

Prediction of Monsoon Rainfall by using Holt-Winters Models for Junagadh (Gujarat-India) Region

Gundalia Manoj J.

Department of Civil Engineering, C. G. Patel Institute of Technology, UTU, Bardoli, Gujarat, 394601, INDIA
mjgundalia@gmail.com

Abstract

Rainfall becomes the lifeline of the most world's population due to its yawning impact on agriculture, drinking water and energy sectors. Monsoon rainfall is a very imperative parameter for agriculture dominant country likes India; its prediction helps in planning of water management, environmental, agricultural and meteorological projects. Rainfall is highly complex, nonlinear, dynamic in nature and is affected by many interrelated meteorological parameters. Further the temporal and spatial variability of these parameters causes more uncertainty in its occurrence. Despite significant contribution of advance computing techniques, the rainfall prediction yet remains a tough challenge. Monsoon rainfall data often exhibits trend and seasonality and hence, Holt-Winters models could be the best choice for its prediction.

This study examined ability of three different types of time series models (Holt-Winters model (HW), Multiplicative Holt-Winters (MHW) and Additive Holt-Winters (AHW)) in predicting monsoon rainfall for Junagadh (Gujarat-India) region. Performances of the models were evaluated at monthly and annual time scales by using refined Willmott's index (d_r) and mean absolute error (MAE) evaluation measures. All the three models gave better results at annual time scale than that of monthly time scale. Among the three models, HW and MHW performed better with d_r value 0.73 and can be recommended for forecasting annual monsoon rainfall in the similar hydro-meteorological region.

Keywords: Monsoon Rainfall, Meteorological Parameters, Holt-Winters Model, Willmott's Index, Mean Absolute Error.

Introduction

Prediction of monsoon rainfall plays a vital role in planning and design of many activities of water resources and agriculture sectors. Dependency of agriculture, drinking water and energy production on monsoon rainfall makes the monsoon rainfall the lifeline of its population. Junagadh region is the semi-arid region of Gujarat state.

The crop production, economy, life and property in the region largely depend on the amount of annual monsoon rainfall. Further, irregular, insufficient and unpredictable

nature of the monsoon rainfall resulted in unequal economic growth in the region. Knowledge of rainfall variability and its trend in the region is pre-requisite for effective decision-making system. Therefore, predicting the monsoon rainfall has a great socio-economic importance in such semi-arid region.

Indian monsoon rainfall forecasting begun with Sir Henry Blanford in 1886. Many researchers^{2,10,15,17} studied variability and trend of rainfall based on annual rainfall data over India. Das and Bhattacharya⁴ reported that there is no distinct change observed in trend of mean annual rainfall.

However, some researchers^{8,11} have stated an increasing trend in rainfall while others^{13,22} have shown significant rainfall decreasing trends over India. Extensive research works on rainfall prediction were carried out by Indian Meteorological Department (IMD) using advanced empirical models since 1932, albeit with limited success.¹⁹⁻²¹

IMD now uses a statistical prediction system to forecast Indian monsoon at longer time scale.¹⁶ However, recent studies^{1,6,7,14} indicate that the time-series analysis was extensively used for predicting environmental events. Komornik et al¹² predicted hydrological variables in the Czech Republic and Elmunim et al⁵ forecasted flooding in the Mississippi River in America by using time series models.

In these recent studies, researchers simulated the pattern of trend and seasonality present in existing time series data, then predicted the hydro-meteorological parameters using several predicting methods. Holt-Winters models are widely used for forecasting purposes. These models are based on the three smoothing parameters which exhibited level, trend and seasonality of the time series.

The main goal of this study was to check the possibility of using three different forms of the Holt-Winters time series models to predict rainfall for monsoon season for Junagadh region at monthly time scale.

Study Area

Geographical region of Junagadh (Gujarat-India) is encompassed between 20° 26' to 21° 24' North latitudes and 69° 24' to 71° 03' East longitudes. The location of the study area is presented in fig. 1. Junagadh District with population of about 27.4 lakh is the 7th most populous district of Gujarat state (India). Total geographical area of Junagadh district is 8831 km².

