Research Article

Rainfall, Wind Speed, and Temperature Forecast Using Triple Exponential Smoothing and Gradient Descent

Dimara Kusuma Hakim[1*], Rahmat Gernowo[2], Anang Widhi Nirwansyah[3]

[1]Doctoral Program of Information Systems, Diponegoro University, 50275, Indonesia
[2]Department of Physic Faculty of Mathematics and Natural Sciences, Diponegoro University, 50275, Indonesia
[3]Department of Geography Education, Universitas Muhammadiyah Purwokerto, 53182, Indonesia

ORCID
Dimara Kusuma Hakim: https://orcid.org/0000-0003-1568-7362

Corresponding Author: Dimara Kusuma Hakim; email: dimarakusumahakim@gmail.com

Published 12 March 2024

Publishing services provided by Knowledge E

Selection and Peer-review under the responsibility of the ICESRE Conference Committee.

Abstract.
The global community strives to minimize the impact of disasters through various actions, for example, mapping flood-prone areas. Flood-prone areas need to be identified correctly, predicted, understood, and socialized to minimize risks when a disaster occurs regarding death, property damage, and socio-economic losses. This type of data-based prediction has been developed and implemented widely and can be applied to predictions related to hydrology. Data mining approaches (estimation, classification, clustering, and time-series forecast) have significantly influenced research related to flood prediction in recent years. The time-series flood forecast has been widely used in previous research using various statistical and data-mining methods. Predicting floods that occur in coastal areas is less discussed than river floods. One method that is often used is exponential smoothing. Determining damping factor values (alpha, beta, and gamma) in the triple exponential smoothing method, in general, is to use all values from 0 to 1 to find the most optimal damping factor, this takes quite a long time and results generally appear with less accuracy. So, a combination of the triple exponential smoothing algorithm is proposed to perform tTimeseries forecast, and the gradient descent algorithm is used as an optimization algorithm to obtain optimal weight values for alpha, beta, and gamma in triple exponential smoothing.

Keywords: triple exponential smoothing, gradient descent, flood forecast, flood prediction, time-series forecast

1. Introduction

The coast is the place where land and sea meet that is in between them [1]. Along the landward shore, there are areas of land that are both dry and underwater that are still affected by things like sea breezes, tides, and saltwater seepage. Along the shore, parts of the sea are still affected by things that happen naturally on land, like sedimentation and freshwater flows, as well as things that people do on ground, like cutting down trees and polluting. One of the risks of disasters that occur in coastal areas is flooding

OPEN ACCESS

events related to tidal waves or floods influenced by tides, of course heavy rain can also be one of the main causal factors.

Flood forecasting is essential to flood management, as there is no more practical measure of prevention or defense [1,2]. A flood forecasting and forecasting system can be developed as a combination of various elements in an integrated flood analysis system [2–5]. The hope is that flood forecasting in the form of a flood susceptibility map can minimize the impact or consequences of flooding. The analysis that produces a flood susceptibility map provides information about the potential for flooding in the next few years. In previous research that has been carried out, the input used can be (1) a combination of static data (slope, elevation, etc.) and dynamic data (rain, wind, etc.), (2) only static data, and (3) only dynamic data. Dynamic data, namely Meteorological Factors, consist of rain, wind, humidity, and temperature [6–8]. Data-based prediction has been developed and implemented widely and can be applied to predictions related to hydrology. Data mining approaches have significantly influenced research related to flood prediction in recent years.

Time-series Flood Forecast has been widely used in previous research using various statistical and data-mining methods. Predicting floods that occur in coastal areas is less discussed than river floods. One often-used method is Exponential Smoothing [9–11]. Exponential smoothing is a way to predict future values for a time series by taking an exponentially weighted average of past readings [12]. With this method, newer findings are given more weight than older ones. This lets the forecast change as the data shows changing trends. When it comes to exponential smoothing, there are different kinds, from the most basic to the most advanced. The first type is called Simple Exponential Smoothing when there is no trend or seasonality. When there is a trend, it is called Double Exponential Smoothing. When there is a trend and seasonality, it is called Triple Exponential Smoothing [13]. Because of its limited accuracy, exponential smoothing is often reserved for making short-term predictions.

