Hybrid ETS-STLM Model for Forecasting IIPs of India

Dipankar Das[1], Awanish Kumar Tripathi[2], Ayushi Shah[3],Samarth Mehta[4]
Department of BCA[1,2,3,4], The Heritage Academy, Kolkata, West Bengal, India[1,2,3,4]
Email: dipankr.das@gmail.com

Abstract - The Index of Industrial Production (IIP) is a univariate time series and one of the most important indicators of economy. Development of forecasting model of the IIP with relatively good accuracy is an interesting and a challenging task. The present work aims to develop forecasting models for different IIPs of India. The authors used hybrid technique to develop the forecasting models and identified that hybrid ETS + STLM model performed better than the base models. We had noted that the MAPE of the hybrid ETS + STLM model for all the 23 IIPs under study was less than 10. For 19 IIPs out of 23 IIPs under study the RMSE of the said hybrid model is less than 10. Therefore we may conclude that the said model may be considered as one of the candidates for making forecasting of the dataset under study.

Index Terms- ARIMA; ETS; Holt – Winters'; STLM; Forecasting; Time Series Analysis; Hybrid Modeling; IIPs of India

I. INTRODUCTION

The Index of Industrial Production is an important Macroeconomic indicator for the country. The data series is a univariate time series in nature, collected over a span of time. This study is a continuation of an existing study where Multi-Layer Perceptrons were used to predict values of the given IIPs of India over a span of 12 months [20]. This study then uses base case time series models namely ETS, ARIMA, STLM, and Holt-Winters and applies them to forward forecasting of 12 months. Once these predictions are achieved, we use the "forecastHybrid" package of R to implement certain hybrid models and check if better results are obtained in comparison to our base case models.

The present study aims at:

Accurately computing a hybrid model using the "forecastHybrid" package in R to achieve a 12 month forward forecasting having an error rate lower than our base time series models (ARIMA, ETS, Holt-Winters, STLM)

In this paper the authors have discussed very briefly about Index of Industrial Productions of India, Time Series models such as ETS, ARIMA, STLM and Holt-Winters and other Hybrid models and stated the research question in the Introduction section (Section I), gave literature review in the Related Work section (Section II), expressed clearly the objectives in the Objectives of the Study section (Section III), briefly mentioned the methodology of the research in the Methodology section (Section IV), elaborated the results of the research with suitable diagrams and table in the Data Analysis & Findings section (Section V) and lastly in the Conclusion section (Section VI) appraised about the final findings of this empirical analysis.

II. RELATED WORKS

Time Series Forecasting is a very important domain in the field of Data Science. Use case of models like ARIMA[8][9][10][11], ETS[12][13], STLM[14][15][16] and Holt Winters[17][18][19] as base case models for forecasting has been done since many years. Application of Hybrid Models and chaining models in the realm of forecasting had also being studied [1][2][3][4]. The usual chaining models are ARIMA + ANN [5][6][7]. In this paper, we develop a hybrid model using ETS and STLM for this purpose.

III. OBJECTIVES OF THE STUDY

- To develop base forecasting time series models of the given 23 IIPs of India.
- To create hybrid models using ETS and STLM to predict these 23 IIPs of India.
- To compare these Hybrid Models with the base models and estimating those models that have a lower error rate than the base models.
- To visualize the forecasting results obtained from the developed hybrid model.

IV. METHODOLOGY

The data (i.e. the IIPs of India) under study was collected from "Open Government Data (OGD) Platform India" [21]. The data ranged from April, 2012 to March, 2017 i.e. sixty (60) months. The data under study was monthly in nature.

