Proceedings of the Jangjeon Mathematical Society
21 (2018), No. 4. pp. 689 - 701

www.jangjeon.or.kr
http://dx.doi.org/10.17777/pjms2018.21.4.689

Modeling the socio-economic waste generation factors: A case study of Faridabad (Haryana State, India)

Ajay Satija[1], Dipti Singh[2], Vinai Kumar Singh[1]

[1] Department of Applied Science & Humanities, Inderprastha Engineering College, Ghaziabad, Uttar Pradesh, India
[2]School of Vocational Studies & Applied Sciences, Gautam Budhha University, Gautam Budh Nagar, Uttar Pradesh, India
{aajaysatija@rediffmail.com;diptipma@rediffmail.com;drvinaiksingh@gmail.com}

Abstract. Municipal Solid Waste Management is an extremely complex task especially in Metropolitan cities. An accurate planning is required for its sustainable development. Such planning requires waste generated data as well as waste collected data. A number of socioeconomic factors are responsible for its generation. The aim of present research is to establish various significant correlations between socioeconomic factors such as population, urban population, literacy rate, per capita income and municipal solid waste (MSW) by regression analysis. The study is based on waste collected data of fastest growing metropolitan city Faridabad (Haryana State, India). The significant correlations between socioeconomic factors and MSW have been validated by p-value (< 0.05), high value of adjusted R^2 and minimum values of root mean square error (RMSE).Further the time series analysis has been performed to forecast the waste (collection) up to year 2019. The present study would be extremely beneficial for waste management authorities as well as policy makers of Municipal Corporation of Faridabad (MCF).

Keywords: Municipal Solid Waste, Municipal Solid Waste Management, Regression Analysis.

1 Introduction

The Municipal Solid Waste (MSW) increases with the increase in population, urbanization, industrialization, tourism and economic growth [17]. If such waste is not managed then this may cause serious environmental issue in the society [1, 3, 4, 5, 10]. Basically Solid Waste Management is related to all activities of managing the waste from its generation to final disposal [18] .Hence better waste management system is essentially required in the society.

The short terms predictions as well as long term predictions both are required for better planning and effective waste management system. Short term prediction is required to plan the waste collection, managing sanitation workers and transportation routes and related staff [13]. While long term prediction is required for managing the waste through recycling, incineration as well as its final disposal [14]. Waste

generation rate in India is less than other developed countries. The urbanization has higher impact on waste generation [2]. Higher per capita income implies higher economic prosperity of the people .A country with high per capita income generates more waste than the country with low per capita income. The seasonal variations can be seen in waste generation data .The time series models have been developed in literature to analyze such seasonal variation [8].

In previous studies the MSW has been forecasted by various methodologies such as non linear auto regressive models, exponential smoothing models [16]. The drawback of such models is that these models only shows seasonal variations and do not include the factors responsible for waste generation with their impact [15, 19]. But the regression models have capability to correlate the MSW and the other factors responsible for its generation. In various studies the Regression models have been developed to correlate the various socioeconomic factors with MSW.

In present research the major MSW generating factors such as population, urban population, literacy rate and per capita income of Faridabad District (Haryana State, India) have been analyzed .Here multiple regression models have been developed to predict the collected waste of Faridabad (within period January 2010 to January 2017) using these factors. Further the time series model has been developed to forecast the waste (collection) for next two years.

The present study will be helpful for the authorities of Municipal Corporation of Faridabad for waste management.

2 Materials and Methods

2.1 Case Study Area and Waste Management issues

The study area of the present research work is Faridabad. Sufi saint Sheikh Farid (in A.D. 1607) was the founder of Faridabad .Faridabad get separated from Gurgaon district and became 12th district of Haryana State on 15th August, 1979. The district has area 742.9 Km^2 with overcrowded population of 1,809,733 people (census India, 2011). The average Literacy rate in the district is 81.70%.

Municipal Corporation came into existence in year 1993 to serve the people of the city with services sewage, health, education, solid waste management. Today Faridabad is one of the core industrial estates of National capital region and highly populated district in Haryana State (India). Faridabad is one of the major industrial hubs of National Capital Region, India. But the waste is not properly treated in the district.