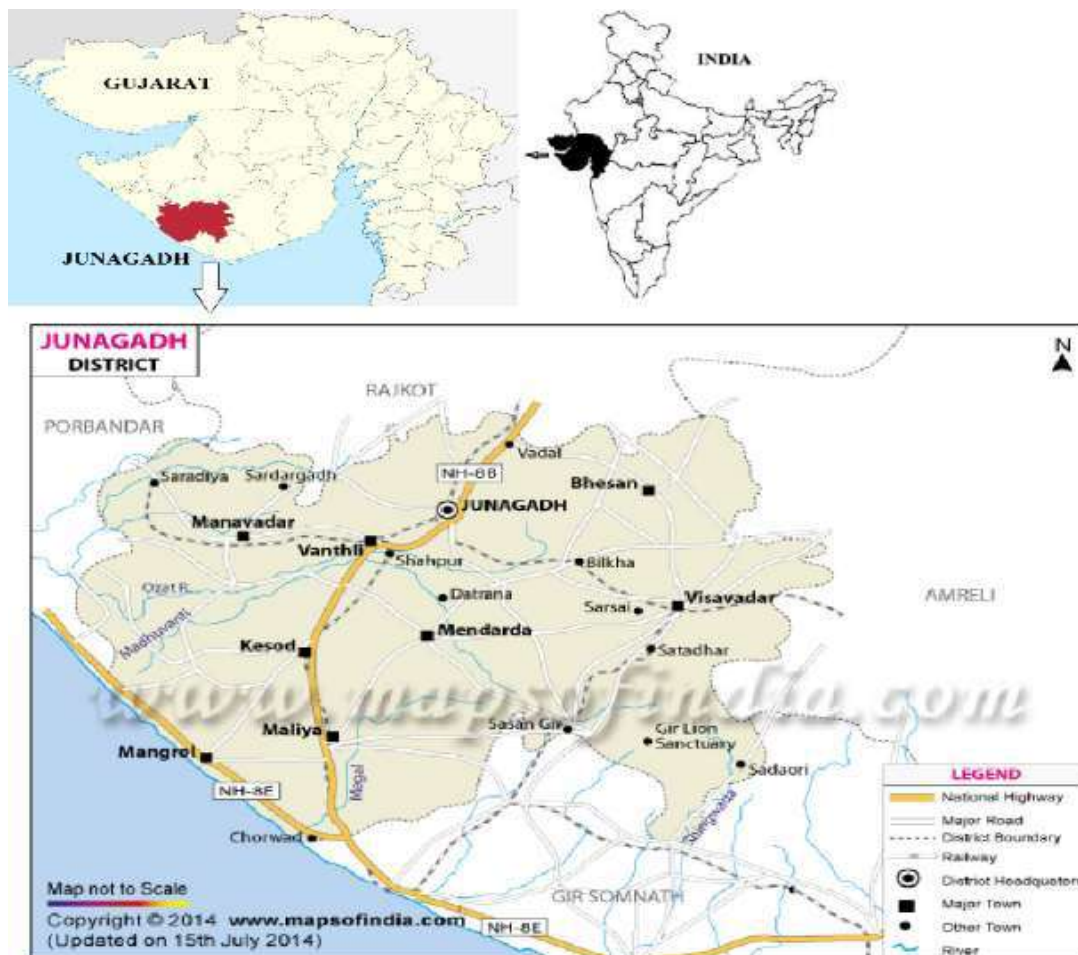


Fig. 1: Junagadh Region (Gujarat-India)

Population density of the district is 311 persons per km². It has mean annual rainfall of 900 mm (semi-arid region), mean maximum temperature 38.56°C and mean minimum temperature 14.82°C. Junagadh region is characterized by its inadequate rainfall pattern and lesser aquifer water storage capacity. Natural water conservation and water scarcity are the prime factors influencing water resource management, cropping pattern, irrigation scheduling and environmental assessment in this region.

Data and Methodology

Data: Monthly rainfall data of Junagadh (Gujarat-India) weather stations were collected from website (<http://www.jau.in/index.php/annual-weather-reports-weather-data>) of Junagadh Agro-meteorological Cell for 36 years (1984-2019) and was used in this study. Junagadh weather station has an altitude of 61m from mean sea level and lies between latitude of 21°31' N and longitude of 70°33' E. The weather parameters are recorded at the Agro-meteorological observatory, Department of Agronomy affiliated with Junagadh Agricultural University, Junagadh and published annually. The annual weather report is published by this department every year using observed weather data at the observatory at Instructional Farm.

The report carries daily, weekly, monthly, seasonal and

annual weather data of all parameters like air and soil temperatures, rainfall, relative humidity, wind speed, wind direction, bright sunshine hour and evaporation. This region has warm and dry summers and mild winter conditions. The highest mean annual wind speed was observed 11.86 km/h in the June while lowest mean annual wind speed was observed 2.85 km/h in the November. Microsoft Excel work sheets were used to organize the weekly and monthly rainfall records.

Methodology: Holt-Winters is a one of the leading time series models. Three basic features of the time series are a typical value (average), a slope (trend) over time, and a cyclical repeating pattern (seasonality). HW model can be classified into MHW and AHW models. These models differ in the nature of the seasonal component. When the size of the seasonal component has a direct relationship with trend level, then MHW model should be preferable. AHW model should be preferable when the seasonal component is directly proportional to trend level.

In this study, the MHW model and AHW model were applied to predict the monthly and seasonal rainfall for Junagadh region. The MHW and AHW models include separate equations for the level, trend, seasonal component, and forecast.