The IIPs of twenty three (23) different items of India was analyzed in this current work which is listed below along with their coded names. Manufacture of

(i) "Food products" coded as fprod,
(ii) "Beverages" coded as bprod,
(iii) "Tobacco products" coded as tpprod,
(iv) "Textiles" coded as tprod,
(v) "Wearing apparel" coded as wprod,
(vi) "Leather and related products" coded as lprod,
(vii) "wood and products of wood and cork, except furniture; manufacture of articles

of straw and plaiting materials" coded as woprod,

(viii) "Paper and paper products" coded as paprod,

(ix) "Coke and refined petroleum products" coded as cprod,

(x) "Chemicals and chemical products" coded as chprod,

(xi) "Pharmaceuticals, medicinal chemical and botanical products" coded as phprod,

(xii) "Rubber and plastics products" coded as ruprod,

(xiii) "Other non-metallic mineral products" coded as nomprod,

(xiv) "Basic metals" coded as bmprod,

(xv) "Fabricated metal products, except machinery and equipment" coded as fmprod,

(xvi) "Computer, electronic and optical products" coded as ceprod,

(xvii) "Electrical equipment" coded as eeprod,

(xviii) "Machinery and equipment n.e.c." coded as meprod,

(xix) "Motor vehicles, trailers and semi-trailers" coded as mvtprod,

(xx) "Other transport equipment" coded as oteprod,

(xxi) "Furniture" coded as furprod.

The other two IIPs are

(xxii) "Printing and reproduction of recorded media" coded as remedia and

(xxiii) "Other manufacturing" coded as othrprod.

The data was divided into two parts – training set (forty eight months) and test set (twelve months).

In the given study we are using the "forecastHybrid" [22] package in R, through which we are developing the hybrid models to forecast the IIPs of India in the form of a forward 12-month progression. The results so obtained are then compared to the results of our base models.

The Accuracy measures used in this process were:

- Root Mean Square Error (RMSE)
- Mean Absolute Percentage Error (MAPE)

V. DATA ANALYSIS AND FINDINGS

The RMSE and the MAPE results of the given 23 IIPs are listed in the tables below:

TABLE I. MAPE of 23 IIPs

Sectors	Model	ARIMA	ETS	Holt-Winte	ETS+STLM
fprod	MAPE	6.119561	7.482446	3.361687	3.251984
bprod	MAPE	9.686169	11.28237	5.10097	4.379443
tprod	MAPE	7.523725	7.704695	7.597761	6.917166
tpprod	MAPE	3.196924	3.24447	1.505398	1.43658
wprod	MAPE	8.804443	8.783012	5.304818	5.117395
lprod	MAPE	5.619017	5.726254	3.029153	2.595068
woprod	MAPE	5.699247	5.699661	3.227725	3.221853
cprod	MAPE	3.203025	3.586542	2.656124	2.690378
paprod	MAPE	6.301513	6.552289	5.045555	5.111293
chprod	MAPE	4.033618	4.277115	2.585439	2.494895
phprod	MAPE	3.424898	3.4723	1.85036	2.961262
ruprod	MAPE	7.050838	7.500941	6.119801	5.761168
nomprod	MAPE	3.822134	4.16765	2.901571	2.58447
bmprod	MAPE	4.464501	5.132235	2.089439	1.87319
fmprod	MAPE	3.289489	3.398066	2.464354	2.307426
ceprod	MAPE	7.942742	8.025675	5.453058	4.84607
eeprod	MAPE	9.285578	8.859928	3.717637	3.690473
meprod	MAPE	8.837291	8.720952	6.655167	6.859429
mvtprod	MAPE	7.228518	7.818504	3.30101	3.199871
oteprod	MAPE	4.511539	4.508394	3.172751	3.492245
furprod	MAPE	4.761209	4.822081	3.434209	2.953035
remedia	MAPE	9.232245	9.31049	8.780519	7.361337
othrprod	MAPE	11.1002	11.46543	8.636395	9.337787