Determining damping factor values (alpha, beta, and gamma) in the Triple Exponential Smoothing method, in general, is to use all deals from 0 to 1 to find the most optimal damping factor, this takes quite a long time and with results which has less accuracy [14,15]. So, a combination of the Triple Exponential Smoothing algorithm is proposed to perform Timeseries Forecast, and the Gradient Descent algorithm is used as an optimization algorithm to obtain optimal weight values for alpha, beta, and gamma in Triple Exponential Smoothing.

New data mining techniques have encouraged researchers to increase research efforts to obtain increasingly accurate flood predictions. Data mining techniques are based on machine learning techniques that can intelligently generate rules and patterns from large amounts of data [16]. Identification and prediction of hydrology-related disasters, which are carried out using time-series data mining analysis methods, are influenced by the availability of historical data that is large enough to present strategic decision alternatives in the future [17,18].

2. Method

2.1. Triple Exponential Smoothing

Triple Exponential Smoothing, also known as Holt's Winter Exponential Smoothing, is a time-series forecasting method that extends Simple Exponential Smoothing by adding trend and seasonal components. This method uses two smoothing constants, one for level and one for trend, to update the components for each period. The equation used is as follows in Eq. (1), L_t refers to Level. In exponential smoothing, the word "level" means a rough guess of the local mean or "level" of the process of making the data at a specific time. This part of the forecast considers long-lasting effects, like the average value. The level is found by taking the weighted average of past readings, where the weights get smaller and smaller as the observations get older. A quantity called α, which can be any number from 0 to 1, controls the weights. If α is high, the most recent measurement is more critical. If it is low, it means that the most recent prediction is more important. This Level component is used in Single Exponential Smoothing. In Eq. (2), T_t (trend) refers to a consistent upward or downward pattern in the data over time, in other word, trend is the increase or decrease in the series. The addition of the Trend component is used in Double Exponential Smoothing. In Eq. (3), it is possible to determine seasonality (S_t), level, and trend (T_t) in triple exponential smoothing. This works well for data that has movements that repeat over time. When exponential smoothing is used, the seasonal part looks at patterns or trends over a specific period, like a day, week, month, season, etc. This approach works best for data that has clear seasonal trends, and it is often used to predict sales, among other things. When discussing exponential smoothing, "seasonal" means that the forecast model uses the seasonality factor. In this way, the model can consider trends in the data that happen at set times. For instance, store sales might rise yearly in December and fall every year after. Exponential smoothing can make better predictions for this kind of data by including this "seasonal" factor in the model. Forecast in the Eq. (4) is a formula used to make future predictions. The definitions of all variables used are: α is the weight (or smoothing constant) for the level, Y_t is the observed value at time t, L_{t-1} is the level at time t-1, T_{t-1} is the trend at time t-1, S_t is the smoothed value at time t, the measure β (beta) smooths out the trend that has a value between 0 and 1 and tells the computer how much weight to give to the trend component.

Table 1:

$L_t = \alpha(y_t / S_{t-s}) + (1-\alpha)(L_{t-1} + T_{t-1})$	(1)
$T_t = \beta(L_t - L_{t-1}) + (1-\beta)\ T_{t-1}$	(2)
$S_t = \gamma(y_t / L_t) + (1-\gamma)\, S_{t-s}$	(3)
$\hat{y}_{t+k} = (L_t + k\, T_t\,)\ \ S_{t+k-s}$	(4)

2.2. Gradient Descent

Optimization is an algorithm that can find optimal values by minimizing or maximizing the objective function (error function). The output of all calculation processes is studied and updated towards the optimal solution, namely by reducing losses with the Training process. Gradient Descent is an algorithm used in Machine Learning. This algorithm is an optimization algorithm to find the minimum value of a function by minimizing the loss function. This process occurs in the backpropagation phase. The gradient is the value of the slope or inclination of a line that compares the y component (ordinate) with the x component (abscissa). The gradient here is the slope value of a function, namely the rate of change of a parameter relative to the number of other parameters. Mathematically, a gradient can be described as a partial derivative of a series of parameters concerning its input—the more gradients, the steeper the slope. Gradient Descent is also described as iteration used to find a parameter's optimal value (Fig. 1) by using calculus to find the minimum value. Gradient Descent is used to update weights by minimizing the loss function. This Gradient Descent process occurs in the backpropagation phase. Gradient Descent is a process that occurs in the backpropagation phase where the goal is to continuously change the gradient of the model parameters in the opposite direction based on the weight w, updating consistently until it reaches the global minimum of the function.

