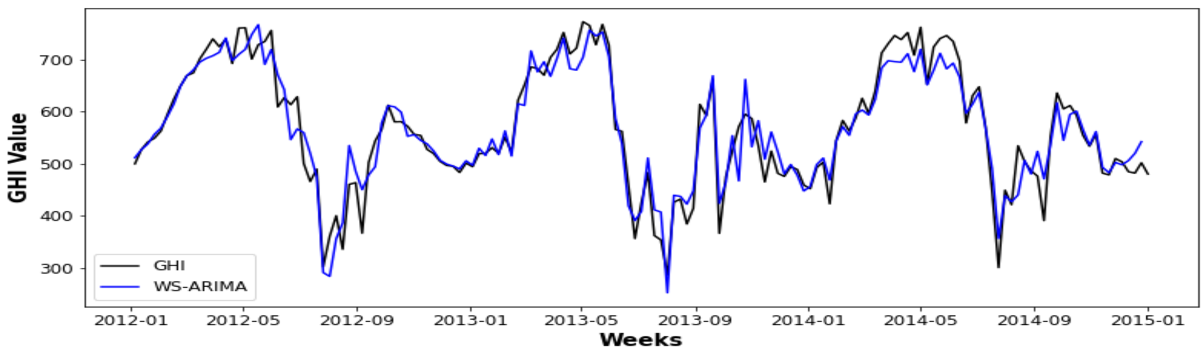
Figure 1: Daily prediction through the WSARIMA method: (A) wind speed in Madhya Pradesh, (B) wind speed in Telangana, (C) GHI in Madhya Pradesh, and (D) GHI in Telangana.



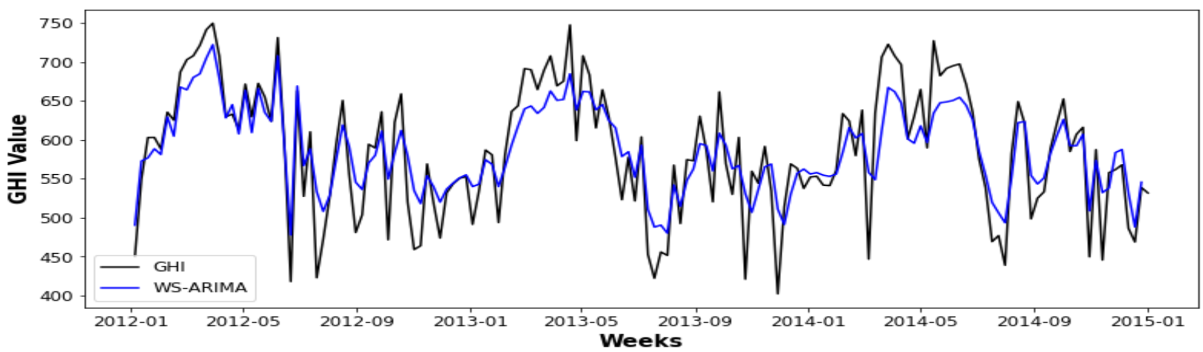
(A)



(B)



(C)



(D)

Figure 2: Weekly prediction through the WSARIMA method: (A) wind speed in Madhya Pradesh, (B) wind speed in Telangana, (C) GHI in Madhya Pradesh, and (D) GHI in Telangana.

Table 2: Optimal parameter values and RMSE measures in wind speed prediction.

Location	Method	Daily		Weekly	
		(p,d,q)	RMSE	(p,d,q)	RMSE
Dhanora (Madhya Pradesh)	ARIMA	(1,0,3)	0.9077	(3,0,2)	0.6014
	WSARIMA		0.2022		0.3463
Nowlaipalle (Telangana)	ARIMA	(2,0,3)	0.9716	(1,0,2)	0.6762
	WSARIMA		0.1735		0.2420

Table 3: Optimal parameter values and RMSE measures in GHI prediction.