Holt-Winters Model (HW): The Holt-Winters is a time-series forecasting model used to predict the future values of the parameters based on existing attribute of the parameters. It is more suitable when the seasonal variability of the series is synchronized with the level of the series. This model permits calibration of its smoothing parameters based on weighted average of past and present observations. However, weights of more recent data keep larger than that of older one. This implies that more recent historical data is dominating in forecasting than the former observations. The forecast of time series was carried out using calibrated smoothing constants. Rainfall prediction in Holt-Winters estimates elements of base, trend and seasonal index. This model relies on four equations viz. level eq. (01), trend eq. (02), seasonal term eq. (03) and forecast eq. (04).¹⁸

Time series base (level) equation:

$$L_t = \alpha \left(\frac{Y_t}{S_{t-s}} \right) + (1 - \alpha)(L_{t-1} + T_{t-1}) \quad (1)$$

where Y_t = Observed rainfall data at time t , L_{t-1} = Smoothed base value for period $t-1$, T_{t-1} = Trend estimate for period $t-1$, S_{t-s} = Sum of seasonality components for s consecutive periods of time, length of the seasonality is s and α = Exponential smoothing parameter for the data ($0 < \alpha < 1$).

Time series trend equation:

$$T_t = \beta(L_t - L_{t-1}) + (1 - \beta)T_{t-1} \quad (2)$$

where β = Smoothing constant for the trend estimate ($0 < \beta < 1$).

Time series season equation:

$$S_t = \gamma \left(\frac{Y_t}{L_t} \right) + (1 - \gamma)S_{t-s} \quad (3)$$

where γ = Smoothing constant for the seasonality estimate ($0 < \gamma < 1$).

The rainfall for a stipulated time future period is estimated by using eq. (04):

$$P_{t+m} = (L_t + mT_t)S_{t-s+m} \quad (4)$$

To calculate the level, $L_s = (Y_1 + Y_2 + \dots + Y_n) / s$ (in our case $s = 3$) is used. For initializing the trend, we used $T_s = (Y_s + 1 - Y_1) / s$ while seasonal index was initialized by using $S_p = Y_p / L_s$, $p = 1, 2, \dots, s$.

Multiplicative Holt Winters Model (MHW): (MHW) model could be the best choice for forecasting when time series data exhibits multiplicative seasonality and the amplitude of the seasonal pattern is proportional to the average level. In the multiplicative model, the sum of the seasonality components is fixed as 1 for c consecutive periods of time. The original HW model initialized the trend

value with 0 while in MHW the same is initialized with 1.

Additive Holt Winters Model (AHW): Sometime, additive seasonality in the time series data detects and the seasonal variations through the series remain constant. In such situation, the AHW model is quite useful in forecasting. AHW fits well for a time series having independent amplitude of the seasonal pattern. It differs from (MHW) model, as MHW considered multiplicative seasonality in place of additive seasonality.

The forecasted value was computed by summing up the baseline, trend and seasonality components. The seasonal indices were divided and multiplied in MHW model, while they added and subtracted in the AHW model eq. (06). The seasonal indices were initialized by using $S_p = Y_p - L_s$, $p = 1, 2, 3, \dots, s$. The sum of the seasonality components for c consecutive periods of time is about c (not 1 as in the multiplicative model).

The forecast value at t periods is given by:

$$P_p = (F_t + mT_t) + S_{t-s+m} \quad (5)$$

where $P_p = P_{(t+m-i)}$ = Population of m steps ahead to get predictions decade ($i = 1, 2, \dots, m$).

The initial value of the trend is taken equal to zero.

$$S_t = \gamma(Y_t - T_t) + (1 - \gamma)S_{t-s} \quad (6)$$

The decision in choosing between additive method and multiplicative method depends on the time series characteristics as different methods will be suitable for different data and each method has its drawbacks.

Results and Discussion

The HW, MHW and AHW models have been applied to the monthly rainfall data from the year 1984 to 2019 to predict the monthly and seasonal rainfall for the Junagadh region in Gujarat state. The rainfall data was downloaded from website of Junagadh Agro-meteorological Cell of Junagadh Agricultural University. The data of the 33 years (1984-2016) were used for calibration of models parameters. The values of these parameters were kept constant during model validation and forecasting procedures. Models were validated by comparing data of the remaining three years (2017-2019) and forecasts were made for the next three years (2020-2022).

The model parameters require to be optimized as the meteorological time series are very dynamic in nature. Once the values of the smoothing constants i.e. α , β , and γ were calibrated, they took as it is. Microsoft Excel add-in program Solver was used for calibration and optimization of the smoothing constants (α , β , and γ). Microsoft Excel spread sheets was used for statistical analysis, evaluation and graph preparation.

Models were evaluated by employing refined Willmott's index (d_r) (Dimensionless statistic) and mean absolute error (MAE) (error index statistic) statistical criteria. d_r indicates the degree of similarity between the model predicted and observed deviations about the observed mean comparative to the sum of the magnitudes of the perfect model and observed deviations about the observed mean. $d_r=0.5$ implies that the sum of the error magnitudes is half of the sum of the perfect model deviation and observed deviation magnitudes.

$d_r=0.0$ indicates that the sum of the magnitudes of the errors and the sum of the perfect model deviation and observed deviation magnitudes are alike. $d_r=-0.5$ indicates that the sum of the error-magnitudes is twice the sum of the perfect model deviation and observed deviation magnitudes. The d_r is applied to quantify the degree to which values of measured rainfall are captured by the applied models. The range of d_r varies from -1.0 to 1.0. A d_r of 1.0 indicates perfect agreement between modeled value and observation and a d_r of -1.0 shows either lack of agreement between the modeled value and observation or insufficient variation in observations to adequately test the model. MAE is the most natural and unambiguous measure of average error scale.