The Municipal Corporation of Faridabad (MCF) seems unable to manage the increasing amount of generated waste day by day. The waste management issues in Faridabad are: overcrowded population, public unawareness regarding energy renewal strategies from waste, improper planning to implement government MSW 2000 and MSW 2016 rules, lack of sanitation funds, advance equipment's and technical knowledge among municipal sanitation workers.

2.2 Material and Methods

The data of the population, urban population (%) and literacy rate have been collected from the website of Census of India. Per capita income data has been collected from Department of Economic and Statistical Analysis, Haryana (India). The MSW data has been gathered from the office of landfill site of village Bandhwari , District Gurugram (Haryana, India).

2.2.1 Population

Faridabad is emerging as prominent industrial city in National Capital Region of India. People are migrating from other districts and villages to the city in search of job. The population density is 2442 people per Km^2. Population wise the district has first place in the Haryana State. The population growth in the city in year 2001 was 58.88% and in 2011 was 32.54%.

A rapid growth has been observed in last two decade, which made district as highly populated. The population rate has been decreased in 2011 than census year 2001. The MSW has been assumed as strong function of the population .Table 1 presents the population rate of District Faridabad. Population has been forecasted for the year 2021 by Arithmetic Increase method.

Table 1. Population of Faridabad District

Census Year	Population Growth (%)	Population
1991	46.51	859431
2001	58.88	1365465
2011	32.54	1809733
2021	26.25	2284884

Source: [6,7]

2.2.2 Urban Population (%)

Urbanization has higher impact on waste generation. Faridabad is highly urbanized district in Haryana. There are various job opportunities in the city. Statistical data shows that this industrial estate has 15000 small scale, medium scale and large scale world famous industries. Better rail road transport facilities have been provided to the people of the district .The district is well connected by better rail road network with all other metropolitan cities of National Capital Region of India. The urban population

(%) has been presented in Table 2. Urban Population (%) for year 2021 has been forecasted by Arithmetic Increase method.

Table 2. Urban Population (%) of Faridabad District

Census Year	Urban Population (%)
1981	41.43
1991	48.57
2001	77.8
2011	79.51
2021	92.20

Source: [6]

2.2.3 Literacy Rate

Literacy in any society creates changes. Literacy among people creates social awareness to clean the environment around them. This is one of the major factors affecting waste management. Faridabad has fourth rank in terms of literacy rate in the Haryana State. The following table illustrates the Faridabad literate population. This was concluded from the table that the literacy rate drastically get increased from census year 1981 to 1991.Literacy rate for year 2021 is measured by Arithmetic Increase method.

Table 3 Literacy rate (%) of Faridabad District

Census Year	Literacy rate (%)
1981	39.11
1991	59.77
2001	76.30
2011	81.70
2021	95.89

Source: [6, 9]

2.2.4 Per Capita Income (PCI)

Faridabad and Gurugram districts contribute about 50% of income tax for their State Government [11] .Faridabad generates highest revenue for Haryana State Government. Basically if a person earns high income he would be considered as healthier, serve better to his family by education, living in a nice and secure home finally with a happier and luxurious life. It is observed that as income level gets increased, per capita waste generation get increased [12].

Table 4 present the per capita income of the District Faridabad. Per capita income for year 2018 has been estimated on average basis.

Table 4. Per capita income of District Faridabad

Financial Year	Per Capita Income (Rupees)
2005	41590
2006	49408
2007	54359
2008	58882
2009	61787
2010	106896
2018	238216

Source: Department of Economic and Statistical Analysis, Haryana

2.2.5 Municipal Solid waste (MSW) data

The MSW data records have been gathered from January 2010 to January 2017 (85 months) from landfill office of Bandhwari village , Gurugram. The fluctuations can be seen in MSW data due to seasonal variations, people migration for job, medical facilities as well as better education. Municipal Corporation of Faridabad (MCF) is the responsible for waste management. Table 5 presents the MSW collected data from the landfill site of Bandhwari village.

Table 5. MSW Collected (Kg) (January 2010- January 2017)

Year	Municipal Solid Waste (Kg)
2010	44631805
2011	70199590
2012	95295780
2013	116555040
2014	133871295
2015	100849555
2016	148782410
January 2017	19917480

2.2.6 Statistical Analysis of Data

Table 6 presents the Statistical Analysis of population, urban population %, literacy rate, per capita income and MSW. The table presents the number of months, mean value, minimum value, maximum value and skewness of each variable involved in the model formation. Table shows that the data is normalized because the standard deviation in each variable is less than half of the mean and the skewness factor of each variable is less than -1.