2.3. Data

Data was taken geometrically on the coast of Pekalongan City, from 2 sources: meteorological data from NASA: GMAO MERRA-2 (satellite), and Sea Level from BIG (Geospatial Information Agency) from tidal measurements at the harbor, over a period of 4

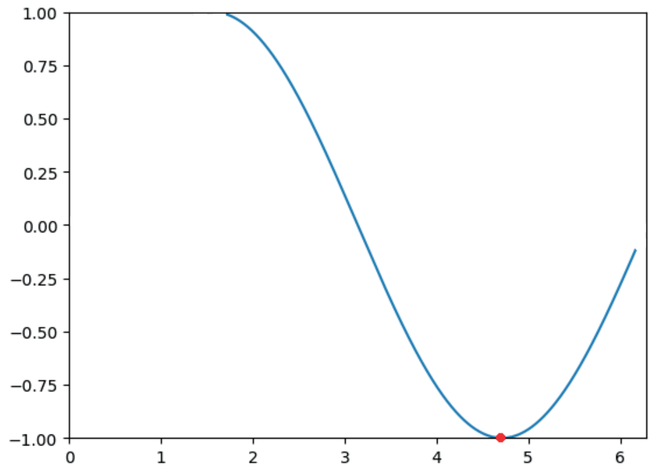

Figure 1: Get Local Minima with Gradient Descent.

years (2019 -2022). Meteorological data includes: Temperature at 2 Meters (C), Specific Humidity at 2 Meters (g/kg), Precipitation Corrected (mm/day), and Wind Speed at 10 Meters (m/s). We use the maximum value for each month in this study.

3. Result and Discussion

3.1. Temperature

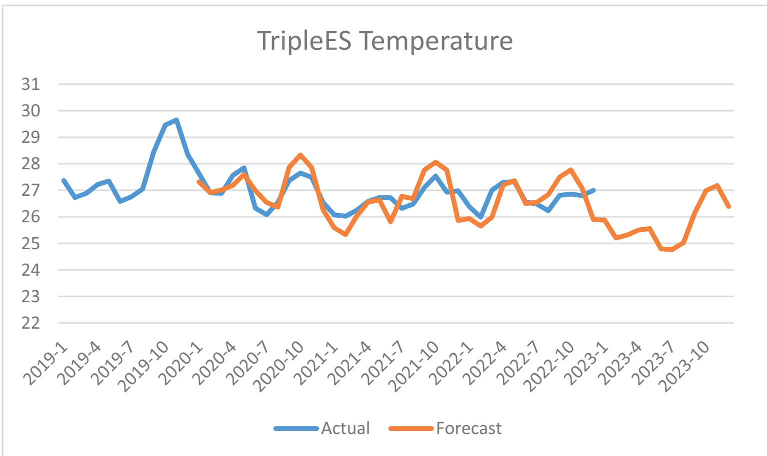

Figure 2: Maximum temperature graph for each month.

Based on Fig. 2, the temperature pattern in Pekalongan varies greatly. Data increased at the end of 2019 and continued in early 2020. The pattern of temperature in the picture experiences fluctuations every year, where the temperature tends to rise in November, December, and January, while for other months, the data experiences a downward trend. This repeats itself every year, so it can indicate that there is a seasonal pattern in the data. In the picture, the data experiences an increase or decrease every year. Monthly observations suggest that levels, trends, and seasonality are formed in some months and other months experience fluctuations. It can be seen in the picture that the data pattern shows. Data patterns tend to fluctuate so that the temperature is calculated. In the calculations, the season's length is 12 because what is used is monthly for one year. Next, analyses were carried out for the initial data smoothing values using Eq (1), (2), [3], [4), and the results obtained were as shown in Table 1.