It provides more robust measure of average model error, since it is not influenced by extreme outliers and estimates model error in the units of the variable. A higher MAE value indicates poor model performance and vice versa. MAE=0 indicates a perfect fit.

The data in table 1 shows how all the three models performed at annual time scale along with calibration values of smoothing constants. The results displayed in table 1 indicate that HW ($d_r=0.73$ and MAE =167.65 mm) and MHW ($d_r=0.73$ and MAE =166.74 mm) models performed better than AHW ($d_r=0.67$ and MAE =205.76 mm) model at annual time scale. This explicates that the annual data do not exhibit additive seasonality. The results are in good agreement with results obtained by Elmunim et al⁵ in their study for forecasting the ionosphere delay. Predicted annual

rainfalls for the year 2020 to 2022 by these models are presented in table 1.

The entire three models predicted highest rainfall (ranges from 1445.45 to 1485.03 mm) in the year 2022 which is more than 1.5 times the average annual rainfall (AAR) (1984-2019) of the region. It was revealed from fig. 2 that all the three methods overestimated annual rainfall in the year 2017 and 2018 while they underestimated the rainfall in the year 2019. The trend smoothing constant calibrated values ($\beta=1$) showed the significant increasing trend when the models were tested at annual time scale.

The models were also tested at monthly time scale for the monsoon season (June to September). It is clear from table 2 that HW during (July ($d_r=0.72$ and MAE =91.82 mm) and August ($d_r=0.42$ and MAE =128.43 mm)) and MHW during (July ($d_r=0.71$ and MAE =92.05 mm), and August ($d_r=0.46$ and MAE =119.52 mm)) models gave better results as compared to other months. On the contrary, the AHW model during June ($d_r=-0.13$ and MAE =138.07 mm) and September ($d_r=0.41$ and MAE =331.17 mm) provided better results as compared to the HW and MHW models. Negative d_r values in the June confirmed poor performance of the model which may be due to modestly very little rainfall (only 7.40 mm) in the June, 2018. It was also observed that HW and MHW models performed better when average monthly rainfall (AMR) exceed 25% of the AAR.

AMR at beginning (June) and ending (September) was observed usually less and hence AHW model did well. The time series data displayed multiplicative seasonality in the month of July and August where nearly 70% of the monsoon season rainfall occurred. It is clear from figs. 3 to 6 that in the year 2017, all the three models overestimated monthly rainfall similarly in the year 2018 except in month of July. In the year 2019 the models overestimated monthly rainfall in June and July months while underestimated the same in August and September months. The results depicted that no single model performed consistently.

Table 1
Performance of HW, MHW and AHW models at annual time scale

Years	Observed Rainfall in mm	Predicted Rainfall in mm		
		HW	MHW	AHW
2017	801.80	1162.88	1159.44	1129.65
2018	789.40	845.60	842.76	956.80
2019	1492.20	1406.54	1402.98	1370.18
2020	AAR = 916.61(1984-2019)	1073.24	1070.51	1089.78
2021		854.55	852.12	960.11
2022		1485.03	1481.64	1445.45
Statistical Measures	d_r	0.73	0.73	0.67
	MAE (mm)	167.65	166.74	205.76
Smoothing Constants	α	0.0030	0.0026	0.0065
	β	1.0000	1.0000	1.0000
	γ	0.3361	0.3364	0.2361