Table 6. Statistical Analysis of Socio-economic factors and MSW

Variable	Number of Months	Mean	Standard Deviation	Minimum	Maximum	Skewness
Population (XPOP)	85	1926246	96081	1764357	2091068	0.02
Urban Population(%) (URB)	85	81.50	1.87	79.24	85.40	0.54
Literacy Rate (LIT)	85	84.43	2.37	80.98	88.97	0.29
Per capita Income (PCI)	85	190409	22257	155359	227916	-0.04
Municipal Solid Waste (MSW)	85	8589	4185	195	19917	0.3

Table 7 presents the coefficient of correlation and p-value of input variables of regression model to the output variables.

Table 7. The coefficient of correlation of collected MSW with other Socio-economic factors

	Population	Urban Population (%)	Literacy Rate	Per Capita Income
MSW	0.649	0.613	0.635	0.639
p-Value	0.000	0.000	0.000	0.000

2.2.7 Proposed Multiple Regression Model

In simple regression only one parameter can be predicted with the help of another parameter but in multiple regression more than one parameters can be used to predict

some another parameter. Let's assume the following parameters for multiple regression models.

YMSW = Per month collection of MSW by sanitation workers of MCF. The unit assigned is Metric ton for variable YMSW

XPOP = A column vector of 85 months representing the population of Faridabad. The corresponding unit assigned for XPOP is person.

XURB = A column vector of 85 months representing the urban population (%) of Faridabad.

XLIT = A column vector of 85 months representing the literacy rate among Population of Faridabad

XPCI = A column vector of 85 months representing the Per capita income of Faridabad. The corresponding unit assigned for XPCI is Rupee.

E_i = Random error caused from unknown factors.

Multiple linear regression models can be expressed as

$$YMSW = b_0 + b_1 XPOP + b_2 XURB + b_3 XLIT + b_4 XPCI + E_i \quad (1)$$

The equation (1) can be extended to second degree as

$$YMSW = b_0 + b_1 XPOP + b_2 XURB + b_3 XLIT + b_4 XPCI + b_1 XPOP^2 + b_2 XURB^2 + b_3 XLIT^2 + b_4 XPCI^2 + b_5 XPOP * XURB + b_6 XPOP * XLIT + b_7 XPOP * XPCI + b_8 XURB * XLIT + b_9 XURB * XPCI + b_{10} XLIT*XPCI + E_i \quad (2)$$

b_0 = Height of the regression plane above the axis. The intercepts $b_1, b_2, b_3 \ldots b_k$ are the k slopes of plane in different 1, 2, 3…..k directions.

2.2.8 Proposed Time Series Model

In this study moving average model and exponential smoothing models have been developed.

(1) Moving Average model: The method smooth the irregularities of the time series model by averaging. The moving averages of n period are the consecutive averages of n terms. Consider the time series with n observations: X_t , X_{t+1},.....X_{t+n-1} . If n is odd say n=2k+1 then the calculated moving average term has to be placed against the middle value of periods i.e. along with t= k+1. If n is odd say n=2k then the calculated moving average term has to be placed against the middle vales of the periods it covers. For example considering 3-Yearly moving averages, then take the

average of three consecutive terms and place this value against second term .It is considered as predicted value. Considering 4-yearly moving averages, then take the average of four consecutive terms and place this value between second term and third term considered as 2.5^{th} period average .Skip first term and repeat the similar process for next four terms . Place it against 3.5^{th} period. Again take average of these two averaged terms (2.5^{th} period and 3.5^{th} period) and place against third term, repeat the process for fifth, seventh and so on as predicted value. Take the difference between actual and predicted value. The corresponding Mean absolute percentage error (MAPE) term can be evaluated. Simple moving average method allocates equal weights (1/n) to all n observation.

(2) Single and double Exponential Smoothing Model: The model allocates the weights to the observations. The value of these weights goes exponentially decreasing with the observations. The general expression for model is given by

$$F_{t+1} = \alpha X_t + (1- \alpha) F_t$$

Here F_t and F_{t+1} are forecast value of previous observation as well as next observation. X_t is the observation at period t and α is smoothing constant.