TABLE II. RMSE of 23 IIPs

Sectors	Model	ARIMA	ETS	Holt-Winte	ETS+STLM
fprod	RMSE	8.593407	10.96834	4.544873	4.321108
bprod	RMSE	13.39609	16.31895	6.957454	6.155892
tprod	RMSE	12.00083	12.44125	11.77571	11.25936
tpprod	RMSE	4.640177	4.638818	2.339457	2.140084
wprod	RMSE	14.61162	14.70082	8.068922	7.663949
lprod	RMSE	8.4001	8.189332	4.551829	3.995015
woprod	RMSE	6.327185	6.327501	3.773319	3.705833
cprod	RMSE	4.498263	4.934439	3.580286	3.763325
paprod	RMSE	7.7943	8.678986	6.542545	7.000345
chprod	RMSE	5.245363	5.510047	3.341174	3.341199
phprod	RMSE	4.567181	4.616492	2.419105	4.062566
ruprod	RMSE	10.33143	10.62898	8.784476	8.159177
nomprod	RMSE	5.631654	6.065271	4.548523	4.321553
bmprod	RMSE	5.748746	6.542965	2.942023	2.400959
fmprod	RMSE	5.087361	5.216542	3.724882	3.288449
ceprod	RMSE	10.59587	10.58852	6.613291	6.051624
eeprod	RMSE	14.754	13.37835	5.75764	5.722615
meprod	RMSE	14.01368	14.00462	9.945352	10.16848
mvtprod	RMSE	10.98346	11.72591	4.20811	4.000401
oteprod	RMSE	5.809671	5.802615	4.101294	4.492936
furprod	RMSE	6.726996	6.724839	4.169798	4.064508
remedia	RMSE	16.23879	16.23785	14.44815	12.65021

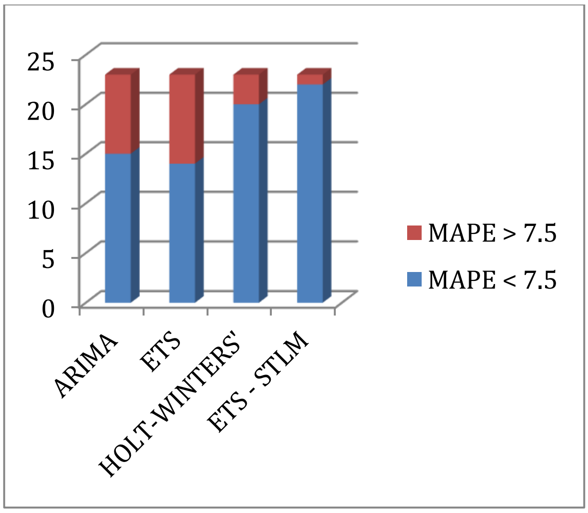

Fig. 1: Comparison between base models and ETS – STLM using MAPE

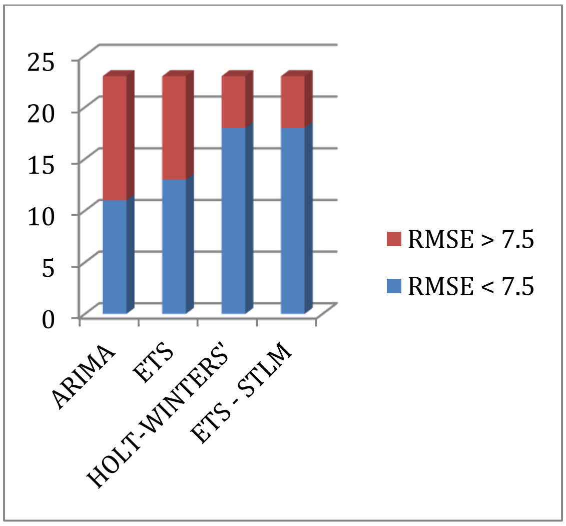

Fig. 2: Comparison between base models and ETS – STLM using RMSE

The Visualization of the 23 IIP models using the ETS + STLM hybrid model is listed below. On further investigation, it was found out that ETS + STLM model is giving us the best results and the error rates obtained through this hybrid model are lower than our base models. The following figures show the ETS + STLM hybrid models for all the IIPs under study:

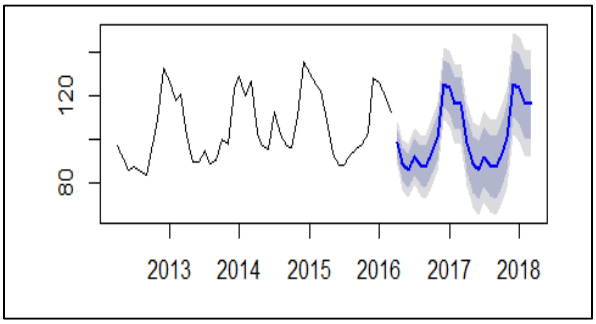

Fig. 3: Forecast of fprod

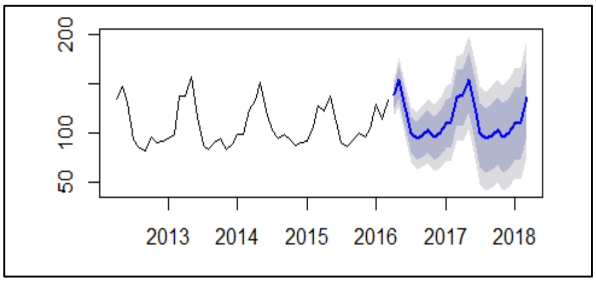

Fig. 4: Forecast of bprod

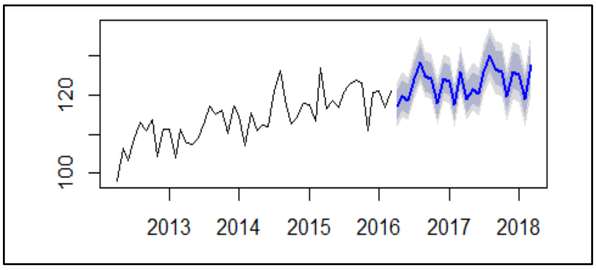

Fig. 5: Forecast of tprod

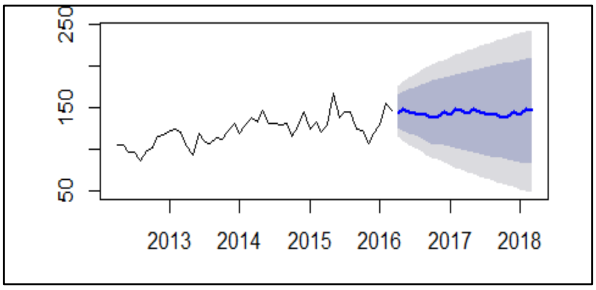

Fig. 6: Forecast of tpprod

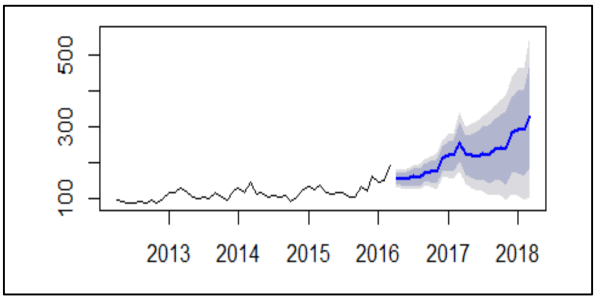

Fig. 7: Forecast of wprod

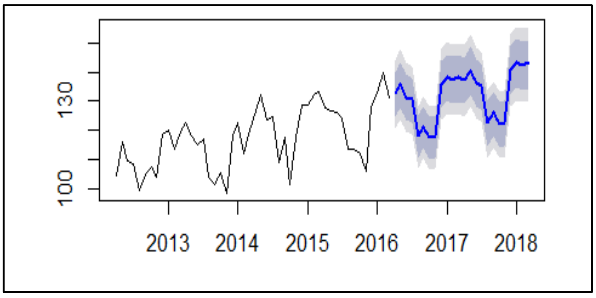

Fig. 8: Forecast of lprod

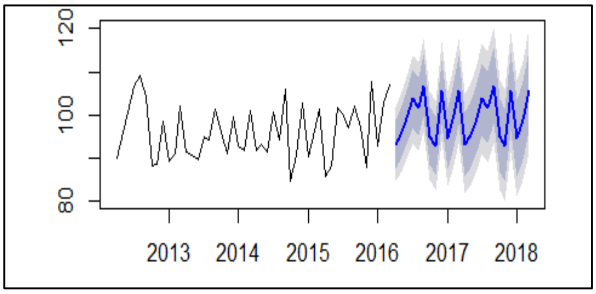

Fig. 9: Forecast of woprod

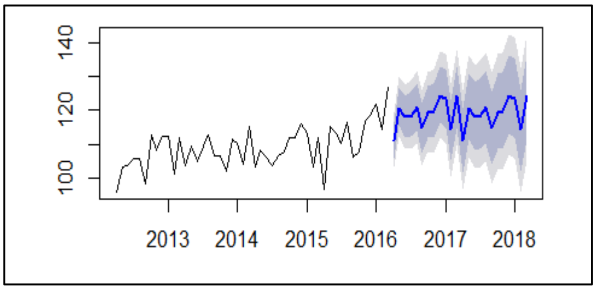

Fig. 10: Forecast of cprod

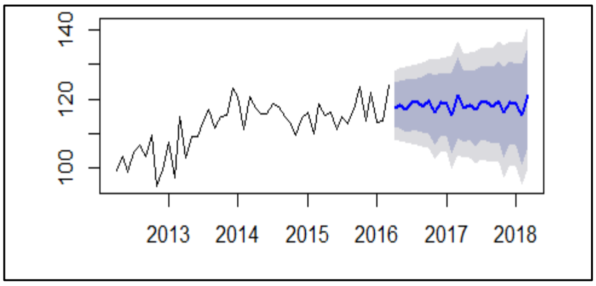

Fig. 11: Forecast of paprod

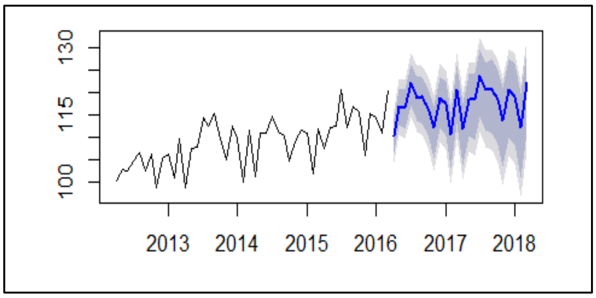

Fig. 12: Forecast of chprod

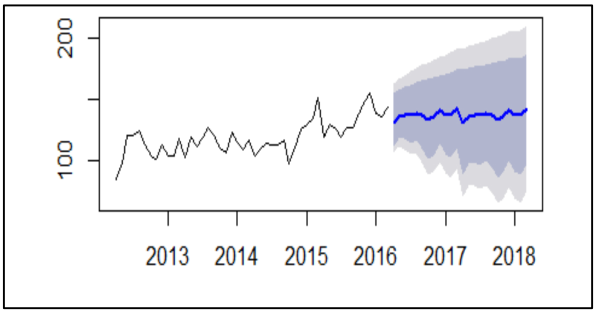

Fig. 13: Forecast of phprod

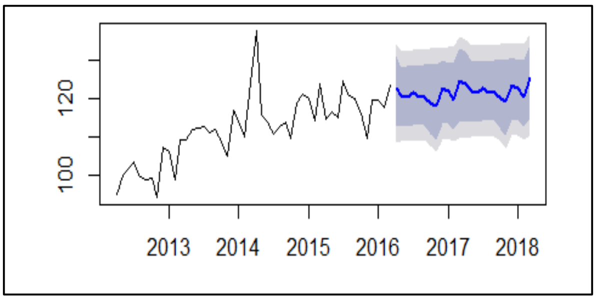

Fig. 14: Forecast of ruprod

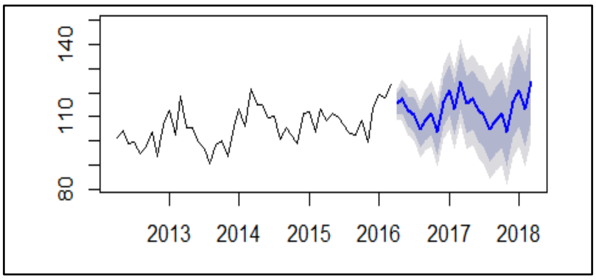

Fig. 15: Forecast of nomprod