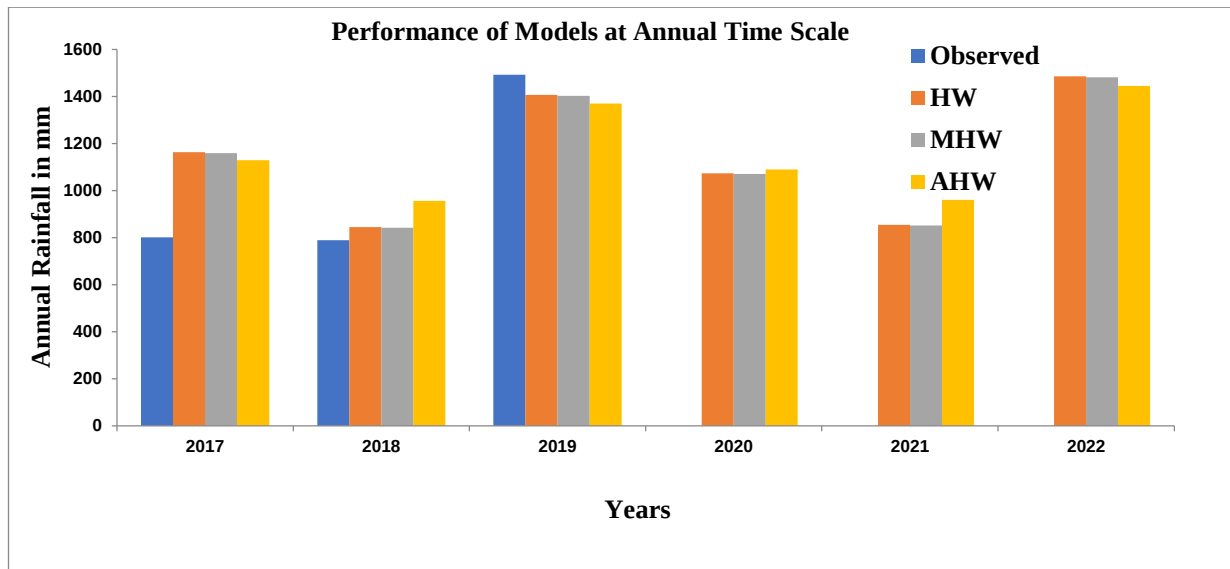


Fig. 2: Performance of HW, MHW and AHW models at Annual Time Scale

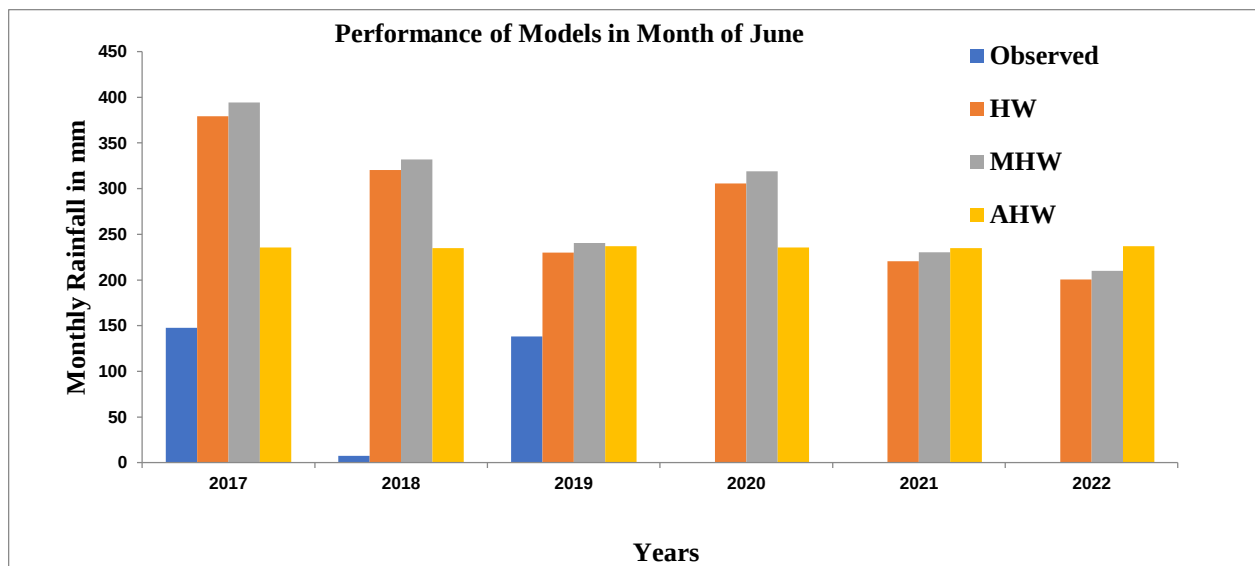


Fig. 3: Performance of HW, MHW and AHW models in Month of June

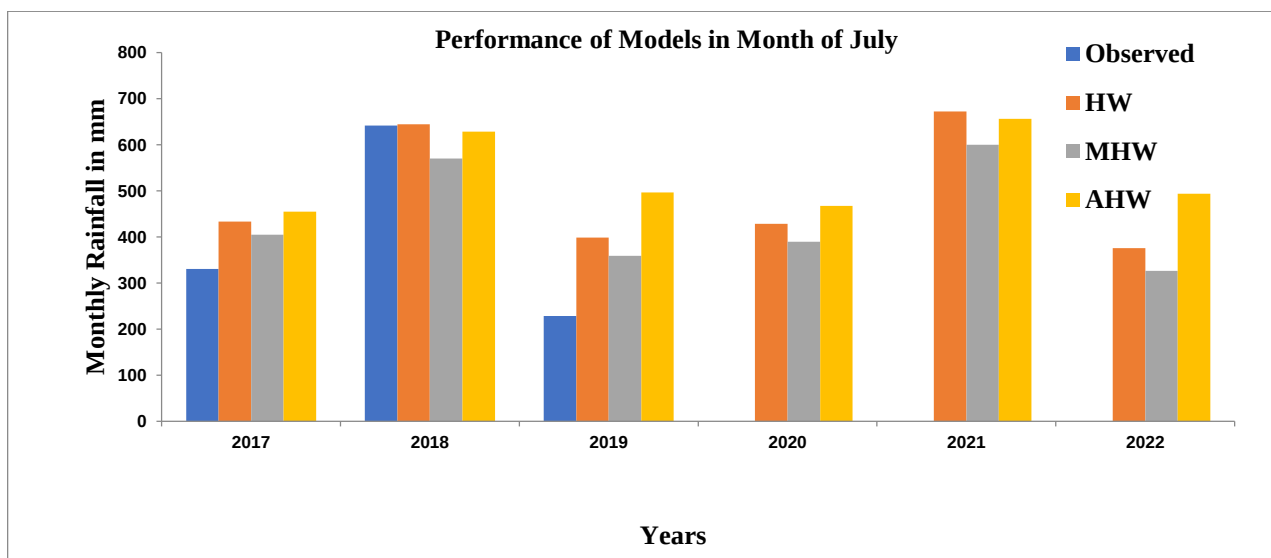


Fig. 4: Performance of HW, MHW and AHW models in Month of July

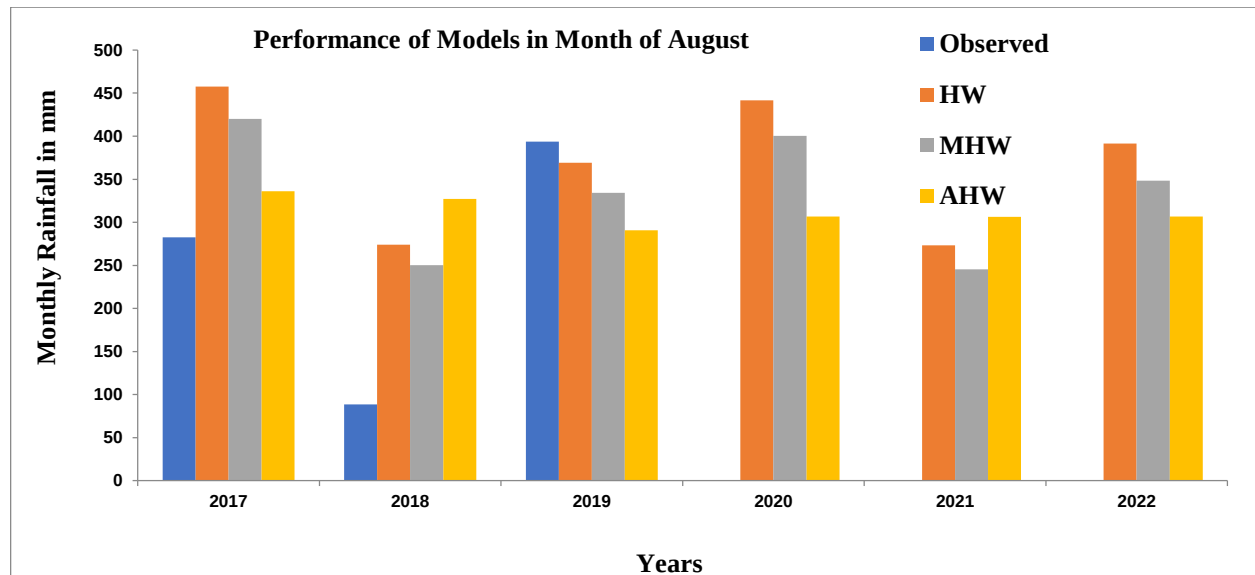


Fig. 5: Performance of HW, MHW and AHW models in Month of August