The training results on the data produced optimal values alpha=0.88228094, beta=0, and gamma=1. Then the evaluation results have MAE=0.45, MSE=0.30, RMSE=0.55, and MAPE=1.66%. It was concluded that there is no trend component in temperature but an absolute seasonality component.

Table 2: Initial Trend of Temperature.

Initial Trend			
Year 1	**Year 2**	**Y2 - Y1**	**(Y2 - Y1) / 12**
27,37	27,64	0,27	0,02
26,73	26,9	0,17	0,01
26,89	26,89	0	0,00
27,22	27,56	0,34	0,03
27,35	27,85	0,5	0,04
26,58	26,33	-0,25	-0,02
26,76	26,08	-0,68	-0,06
27,05	26,54	-0,51	-0,04
28,47	27,37	-1,1	-0,09
29,46	27,65	-1,81	-0,15
29,65	27,49	-2,16	-0,18
28,35	26,55	-1,8	-0,15
			-0,05

3.1.1. Humidity

Figure 3 shows that the pattern of humidity in Pekalongan is very different. Data improved at the end of 2019 and is still going strong in early 2020. When you look at the picture, the pattern changes every year. The data tends to go up in November, December, January, February, March and April, but in other months, it goes down. It does this same thing every year, which could mean the data has a seasonal trend. The info in the picture looks up or down every year. Based on monthly observations, levels, trends, and patterns are formed in some months while they change in others. The picture shows the data trend by looking at how the data trends change over time.

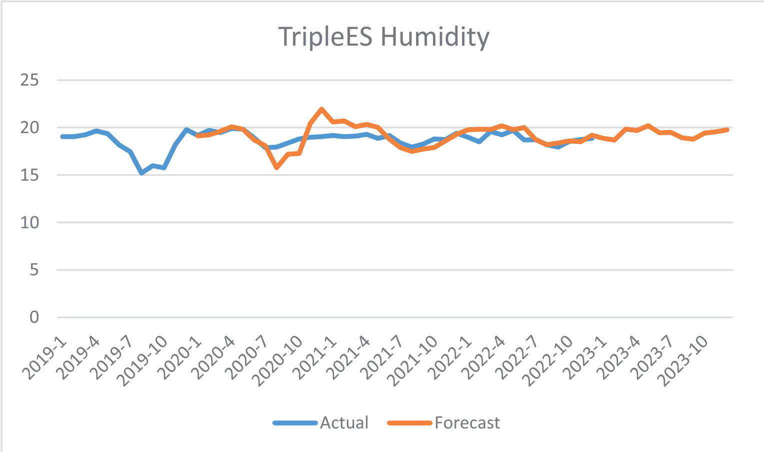

Figure 3: Maximum Humidity graph for each month.

The calculations say the season lasts 12 months because that's how long a year is. Next, equations (1), (2), [3, 4], and (9) were used to figure out the initial data smoothing values. The outcomes are shown in Table 2.

After training on the data, the best numbers were alpha=0.2, beta=0, and gamma=1. The study then shows that MAE=0.71, MSE=0.96, RMSE=0.98, and MAPE=3.77%. To sum up, it was found that humidity does not have a trend component but an absolute annual component.

Table 3: Initial Trend of Humidity.

Initial Trend			
Year 1	**Year 2**	**Y2 - Y1**	**(Y2 - Y1) / 12**
19,04	19,17	0,13	0,01
19,04	19,71	0,67	0,06
19,23	19,47	0,24	0,02
19,65	19,9	0,25	0,02
19,35	19,84	0,49	0,04
18,19	18,92	0,73	0,06
17,46	17,88	0,42	0,03
15,2	17,94	2,74	0,23
15,99	18,37	2,38	0,20
15,75	18,8	3,05	0,25
18,19	18,98	0,79	0,07
19,78	19,04	-0,74	-0,06
			0,08

DOI 10.18502/kss.v9i6.15311

3.2. Precipitation

Based on Figure 2, the rainfall pattern in Pekalongan varies greatly. Rainfall data increased in November but decreased in December and continued to increase again in January, February, March and April. Seen in the image which shows the data pattern. The data pattern tends to fluctuate so the data is calculated. By using Equations (1), (2), [3], [4), and the results obtained are as in Table 3.

