Weighted Exponential Regression Model for Intraday Data

S.C. Thasleema[1]
Department of Statistics, S.V. University
Tirupati- 517502.

B.Sarojamma[2]
Department of Statistics, S.V. University,
Tirupati- 517502

Abstract- **In this paper weighted exponential regression model for intraday data is introduced. Constants are fixed in one model and estimated using OLS estimation in another model. These models are compared by using MSE and RMSE. These models are empirically proved by using wind speed data.**

Keywords- Weighted exponential regression model, Intraday data, OLS estimation, MSE, RMSE, Wind speed data.

I. INTRODUCTION

Time series analysis and forecasting methods play an important role in studying of temperature, wind speed, moisture in weather, etc. and are some of the variables of atmospheric data that change according to time with in a day. Intraday data is now widely establishing time series data model. Intraday data is the best example for working days of stock market values. Now a days intraday data analysis in fields of economic, business, stock markets are growing rapidly due to enthusiastically of leading about profits and losses in a day during hours. James W. Taylor in the article, "Exponentially weighted methods for forecasting intraday time series with multiple seasonal cycles", introduced five new univariate exponentially weighted methods for forecasting intraday time series data, contains intraweek as well as intraday seasonal cycles. For call centre intraday data, James W. Taylor introduced Holt winter Exponential smoothing model extended by adding intraday and intraweek cycles in intraday and intraday; a new double seasonal total and spilt exponential smoothing formulation is introduced by capturing the intraday and intraweek cycles to intraday data; using trigonometric functions discounted weighted regression with trigonometric terms for intraday data, splines based on discounted weighted regression. DWR and exponential smoothing established; singular value decomposition is introduced for time series. Gould, P.G. Koehler, A.B., ord, S.K., Snyder, R.D., Hyndman, R.J. & Vahid - Araghi, F. explains about Forecasting time series with multiple seasonal patterns. Harvey, A. & Koopman, S.J. (1993) says about "Forecasting hourly electricity demand by using time – varying splines". Poirier, D.J. explains about "Piece wise regression using cubic splines".

II. METHODOLOGY

By taking α, β and γ are constants and St, Wt and R_t are time series of summer season, winter season and rainy seasons.

$$Y_t = C + \alpha S_t + \beta W_t + \gamma R_t \quad \ldots\ldots\ldots\ldots\ldots(1)$$

where Y_t is time series values
S_t is summer season
W_t is winter season
R_t is rainy season
α, β and γ are constants.

Summer season (S_t) from March, April, May and June months, Winter season (W_t) from January, February, November and December months and Rainy seasons (R_t) from July, August, September and October months.
Estimation of the constants α, β and γ.

A. *Model 1:*

Compute St, Wt and R_t using the following equation.

$$S_t^{'} = C_s + \alpha^{'}(Y_t) \ldots\ldots\ldots(2)$$

$$W_t^{'} = C_W + \beta^{'}(Y_t) \ \ldots\ldots.(3)$$

$$R_t^{'} = C_R + \gamma^{'}(Y_t) \ \ldots\ldots.(4)$$

C_S, C_W, C_R, α, β and γ are coefficients and are estimated by solving three regression equations (2), (3) & (4) and using OLS estimations with each one for one season. Estimated equations for seasons by using exponential smoothing are

$$S_t' = C_S + \alpha'(Y_t) + (1-\alpha')\,S_{t-1} \ldots\ldots.(5),\ \alpha' + (1-\alpha') = 1$$

$$W_t' = C_W + \beta'(Y_t) + (1-\beta')\,W_{t-1} \ldots\ldots(6),\ \beta' + (1-\beta') = 1$$

$$R_t' = C_R + \gamma'\,(Y_t) + (1-\gamma')\,R_{t-1} \ldots\ldots(7),\ \gamma' + (1-\gamma') = 1$$

Multiple Regression line fitted by taking S_t, W_t, R_t as follows.

$$Y_t = C + \alpha S_t + \beta W_t + \gamma R_t$$

By using OLS estimation, estimated equation is as follows

$$\hat{Y}_t = C + \hat{\alpha}\,S_t + \hat{\beta} W_t + \hat{\gamma}\,R_t$$

B. Model 2:

In model 2, α, β and γ are taking equal value i.e., $\frac{1}{3}$ each and constant becomes zero i.e., $Y_t = \frac{1}{3} S_t + \frac{1}{3} W_t + \frac{1}{3} R_t$

RMSE: Positive square root of Mean square error gives

$$\text{RMSE} = \frac{(Y_t - \hat{y}_t)^2}{n}$$

where Y_t = time series value at time‘t’.