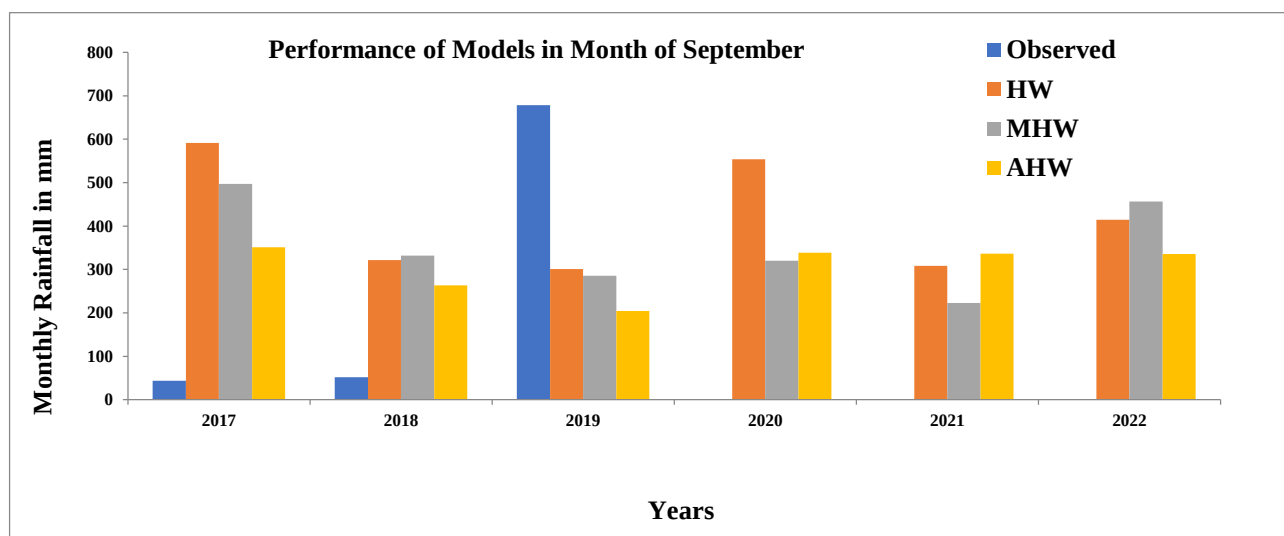


Fig. 6: Performance of HW, MHW and AHW models in Month of September

This supported the finding of Hassani et al⁹. The data in table 2 depicted that the values of smoothing constant α were nearly zero if AMR is less which indicates no significant contribution of average (level) value of series in forecasting the monthly rainfall. It was also noted that when AMR is more, HW and MHW performed better than AHW.