The training results on the data produce optimal values alpha=0, beta=0.4, and gamma=0. Then the evaluation results have MAE=9.31, MSE=275.33, RMSE=16.59, and MAPE=72.44%. It was concluded that there is no seasonality component in rainfall.

Table 4: Initial Trend of Precipitation.

Initial Trend			
Year 1	**Year 2**	**Y2 - Y1**	**(Y2 - Y1) / 12**
9,19	7,99	-1,2	-0,10
9,76	8,48	-1,28	-0,11
9,44	9,86	0,42	0,04
9,64	8,07	-1,57	-0,13
9,3	8,47	-0,83	-0,07
0,53	6,96	6,43	0,54
2,06	34,73	32,67	2,72
1,58	9,39	7,81	0,65
6,8	6,01	-0,79	-0,07
3,94	9,91	5,97	0,50
9,77	9,39	-0,38	-0,03
7,58	7,6	0,02	0,00
			0,33

3.3. Wind Speed

Figure 4 shows that the wind speed trend in the Pekalongan changes over time. It's getting better at the start of the year. Every year, the pattern of data in the picture calls bigger. The data got bigger in January, but in other months, it got smaller. In this case, it is repeated every year to show that the data has seasonal trends. It's still possible for the info in this picture to go up or down. Monthly data sets levels, trends, and seasonality in some months, while gains happen in others. The next step was to use Equations (1), (2), [3, 4], and the results were shown in Table 4 to make sure the original data was correct.

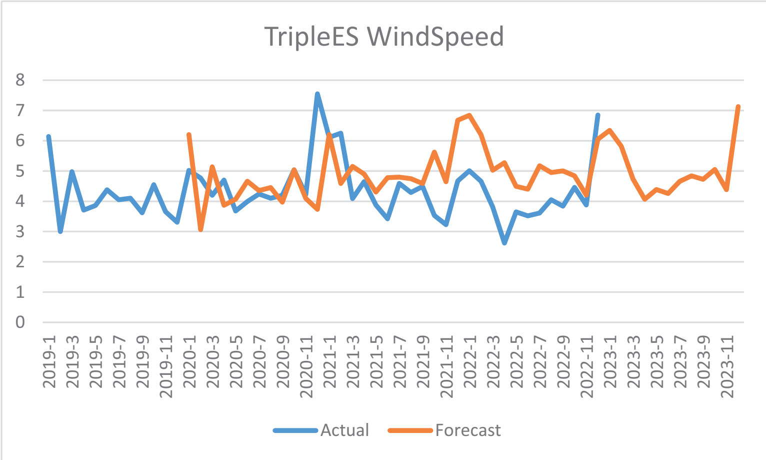

Figure 4: Maximum Wind Speed graph for each month.

It was found that 0 for alpha, 0.4 for beta, and 0.6 for gamma were the best numbers. It was then evaluated and found that MAE=0.99, MSE=1.64, RMSE=1.28, and MAPE=23.51%. This shows the existence of Trend and Seasonal patterns.

Table 5: Initial Trend of Wind Speed.

Initial Trend			
Year 1	**Year 2**	**Y2 - Y1**	**(Y2 - Y1) / 12**
6,14	5,02	-1,12	-0,09
3	4,77	1,77	0,15
4,98	4,2	-0,78	-0,07
3,71	4,7	0,99	0,08
3,86	3,68	-0,18	-0,02
4,38	3,99	-0,39	-0,03
4,05	4,24	0,19	0,02
4,1	4,1	0	0,00
3,62	4,2	0,58	0,05
4,55	5,04	0,49	0,04
3,66	4,22	0,56	0,05
3,31	7,55	4,24	0,35
			0,04

3.4. Sea Level

Figure 5 shows that the Sea Level in Pekalongan has a specific pattern. Data increased in April and May. If you look at the picture, the pattern changes every year. The same thing happens yearly, meaning the data has a seasonal trend. The info in the shot goes up or down every year. Based on observations made every month, it can be seen that