Location	Method	Daily		Weekly	
		(p,d,q)	RMSE	(p,d,q)	RMSE
Dhanora (Madhya Pradesh)	ARIMA	(1,0,2)	100.3923	(3,0,2)	68.3457
	WSARIMA		33.7763		33.6790
Nowlaipalle (Telangana)	ARIMA	(2,0,4)	114.6342	(1,0,1)	64.8810
	WSARIMA		52.6336		39.1156

Acknowledgements

Careful and thoughtful reviews from two anonymous reviewers are greatly appreciated. Sarita and Sakshi acknowledge the financial assistance in terms of SRF and JRF, respectively.

References

[1] S. S. Soman, H. Zareipour, O. Malik, and P. Mandal, "A review of wind power and wind speed forecasting methods with different time horizons," *North American Power Symposium 2010*, pp. 1–8, 2010.

[2] G. Reikard, "Predicting solar radiation at high resolutions: A comparison of time series forecasts," *Solar Energy*, vol. 83, pp. 342–349, 2009.

[3] S. Pasari and A. Shah, "Time series auto-regressive integrated moving average model for renewable energy forecasting," *Sustainable Production, Life Cycle Engineering and Management*, 2020.

[4] S. Sheoran, R. Bavdekar, S. Pasari, and R. Kulshrestha, "Wind speed forecasting using time series methods: a case study," . *In Chamola BP, Kumari P, Kaur L (eds) Emerging Advancements in Mathematical Sciences, Nova publishing*, pp. 125– 133, 2022.

[5] A. K. Barik., S. Malakar, S. Goswami, B. Ganguli, S. R. Sen and A. Chakrabarti, "Analysis of GHI Forecasting Using Seasonal ARIMA" In *Data Management, Analytics and Innovation*, Singapore, Springer, pp. 55-69, 2021.

[6] M. H. Alsharif, M. K. Younes, and J. Kim, "Time series ARIMA model for prediction of daily and monthly average global solar radiation: The case study of Seoul, South Korea," *Symmetry*, vol. 11, pp. 240, 2019.

[7] H. Sharadga, S. Hajimirza, and R. S. Balog, "Time series forecasting of solar power generation for large-scale photovoltaic plants," *Renewable Energy*, vol. 150, pp. 797–807, 2020.

[8] L. Martín, L. F. Zarzalejo, J. Polo, A. Navarro, R, Marchante, and M. Cony, "Prediction of global solar irradiance based on time series analysis: Application to solar thermal power plants energy production planning, ʻʼ *Solar Energy*, vol. 10, pp. 1772-1781, 2010.

[9] H. Dong, X. Guo, H. Reichgelt, and R. Hu, "Predictive power of ARIMA models in forecasting equity returns: a sliding window method," *Journal of Asset Management*, vol. 21, pp. 549–566, 2020.

[10] D. Alberg and M. Last, "Short-term load forecasting in smart meters with sliding window-based ARIMA algorithms," *Vietnam Journal of Computer Science*, vol. 5, pp. 241–249, 2018.

[11] Q. Yu, L. Jibin, and L. Jiang, "An improved ARIMA-based traffic anomaly detection algorithm for wireless sensor networks," *International Journal of Distributed Sensor Networks*, vol. 12, 2016.

[12] G. Reikard and C. Hansen, "Forecasting solar irradiance at short horizons: Frequency and time domain models," *Renewable Energy*, vol. 135, pp. 1270–1290, 2019.

[13] S. Sheoran, R. S. Singh, S. Pasari, and R. Kulshrestha, "Forecasting of solar irradiances using time series and machine learning models: A case study from India," *Applied Solar Energy*, vol 58, pp. 137-135, 2022.

[14] S. Sheoran and S. Pasari, "Efficacy and application of the window-sliding ARIMA for daily and weekly wind speed forecasting", Journal *of Renewable and Sustainable Energy*, vol. 14, pp. 053305, 2022.

[15] A. Inove, L. Jin, and B. Rossi, "Rolling window selection for out-of-sample forecasting with time-varying parameters," Journal *of Econometrics*, vol. 196, no. 1, pp. 55–67, 2017.