$\hat{y}_t$ = estimated time series value at time‘t’.

n = number of observations.

A model 1 which possesses lower value compared with model 2 RMSE value then model 1 is the best model compared with model 2.

III. EMPIRICAL INVESTIGATIONS

A. Model 1:

By taking hourly data of wind speed from Gadanki, Chittoor district, Andhra Pradesh, India. We perform weighted exponential regression equation as follows.

$$Y_t = C + \alpha S_t + \beta W_t + \gamma R_t$$

S_t, W_t and R_t are time series values of summer, winter and rainy seasons.

$$S_t' = C + \alpha' Y_t$$
$$W_t' = C + \beta' Y_t$$
$$R_t' = C + \gamma' Y_t$$

By using ordinary least squares method, the fitted equations are

$$S_t' = 216.6274 - (0.2265)\, Y_t$$
$$W_t' = 223.0351 + (0.6434)\, Y_t$$
$$R_t' = 237.8974 + (0.3218)\, Y_t$$

By taking coefficient of Y_t as α', β', γ' for S_t', W_t' *and* R_t' respectively. S_t', W_t' *and* R_t' are converted in exponential equation with constant as follows.

$$S_t = S_t' + (1 - \alpha')\, S_{t-1}$$
$$W_t = W_t' + (1 - \beta')\, W_{t-1}$$
$$R_t' = R_t' + (1 - \gamma')\, R_{t-1}$$

Therefore, equations becomes

$$S_t = 216.6274 - (0.2265)\, Y_t + (1.2265)\, S_{t-1}$$
$$W_t = 223.0351 + (0.6434)\, Y_t + (0.3566)\, W_{t-1}$$
$$R_t = 237.8974 + (0.3218)\, Y_t + (0.6782)\, R_{t-1}$$

The fitted weighted exponential regression for data is

$$Y_t = 11.2025 - (0.0071)\, S_t + (0.0119)\, W_t - (0.0016)\, R_t$$

B. Model 2:

By taking α, β and γ as 1/3 and constant zero then the model 2 becomes

$$Y_t = \frac{1}{3} S_t + \frac{1}{3} W_t - \frac{1}{3} R_t$$

where S_t, W_t and R_t are estimated in the above model.

RMSE: Root Mean Square error for both models is
RMSE Model 1: 3.3876
RMSE Model 2: 18.8083
Therefore, model 1 is the best compared with model 2.

IV. SUMMARY AND CONCLUSIONS

We are fitted two models for intraday data of wind speed of Gadanki, Andhra Pradesh.

$$Model\ 1: Y_t = 11.2025 - (0.0071)\, S_t + (0.0119)\, W_t - (0.0016)\, R_t$$

$$Model\ 2: Y_t = \frac{1}{3} S_t + \frac{1}{3} W_t + \frac{1}{3} R_t$$

Root Mean Square error for model 1 and model 2 are

RMSE Model 1: 3.3876
RMSE Model 2: 18.8083
Therefore, Model 1 is the best compared with Model 2.

V. ACKNOWLEDGEMENTS

The authors are thankful to Dr. T. Narayana Rao, Scientist, NARL, Gadanki for providing the wind speed data.

REFERENCES

[1] Gould, P.G., Koehler, A.B., Ord, J. k., Snyder, R.D., Hyndman, R.J & Vahid-Araghi, f. (2008): Forecasting time series with multiple seasonal patterns, European Journal of Operational Research, 191, 207-222.

[2] Harvey, A. & Koopman, S.J. (1993): Forecasting hourly electricity demand using time-varying splines, Journal of the American Statistical Association, 88, 1228-1236.

[3] James W Taylor 2010: Exponentially weighted methods for forecasting intraday time series with multiple seasonal cycles. International Journal of Forecasting 26, 627-646.

[4] James W Taylor 2006: Forecasting daily supermarket sales using exponentially weighted quantile regression. European Journal of Operational Research 178, 154-167.

[5] Spyros Makridakis, Steven C. Wheelwright, Rob J. Hyndman: Forecasting Methods and Applications, 3rd Ed., John Wiley & Sons.

[6] Poirier, D.J. (1973): Piecewise regression using cubic splines, Journal of the American Statistical Association, 68, 515-524.