The double exponential smoothing handles both level (L_t at period t) and trends (T_t at period t). There are two smoothing parameters have been used to forecast next observation.

$$L_t = \alpha X_t + (1 - \alpha) [L_{t-1} + T_{t-1}]$$
$$T_t = \beta [L_t - L_{t-1}] + (1 - \beta) T_{t-1}$$
$$F_{t+1} = L_t + T_t$$

For next n-period forecast is given by

$$F_{t+1} = L_t + n T_t$$

The value of α and β must have to be chosen between 0 and 1.

3 Results

Table 8 presents the proposed regression models .Various correlations have been established between population, urban population (%), literacy rate, per capita income and MSW. The significance of these correlations has been tested by p-value (< 0.05). These results have been validated with Root mean square error (RMSE) and adjusted R^2.The first model can be presented by equation (3)

$$YMSW = 24402 + 0.0749 XPOP - 1896 XLIT \quad (3)$$

The model can be interpreted as keeping XLIT fixed if one person in population is increased then MSW will be increased by 0.0749 Metric tons. Similarly if XPOP is assumed constant then 1% increase in XLIT will decrease YMSW by 1896 Metric tons on average basis.

Table 8. Regression Models Analysis

S.N.	MODEL	RMSE	R^2 adjusted	p-value
1	YMSW = 24402 + 0.0749 XPOP - 1896 XLIT	3187	42.0	0.000
2	YMSW = 8654 + 0.0515 XPOP - 1217 XURB	3186	42.0	0.000
3	YMSW = 3500 - 270 XURB + 0.142 XPCI	3253	39.5	0.000
4	YMSW = 7455816 - 2.40 XPOP - 107587 XLIT + 0.0384 XPOP*XLIT	2951	50.3	0.000
5	YMSW = 17585429 + 165 XPOP - 4473424 XLIT - 2.15 XPOP*XLIT + 54956 $XLIT^2$	2564	62.4	0.000
6	YMSW = 17955449 + 157 XPOP - 4267877 XLIT - 2.04 XPOP*XLIT + 52201 $XLIT^2$ - 0.000001 $XPCI^2$	2524	63.6	0.000
7	YMSW = 1.77 x 10^{08} - 41.1 XPOP + 5146913 XURB - 7728923 XLIT - 0.056 XPCI – 2.64 XPOP*XURB + 3.38 XPOP*XLIT	2497	64.4	0.000
8	YMSW = 51201282 + 84.1 XPOP - 521181 XURB - 2673070 XLIT - 0.119 XPCI – .04 XPOP*XURB + 31784 XURB*XLIT	2515	63.9	0.000
9	YMSW = 55814889 - 31.4 XPOP - 228410 XURB - 467240 XLIT + 65.9 XPCI + 0.401 XPOP*XURB + 3.43 XLIT*XPCI - 4.38 XURB*XPCI	2556	62.7	0.000
10	YMSW = 3.02 x 10^{08} - 221 XPOP + 12696477 XURB - 13106728 XLIT - 81 XPCI – 3.5 XPOP*XURB + 7.0 XPOP*XLIT – 0.00014 XPOP*XPCI - 57248 XURB*XLIT - 8 XURB*XPCI + 12 XLIT*XPCI	2553	62.8	0.000
11	YMSW = 1.83 x 10^{08} - 42.2 XPOP + 5356200 XURB - 8044237 XLIT + 2.5 XPCI – 2.73 XPOP*XURB + 3.50 XPOP*XLIT - 0.000001 XPOP*XPCI	2512	63.9	0.000

The model is logically true because literacy rate increment decreases YMSW and increase in population implies more generation of waste. The RMSE of model is 3187 Metric tons and value of adjusted R^2 is 42 %.
It is clear that MSW is strong function population. The second correlation has been established between variables MSW, population and urban population.

It is presented by equation (4).

$$YMSW = 8654 + 0.0515\ XPOP - 1217\ XURB \quad (4)$$

The model indicates that if one person in population is increased then there will be increase of 0.0515 Metric tons in MSW keeping XURB fixed on average basis. Similarly if 1% of urban population is increased then there will be decrease of 1217 Metric tons of MSW keeping population fixed. The RMSE value of MSW is 3186 Metric tons and adjusted R^2 value is 42%. Similar interpretation can be done for third model.