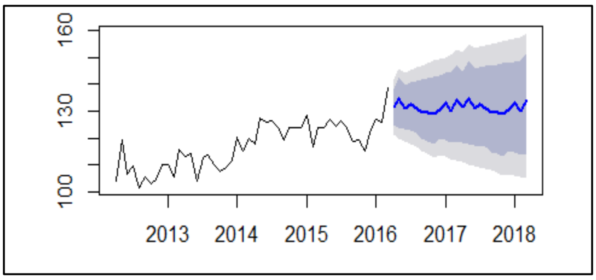

Fig. 16: Forecast of bmprod

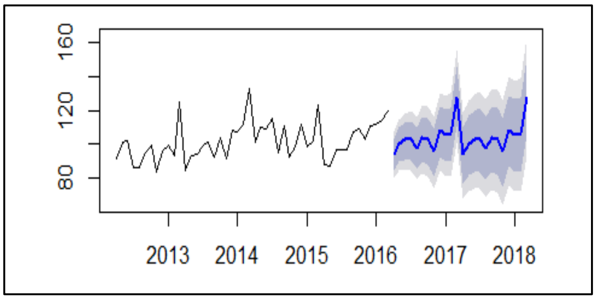

Fig. 17: Forecast of fmprod

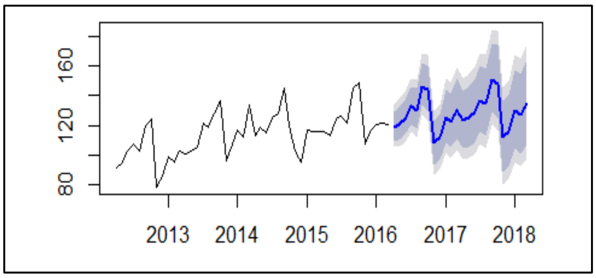

Fig. 18: Forecast of ceprod

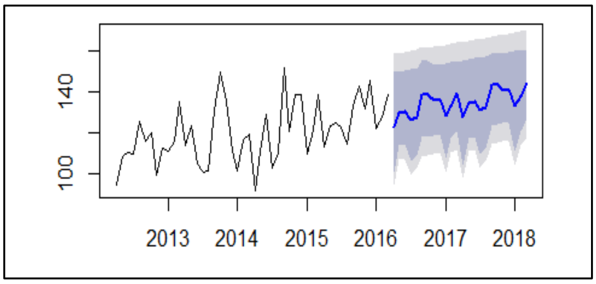

Fig. 19: Forecast of eeprod

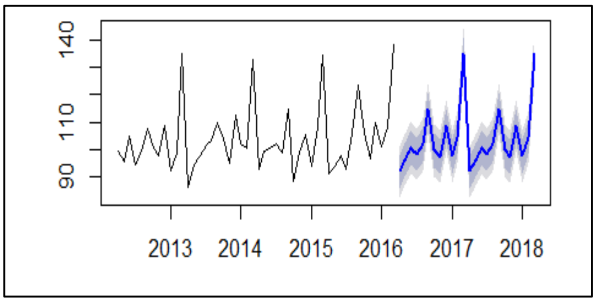

Fig. 20: Forecast of meprod

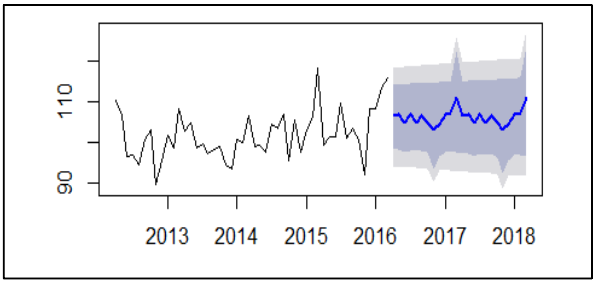

Fig. 21: Forecast of mvtprod

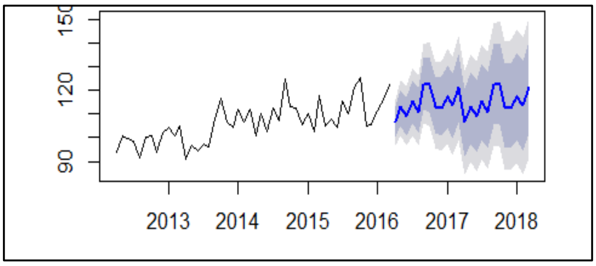

Fig. 22: Forecast of oteprod

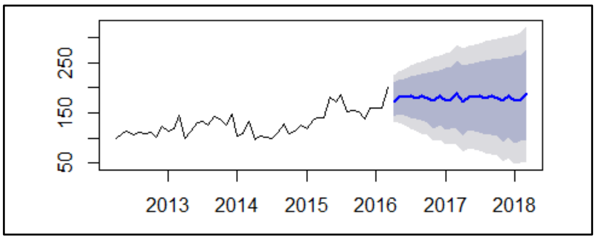

Fig. 23: Forecast of furprod

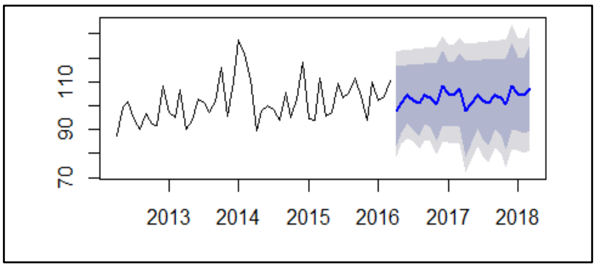

Fig. 24: Forecast of remedia

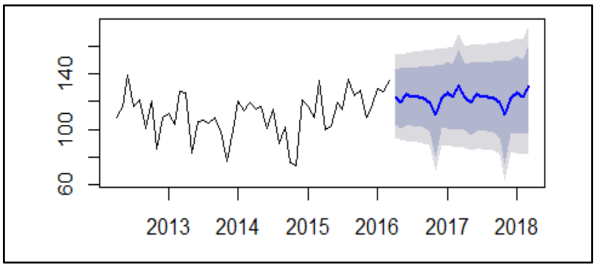

Fig. 25: Forecast of othrprod

VI. CONCLUSION

Above we have constructed a table showing the error values that are obtained from our base models and the error values that are obtained from our Hybrid model for 12-months forward forecasting. Further analysis shows that