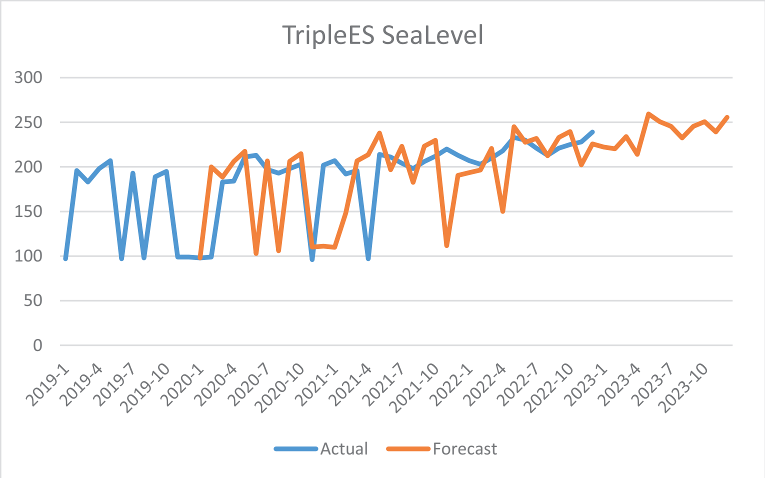

Figure 5: Maximum Sea Level graph for each month.

levels, trends and patterns form in some months and change in others. The figure shows the data trend, looking at how the data trend changes over time. The calculations say the season lasts 12 months because that's how long a year is. Next, equations (1), (2), [3, 4], and (9) are used to determine the smoothing value of the initial data. The results are shown in Table 5.

After training the data, the best numbers were found at alpha=0, beta=0, and gamma=0,7. Research then shows that MAE=32.33, MSE=2359.84, RMSE=48.58, and MAPE=19.02%. In conclusion, it is found that Sea Level does not have a trend component but a seasonal component.

Table 6: Initial Trend of Sea Level.

Initial Trend			
Year 1	**Year 2**	**Y2 - Y1**	**(Y2 - Y1) / 12**
97	98	1	0,08
196	99	-97	-8,08
183	183	0	0,00
198	184	-14	-1,17
207	211	4	0,33
97	213	116	9,67
193	197	4	0,33
98	193	95	7,92
189	198	9	0,75
195	203	8	0,67
99	96	-3	-0,25
99	202	103	8,58
			1,57

DOI 10.18502/kss.v9i6.15311

4. Conclusion

The conclusion obtained from the research that has been carried out is that the maximum rainfall, sea level, humidity, wind speed and monthly temperature data on the Pekalongan Coast have varying Trend and Seasonal components. In the next research, it is hoped that we will be able to conduct research on large amounts of data, namely daily data, and use a Hybrid Method approach. The forecast results with Triple Exponential Smoothing show good MAPE values for the variables: Temperature (1.66%), Humidity (3.77%), Wind Speed (23.51%) and Sea Level (19.02%), but have bad forecasting results in the Precipitation variable (72.44%). The many variables involved in rain prediction in this process are the reason why rain is often difficult to predict. To determine tomorrow's weather, experts usually use today's weather guidelines, this indicates that it is still difficult to predict rain for long-term forecasts. The implication of this research is that it is one of the stages of research in predicting floods, starting with a univariate time-series approach. It is hoped that future research can describe all the variables in this research in a Multivariate Time-series approach.

Acknowledgements

This study received support from Doctoral Program of Information System, School of Postgraduate Studies, Universitas Diponegoro

References

[1] Marijnissen RJC, Kok M, Kroeze C, van Loon-Steensma JM. Flood risk reduction by parallel flood defences – Case-study of a coastal multifunctional flood protection zone. Coast Eng 2021;167. https://doi.org/10.1016/j.coastaleng.2021.103903.

[2] Koks EE, Le Bars D, Essenfelder AH, Nirandjan S, Sayers P. The impacts of coastal flooding and sea level rise on critical infrastructure: a novel storyline approach. Sustain Resilient Infrastruct 2023;8:237–61. https://doi.org/10.1080/23789689.2022.2142741.

[3] Panezai S, Moniruzzaman, Saqib SE, Rahman MS, Ferdous Z, Asghar S, et al. Rural households' food security and its determinants in coastal regions of Bangladesh. Nat Resour Forum 2022;46:200–20. https://doi.org/10.1111/1477-8947.12250.