The fourth model has been presented by equation (5)

$$YMSW = 7455816 - 2.40\ XPOP - 107587\ XLIT + 0.0384\ XPOP*XLIT \quad (5)$$

The RMSE value is 2951 Metric tons. The value of adjusted R^2 is 50.3%. In this model the interaction term XPOP*XLIT has been added. Now the meaning of all coefficients of predictor variables has been changed with response variable. The model can be interpreted as the effect of population on MSW is (0.0384*XLIT -2.40).

The fifth model has been presented by equation (6)

$$YMSW = 17585429 + 165\ XPOP - 4473424\ XLIT - 2.15\ XPOP*XLIT + 54956\ XLIT^2 \quad (6)$$

The model contains interaction term of population and literacy (XPOP*XLIT) along with quadratic term of literacy rate ($XLIT^2$). It is observed that now the non linearity get added in the model. Positive value of coefficient of $XLIT^2$ suggests that the relation now becomes convex. The negative value of quadratic terms suggests that the relation would be concave. The positive coefficient of $XLIT^2$ suggests that extra unit of XLIT has larger impact on YMSW. The RMSE value is 2564 Metric tons. The value of adjusted R^2 is 62.4.The value of adjusted R^2 now get increased means that new added terms have improved the predictability of the model. The rest regression models can be predicted in similar manners.

In this study the time series forecasting models (Moving average model and the exponential models) have been developed. The results of collected MSW have been compiled in table 9. In the moving average model the moving average length is chosen as 3. The forecasted value of MSW collected in January 2019 is 18268.5 Metric tons. The 95% prediction interval is (12219.4, 24317.5). The Mean Absolute percentage error (MAPE) and Mean Absolute Deviation (MAD) are 54% and 2137.

In the single exponential model the value of smoothing parameter α is selected as 0.394531. The forecasted value of MSW collected in January 2019 is 17910.3 Metric tons. The 95% prediction interval is (13065.8, 22754.8). The MAPE and MAD are 55% and 1977 respectively.

In the double exponential model the value of smoothing parameters α and β are selected as 0.824232 and 0.017857. The forecasted value of MSW collected in January 2019 is 27160.0 Metric tons. The MAPE and MAD are 61 and 2003 respectively. The results of the double exponential Smoothing model appears more realistic as it handles both level and trend factor of time series model.

Table 9: The results of time series models

Model Name	Forecasted Value of MSW (collected) in January 2019 (Metric ton)	Mean Absolute percentage error (MAPE) (%)	Mean Absolute Deviation (MAD) (Metric ton)	Forecasted Value of MSW (collected) in 2019 on average basis (Metric ton)
Moving Average Model	18268.5	54	2137	219222
Single Exponential Smoothing model	17910.3	55	1977	214923.6
Double Exponential Smoothing model	27160.0	61	2003	325920

4 Conclusion

Faridabad is rapidly growing, highly urbanizing city of Haryana State of India. The Municipal Solid Waste generation and its management is now becoming an environmental issue for the authorities of Municipal Corporation of Faridabad. Hence Proper planning is required to handle this .Accurate planning requires accurate waste generated and collected data. The planning also requires how much waste in future will be generated and how much waste will be collected .Hence a predictive model is required for such planning. In this study eleven predictive models based on regression analysis have been proposed. Various correlations between socioeconomic factors and MSW have been established. First three models have shown significant relationships but they have low values of adjusted R^2. Hence to enhance the predictability various quadratic terms as well as interaction terms of variables have been added and

significant relations have been found. The Model 7 has shown highest value of adjusted R^2 (64.4%) and RMSE value is lowest (2497 Metric tons) than other models. The predictive results can be enhanced by applying other approaches such as Artificial Neural Network Models (ANN) and Adaptive Neuro Fuzzy Inference System (ANFIS) Models. The proposed time series model (Double exponential smoothing model) has shown that the forecasted value of collected MSW in Month January 2019 is 27160.0 Metric tons. The MSW collected will be 325920 Metric tons in year 2019 on its average basis. Hence Municipal Corporation of Faridabad (MCF) has to plan to manage the increased amount of waste