our hybrid model consisting of Exponential Smoothing and Short Term Long Memory is giving us significant results, better than our selected base models.

It may be noted that from the 23 IIP models we took into consideration, all of them i.e. 23 have performed exceptionally well giving the MAPE of less than 10 under the hybrid ETS + STLM model. We had also noted that for 19 IIPs the RMSE values obtained from the forecasted and observed values are less than 10. Therefore, we may conclude that the hybrid ETS + STLM model is one of the suitable candidates for making forecasting of the IIPs under study.

REFERENCES

[1] Xiao, Y., Xiao, J., & Wang, S. (2012). A Hybrid Forecasting Model for Non-Stationary Time Series. *International Journal of Knowledge and Systems Science*,*3*(2), 67-82. doi:10.4018/jkss.2012040105

[2] Suhartono, Rahayu, S., Prastyo, D., Wijayanti, D., & Juliyanto. (2017). Hybrid model for forecasting time series with trend, seasonal and salendar variation patterns. *Journal of Physics: Conference Series*,*890*, 012160. doi:10.1088/1742-6596/890/1/

[3] Khashei, M., & Bijari, M. (2012). A new class of hybrid models for time series forecasting. *Expert Systems with Applications*,*39*(4), 4344-4357. doi:10.1016/j.eswa.2011.09.157

[4] Khashei, M., & Bijari, M. (2012). A new class of hybrid models for time series forecasting. *Expert Systems with Applications*,*39*(4), 4344-4357. doi:10.1016/j.eswa.2011.09.157

[5] Khandelwal, I., Adhikari, R., & Verma, G. (2015). Time Series Forecasting Using Hybrid ARIMA and ANN Models Based on DWT Decomposition. *Procedia Computer Science*,*48*, 173-179. doi:10.1016/j.procs.2015.04.167

[6] Xiao, Y., Xiao, J., & Wang, S. (2012). A Hybrid Forecasting Model for Non-Stationary Time Series. *International Journal of Knowledge and Systems Science*,*3*(2), 67-82. doi:10.4018/jkss.2012040105