[4] Saravanan S, Abijith D. Flood susceptibility mapping of Northeast coastal districts of Tamil Nadu India using Multi-source Geospatial

data and Machine Learning techniques. Geocarto Int 2022;37:15252–81. https://doi.org/10.1080/10106049.2022.2096702.

[5] Shan X, Wang J, Wen J, Hu H, Wang L, Yin J, et al. Using Multidisciplinary Analysis to Develop Adaptation Options against Extreme Coastal Floods. Int J Disaster Risk Sci 2022;13:577–91. https://doi.org/10.1007/s13753-022-00421-6.

[6] Latif S, Simonovic SP. Trivariate Probabilistic Assessments of the Compound Flooding Events Using the 3-D Fully Nested Archimedean (FNA) Copula in the Semiparametric Distribution Setting. Water Resour Manag 2023;37:1641–93. https://doi.org/10.1007/s11269-023-03448-6.

[7] Khouakhi A, Villarini G. On the relationship between atmospheric rivers and high sea water levels along the U.S. West Coast. Geophys Res Lett 2016;43:8815–22. https://doi.org/10.1002/2016GL070086.

[8] Lavers DA, Villarini G. The contribution of atmospheric rivers to precipitation in Europe and the United States. J Hydrol 2015;522:382–90. https://doi.org/10.1016/j.jhydrol.2014.12.010.

[9] Kamath RS, Kamat RK. Time-series analysis and forecasting of rainfall at Idukki district, Kerala: Machine learning approach. Disaster Adv 2018;11:27–33.

[10] Hayder IM, Al-Amiedy TA, Ghaban W, Saeed F, Nasser M, Al-Ali GA, et al. An Intelligent Early Flood Forecasting and Prediction Leveraging Machine and Deep Learning Algorithms with Advanced Alert System. Processes 2023;11. https://doi.org/10.3390/pr11020481.

[11] de Faria EL, Albuquerque MP, Gonzalez JL, Cavalcante JTP, Albuquerque MP. Predicting the Brazilian stock market through neural networks and adaptive exponential smoothing methods. Expert Syst Appl 2009;36:12506–9. https://doi.org/10.1016/j.eswa.2009.04.032.

[12] Ostertagová E, Ostertag O. Forecasting using simple exponential smoothing method. Acta Electrotech Inform 2013;12:10198. https://doi.org/10.2478/v10198-012-0034-2.

[13] Mi J, Huang G. Dynamic Prediction of Performance Degradation Characteristics of Direct-Drive Electro-Hydraulic Servo Valves. Appl Sci 2023;13:7231. https://doi.org/10.3390/app13127231.

[14] Susila IMD, Atmojo YP, Srinadi NLP, Suradarma IB, Yuningsih L, Rini ES. Performance Analysis of the Triple Exponential Smoothing Method During the Covid19 Pandemic on Tourist Visit Data. 2022 1st Int. Conf. Inf. Syst. Inf. Technol., IEEE; 2022, p. 358–63. https://doi.org/10.1109/ICISIT54091.2022.9872863.

[15] Djakaria I. Djalaluddin Gorontalo Airport Passenger Data Forecasting with Holt's-Winters' Exponential Smoothing Multiplicative Event-Based Method. J Phys Conf Ser 2019;1320:012051. https://doi.org/10.1088/1742-6596/1320/1/012051.

[16] Halim MH, Wook M, Hasbullah NANAA, Razali NNAMAM, Hamid HECECHHECEC. Comparative Assessment of Data Mining Techniques for Flash Flood Prediction. Int J Adv Soft Comput Its Appl 2022;14:125–45. https://doi.org/10.15849/IJASCA.220328.09.

[17] Ayyad M, Hajj MR, Marsooli R. Machine learning-based assessment of storm surge in the New York metropolitan area. Sci Rep 2022;12. https://doi.org/10.1038/s41598-022-23627-6.

[18] Xiang Z, Yan J, Demir I. A Rainfall-Runoff Model With LSTM-Based Sequence-to-Sequence Learning. Water Resour Res 2020;56:1–17. https://doi.org/10.1029/2019WR025326.