References

1 Al-Khatib Issam, A., Monou, M., Abu Zahra Abdul Salam, F., Shaheen Hafez, Q., Kassinos, D.: Solid waste characterization, quantification and management practices in developing countries. A case study: Nablus district—Palestine. J Environ Manag 91,1131–1138 (2010)

2 Asnani, P.U.: India infrastructural report. Solid Waste Management. http://www.iitk.ac.in/3inetwork/html/reports/IIR2006/Solid Waste.pdf (2006) (accessed 06.08.2017)

3 Batool ,S.A., Chaudhry ,M.N. : Municipal solid waste management in Lahore city district, Pakistan. Waste Manage (Oxford) 29:1971–1981(2009)

4 Burnley, S.J., Ellis, J.C., Flowerdewc, R., Polld, A.J., Prosser, H.: Assessing the composition of municipal solid waste in Wales. Resour Conserv Recycl 49, 264–283 (2007)

5 Burnley, S.J.: A review of municipal solid waste composition in the United Kingdom. Waste Manag. 27,1274–1285 (2007)

6 Census of India, 2011. Ministry of Home Affairs, Government of India, New Delhi, India. http://www.census2011.co.in/census/district/227-faridabad.html. (accessed 06.08.2017)

7 City population information , http://www.citypopulation.de/php/india-haryana.php?adm2id=0620. (accessed 06.08.2017)

8 Denafas, G., Ruzgas, T., Martuzevicˇius, D., Shmarin, S., Hoffmann, M., Mykhaylenko, V., Ogorodnik, S., Romanov, M., Neguliaeva, E., Chusov, A., Turkadze, T., Bochoidze, I., Ludwig, C.: Seasonal variation of municipal solid waste generation and composition in four East European cities. Resour. Conserv. Recycl. 89, 22–30 (2014)

9 Education for all in India, Sarva Shiksha Abihiyan information, http://www.educationforallinindia.com/page157.html.(accessed 06.08.2017)

10 Go´mez, G., Meneses, M., Ballinas, L., Castells, F.: Characterization of urban solid waste in Chihuahua, Mexico. Waste Manage (Oxford) 28,2465–2471 (2008)

11 Infogalactic, The Planetary Knowledge Core information. https://infogalactic.com/info/Faridabad. (accessed 06.08.2017)

12 Liu, C., Wu, X. W.: Factors influencing municipal solid waste generation in China: a multiple statistical analysis study. Waste Management and Research 29, 371–378 (2010)

13 Matsuto, T., Tanaka, N.: Data analysis of daily collection tonnage of residential solid waste in Japan. Waste Manag Res 11, 333–343 (1993)

14 Navarro-Esbrı´, J., Diamadopoulos, E., Ginestar, D.: Time series analysis and forecasting techniques for municipal solid waste management. Resour Conservat Recycl. 35, 201–214 (2002)

15 Noori, R., Abdoli, M.A., Farokhnia, A., Abbasi, M.: Results uncertainty of solid waste generation forecasting by hybrid of wavelet transform-ANFIS and wavelet transform-neural network. Expert Syst. Appl. 36, 9991–9999 (2009)

16 Rimaityte, I., Ruzgas, T., Denafas, G., Racys, V., Martuzevicius, D.: Application and evaluation of forecasting methods for municipal solid waste generation in an Eastern-European city. Waste Management and Research 30, 89–98 (2012)

17 Saeed, M.O., Hassan, M.N., Mujeebu, M.A.: Assessment of municipal solid waste generation and recyclable materials potential in Kuala Lumpur, Malaysia. Waste Management 29,2209–2213 (2009)

18 Satija, A., Singh, D., Singh, V. K.: Modeling the socio-economic waste generation factors using artificial neural network: A case study of Gurugram (Haryana State, India). In: Proceeding of second International Conference on Modern Mathematical Methods and High Performance Computing in Science and Technology (M3HPCST 2018) Ghaziabad, Uttar Pradesh, India (2018)(Communicated).

19 Shan, C.S.: Projecting municipal solid waste: the case of Hong Kong SAR. Resour. Conserv. Recycl. 54, 759–768 (2010)