[7] Zhang, G. (2003). Time series forecasting using a hybrid ARIMA and neural network model. *Neurocomputing,50*, 159-175. doi:10.1016/s0925-2312(01)00702-0

[8] Mondal, P., Shit, L., & Goswami, S. (2014). Study of Effectiveness of Time Series Modeling (Arima) in Forecasting Stock Prices. *International Journal of Computer Science, Engineering and Applications,4*(2), 13-29. doi:10.5121/ijcsea.2014.4202

[9] Etuk, E. H., & Mohamed, T. M. (2014). Time Series Analysis of Monthly Rainfall data for the Gadaref rainfall station, Sudan, by Sarima Methods. *International Journal of Scientific Research in Knowledge,*320-327. doi:10.12983/ijsrk-2014-p0320-0327

[10] Z. Asha, B. Arputhamary(2016). A survey on ARIMA Forecasting using Time Series Model. *International Journal of Computer Science and Mobile Computing, Vol 5 Issue 8.*

[11] Bandyopadhyay, G. (2016). Gold Price Forecasting Using ARIMA Model. *Journal of Advanced Management Science,*117-121. doi:10.12720/joams.4.2.117-121

[12] Jofipasi, C. A., Miftahuddin, & Hizir. (2018). Selection for the best ETS (error, trend, seasonal) model to forecast weather in the Aceh Besar District. *IOP Conference Series: Materials Science and Engineering,352*, 012055. doi:10.1088/1757-899x/352/1/012055

[13] Jain, G., & Mallick, B. (2017). A Study of Time Series Models ARIMA and ETS. *SSRN Electronic Journal*. doi:10.2139/ssrn.2898968

[14] Majumdar, K., Vivekanand, S., Huffman, C., Matthews, K., Ngai, T., Chen, C. H., . . . Kirsch, P. D. (2013). STLM: A Sidewall TLM Structure for Accurate Extraction of Ultralow Specific Contact Resistivity. *IEEE Electron Device Letters,34*(9), 1082-1084. doi:10.1109/led.2013.2271032

[15] Pace, R., Barry, R., Gilley, O. W., & Sirmans, C. (2000). A method for spatial–temporal forecasting with an application to real estate prices. *International Journal of Forecasting,16*(2), 229-246. doi:10.1016/s0169-2070(99)00047-3

[16] Guido Guizi, R. Revetria, Claudio Silvestri(2015), A comparison of Forecast models to predict weather *Advances in Energy and Environmental Sciences and Engineering*

[17] Paul Godwin, The Holt-Winters Approach to Exponential Smoothing: 50 Years Old and Going Strong, *Hot New Research*

[18] Rahman, Md. Habibur & Salma, Umma & Hossain, Md & Khan, Md Tareq Ferdous. (2016). Revenue Forecasting using Holt–Winters Exponential Smoothing. *Research & Reviews: Journal of Statistics. 5. 19-25.*

[19] Pongdatu, G. A., & Putra, Y. H. (2018). Seasonal Time Series Forecasting using SARIMA and Holt Winter's Exponential Smoothing. *IOP Conference Series: Materials Science and Engineering,407*, 012153. doi:10.1088/1757-899x/407/1/012153

[20] Das, D., Tripathi, A. K., Shah, A., & Mehta, S. (2018). Application of Multilayer Perceptron for Forecasting of Selected IIPs of India An Empirical Analysis. *International Journal of Computer Sciences and Engineering,6*(11), 400-406. doi:10.26438/ijcse/v6i11.400406

[21] The data has been published by Ministry of Statistics and Programme Implementation and sourced from Open Government Data (OGD) Platform of India [https://data.gov.in/resources/monthly-indices-all-india-index-industrial-production-nic-2008-2-digit-and-sectoral-leve-0]. Released under National Data Sharing and Accessibility Policy (NDSAP): https://data.gov.in/sites/default/files/NDSAP.pdf

[22] David Shaub and Peter Ellis (2019). forecastHybrid: Convenient Functions for Ensemble Time Series Forecasts. R package version 4.2.17. https://CRAN.R-project.org/package=forecastHybrid