The rainfall prediction is a tricky job and there will not be one correct model for all datasets. The rainfall is highly complex natural phenomenon influenced by many meteorological factors and hence does not follow any deterministic trend, pattern and seasonality. However, all the three models explicitly performed well at annual time scale indicating that prediction of total monsoon rainfall is more accurate than monthly rainfall prediction.

Conclusion

The present study is conducted to investigate ability of the HW, MHW and AHW models in predicting the monsoon

rainfall for Junagadh (Gujarat-India) region. The models performances was evaluated for monthly as well as for annual time scales by using two statistical measures d_r and MAE. The results of annual time scale rainfall prediction indicated that the HW and MHW performed better than the AHW model. The trend smoothing constant calibrated value ($\beta=1$) pointed out a significant increasing trend in rainfall time series when the models tested at annual time scale.

In this study, the HW and MHW models showed better accuracy than that of AHW model when tested at monthly time scale. It was observed from the results that HW and MHW models performed pre-eminent while AHW model seem a bit truthful for July and August months. This study also indicates that HW and MHW models make reliable prediction for monsoon rainfall. Therefore, from a practical point of view, the HW and MHW models are recommended for predicting monsoon rainfall in the similar hydro-meteorological region.

Table 2A
Performance of HW, MHW and AHW models at monthly time scale

Years	June				July			
		Predicted Rainfall in mm				Predicted Rainfall in mm		
	Observed Rainfall in mm	HW	MHW	AHW	Observed Rainfall in mm	HW	MHW	AHW
2017	147.80	379.37	394.15	235.54	330.50	433.09	404.48	454.80
2018	7.40	320.49	332.07	234.84	641.90	644.13	570.41	628.71
2019	138.10	229.91	240.35	237.12	228.20	398.83	358.89	496.83
2020	AMR206.31	305.52	318.96	235.54	AMR 376.67	428.21	389.29	467.33
2021		220.64	230.40	234.84		672.48	600.25	656.39
2022		200.63	210.17	237.12		375.76	326.39	493.68
Statistical Measures	d_r	-0.43	-0.46	-0.13	d_r	0.72	0.71	0.58
	MAE (mm)	212.16	224.42	138.07	MAE (mm)	91.82	92.05	135.38
Smoothing Constants	α	0.0000	0.0000	0.0000	α	0.0010	0.0000	0.0057
	β	0.0091	0.0091	0.0050	β	1.0000	0.0000	1.0000
	γ	0.3189	0.3210	0.0000	γ	0.2233	0.2854	0.1048

Table 2B
Performance of HW, MHW and AHW models at monthly time scale

Years	August				September			
		Predicted Rainfall in mm				Predicted Rainfall in mm		
	Observed Rainfall in mm	HW	MHW	AHW	Observed Rainfall in mm	HW	MHW	AHW
2017	282.60	457.61	420.16	335.94	43.50	591.77	496.93	350.94
2018	88.60	274.18	250.10	327.21	51.50	321.33	331.68	263.55
2019	393.80	369.10	334.30	290.60	678.70	301.15	285.46	204.69
2020	AMR 243.43	441.55	400.19	306.74	AMR 205.18	554.06	320.21	338.35
2021		273.28	245.41	306.27		308.49	223.08	336.10
2022		391.52	348.40	306.57		414.38	456.85	335.96
Statistical Measures	d_r	0.42	0.46	0.41	d_r	0.29	0.33	0.41
	MAE (mm)	128.43	119.52	131.72	MAE (mm)	398.55	375.62	331.17
Smoothing Constants	α	0.0824	0.0888	0.1547	α	0.0032	0.0000	0.2769
	β	0.0482	0.0114	0.0000	β	1.0000	0.0088	0.0000
	γ	0.0000	0.0000	0.0000	γ	0.1638	0.4041	0.0000

Acknowledgement

The authors are grateful to Junagadh Agro Meteorological Cell of Junagadh Agricultural University, Junagadh (Gujarat) for providing all necessary meteorological data.

Future scope

Holt-Winters forecasting is simple but powerful method. It can adequately address the complicated seasonal patterns and the effects of slope. This work will be useful to predict monsoon rainfall in the similar hydro-meteorological region.

References

1. Bernal J.L., Cummins S. and Gasparrini A., Interrupted time series regression for the evaluation of public health interventions:

a tutorial, *International Journal of Epidemiology*, **46(1)**, 348 (2017)

2. Chand R., Singh U.P., Singh V.P., Siddique L.A. and Kore P.A., Analysis of weekly rainfall of different period during rainy season over Safdarjung airport of Delhi for 20th century - A study on trend, decile and decadal analysis, *Mausam*, **62(2)**, 197-204 (2011)

3. Damle C. and Yalcin A., Flood prediction using timeseries data mining, *Journal of Hydrology*, **333(2)**, 305 (2007)

4. Das J. and Bhattacharya S.K., Trend analysis of long-term climatic parameters in Dinhat of Koch Bihar district, West Bengal, *Spatial Information Research*, DOI: 10.1007/s41324-018-0173-3 (2018)

5. Elmunim N.A., Abdullah M., Hasbi A.M. and Bahari S.A., The Comparison of Statistical Holt-Winter Models for Forecasting the Ionospheric Delay Using GPS Observation, *Indian Journal of Radio and Space Physics*, **44**, 28-34 (2015)
6. Faust K., Lahti L., Gonze D., De Vos W.M. and Raes J., Metagenomics meets time series analysis: unraveling microbial community dynamics. *Current Opinion in Microbiology*, **25**, 56-66, DOI: 10.1016/j.mib.2015.04.004 (2015)
7. Gocheva-Ilieva S.G., Ivanov A.V., Voynikova D.S. and Boyadzhiev D.T., Time series analysis and forecasting for air pollution in small urban area: an SARIMA and factor analysis approach, *Stochastic Environmental Research and Risk Assessment*, **28(4)**, 1045 (2014)
8. Goswami B.N., Venugopal V., Sengupta D., Madhusoodanam M.S. and Xavier P.K., Increasing trends of extreme rain events over India in a warming environment, *Current Science*, **314**, 1442-1445 (2006)
9. Hassani H., Silva E.S., Antonakakis N., Filis G. and Gupta R., Forecasting accuracy evaluation of tourist arrivals, *Annals of Tourism Research*, **63**, 112-127 (2017)
10. Jain S.K., Kumar V. and Saharia M., Analysis of rainfall and temperature trends in northeast India, *International Journal of Climatology*, DOI: 10.1002/joc.3483 (2012)
11. Jaswal A.K., Roa P.C.S. and Singh V., Climatology and trends of summer high temperature days in India during 1969–2013, *J Earth Syst Sci*, DOI: 10.1007/s12040-014-0535-8 (2015)
12. Komornik J., Komornikova M., Mesiar R., Szokeova D. and Szolgay J., Comparison of forecasting performance of nonlinear models of hydrological time series, *Physics and Chemistry of the Earth*, **31(18)**, 1127-1145, DOI: 10.1016/j.pce.2006.05.006 (2006)
13. Meshram S.G., Singh S.K., Meshram C., Deo R.C. and Ambade B., Statistical evaluation of rainfall time series in concurrence with agriculture and water resource of Ken River Basin, Central India (1901-2010), *Theoretical and Applied Climatology*, DOI: 10.1007/s00704-0172335-y (2017)
14. Montgomery D.C., Jennings C.L. and Kulahci M., Introduction to time series analysis and forecasting, John Wiley & Sons (2015)
15. Pingale S., Khare D., Jat M. and Adamowski J., Spatial and temporal trends of mean and extreme rainfall and temperature for the 33 urban centres of the arid and semi-arid state of Rajasthan, India, *Atmospheric Research*, **138**, 73-90 (2014)
16. Rajeevan M. and Nanjundiah R.S., Coupled model simulations of twentieth century climate of the Indian summer monsoon, In Platinum Jubilee Special Volume of the Indian Academy of Sciences, Indian Academy of Sciences, Bangalore, 537–568 (2009)
17. Rani Arthi, Manikandan and Maragathan N., Trend analysis of rainfall and frequency of rainy days over Coimbatore, *Mausam*, **65(3)**, 379-38 (2014)
18. Rossi M. and Brunelli, Forecasting data centers power consumption with the Holt-Wintersmethod, In Environmental, Energy and Structural Monitoring Systems (EESMS), IEEE Workshop, 210-214 (2015)
19. Gadgil Sulochana and Gadgil Siddharth, The Indian monsoon GDP and agriculture, *Econ Politwky*, **41(47)**, 4887–4895 (2006)
20. Gadgil Sulochana and Srinivasan J., Foretelling the monsoon: Freedom and responsibility, *Current Sci.*, **89**, 238–239 (2005)
21. Wang B., Kang I.S. and Lee Y.J., Ensemble simulations of Asian–Australian monsoon variability during 1997/1998 El Niño by 11 AGCMs, *J. Climate*, **17(4)**, 803–818 (2004)
22. Yin Y., Xu C.Y., Chen H., Li L., Xu H., Li H. and Jain S.K., Trend and concentration characteristics of precipitation and related climatic telecommunications from 1982 to 2010 in the Beas River basin, India, *Global and Planetary Change*, **145**, 116-129 (2016).

(Received 17th January 2021, accepted 31st March 2021)