An ARIMA model for the forecasting of healthcare waste generation in the Garhwal region of Uttarakhand, India

Ankur Chauhan* and Amol Singh

Department of Operations Management,
Indian Institute of Management Rohtak,
MD University Campus,
Rohtak, Haryana 124001, India
Email: Chauhan.ankur2903@gmail.com
Email: amolasingh2007@rediffmail.com
*Corresponding author

Abstract: This study has been carried out to analyse and forecast the quantities of healthcare waste generated from the hospitals of Garhwal region of Uttarakhand, India. In this study, a suitable autoregressive integrated moving average (ARIMA) model has been developed, on the basis of different statistical parameters, for the forecasting of healthcare waste. The analysis of results on the basis of the statistical parameters such as adjusted *R*-square value, mean square error and mean absolute percentage error; the AR(1)MA(1) model has been found as the best ARIMA model for the forecasting of healthcare waste generation. The daily data of healthcare waste generation has been used to develop the ARIMA model in this study. The ARIMA model developed in this study would help the waste disposal firm to plan its waste collection and disposal strategy-related decisions in future.

Keywords: ARIMA model; capacity planning; forecasting; healthcare waste; time series.

Reference to this paper should be made as follows: Chauhan, A. and Singh, A. (201X) 'An ARIMA model for the forecasting of healthcare waste generation in the Garhwal region of Uttarakhand, India', *Int. J. Services Operations and Informatics*, Vol. XX, No. YY, pp.XXX–XXX.

Biographical notes: Ankur Chauhan received MTech degree from the National Institute of Technology, Jalandhar. He is a Doctoral student at Indian Institute of Management Rohtak, India. His research interest lies in various waste management areas such as solid waste, electronic waste and healthcare waste. Currently, he is carrying out his PhD thesis work in the area of healthcare waste management. He has worked on research articles using multi-criteria decision-making techniques, i.e. ISM, DEMATEL, AHP, ANP and TOPSIS. He is also comfortable with various statistical methods such as regression and ARIMA modelling.

Amol Singh received PhD in Industrial Engineering from IIT Roorkee in 2006 and his ME in Production Engineering from Moti Lal Nehru National Institute of Technology, Allahabad, in 2000. He is a faculty in the Area of Operations at IIM Rohtak. His research interests include various aspects of operations management such as project management and supply chain management. He has published several research papers in International Journals and conferences.

1 Introduction

Healthcare waste (HCW) has been defined as a waste that is generated during different patient care activities performed in hospitals, nursing homes, clinics, home care, pathology labs, research centres and veterinary centres (WHO, 2013). The waste generated in patient care activities may consist of syringes, scalpels, bandages, blooded cottons, blood, body parts, chemicals and cytotoxics. Because the exposure of this waste may cost heavily to living beings and environment for being prone in spreading disease, polluting air with foul smell, contaminating soil and ground water, so, it has been named and categorised in infectious and hazardous waste by practitioners and different environmental bodies such as SBC (2013) and USEPA (2016). The variability in the composition (percentage of hazardous and infectious waste) and constituents (types of waste products included in HCW) of HCW has been studied by various researchers across the world (Ananth et al., 2010; Komilis et al., 2012).

According to Pruss (2014) and WHO (2013), the composition of HCW should be such that it consists of 80–85% non-hazardous waste and 15–20% hazardous and infectious components, if segregated properly. The use of colour-coded bags, training and awareness programs for housekeeping staff and attitude of the workers have a co-relation with the segregation of HCW (Komilis et al., 2012; Nema et al., 2011; Oroei et al., 2014; Tudor et al., 2007). However, the lack of implementation of such segregation practices along with the increased usage of disposable custom packs, the increment in hazardous and infectious waste quantities has been described by researchers in their studies (Campion et al., 2015; Caniato et al., 2015; Chauhan et al., 2016). Campion et al. (2015) revealed the increased usage of disposable custom packs in healthcare activities of patient care across the world. The custom packs include a set of disposable products used for a precise procedure to reduce any difficulties or mistake in a patient care activity. Because the usage of a custom pack is limited to one healthcare treatment procedure, so it converts into waste after its usage and leads to the huge amount of waste generated such as 5.9 million tons per year in USA. Likewise, the developing countries such as India also generate a huge amount of HCW, i.e. 17.6 million tons per year (IndiaStat, 2013).

As described previously, HCW consists of hazardous and infectious waste along with the general waste. Because the hazardous and infectious components of HCW could be very harmful to environment and society, therefore, hazardous and infectious components of HCW should not be left unattended for the safety of society and environment (Caniato et al., 2015). In addition, in 2000, world health organization (WHO) estimated the gravity of the issue by revealing that injections with contaminated syringes caused 21 million hepatitis B virus (HBV) infections (32% of all new infections), two million hepatitis C virus (HCV) infections (40% of all new infections) and 260 000 HIV infections (5% of all new infections) (WHO, 2013). Thereby, compelling the world forums such as WHO and the United States Environmental Protection Agency (USEPA) to formulate and implement the international guidelines for better HCW treatment and disposal across the globe (Pruss, 2014; Pruss et al., 1999). The treatment and disposal of HCW should be carried out using one of the three known methods, i.e. landfilling, incineration and recycling (Alvim-Ferraz and Afonso, 2005; Caniato et al., 2015; Chauhan and Singh, 2016a). Hence, a disposal facility with a sufficient capacity is needed for the treatment and disposal of HCW. Because the capacity planning of a treatment and disposal facility is a long-term decision, therefore, the availability or demand of raw material (HCW in this case), in terms of present and future scenario, for processing in a disposal facility should be forecasted for HCW disposal facility (Dyson and Chang, 2005; Martínez-Costa et al., 2014; Mula et al., 2006; Schuh et al., 2011).

The present study has been carried out to formulate a forecasting model, an autoregressive integrated moving average (ARIMA) model, to predict the generation of HCW in future. The organisation of paper is as follows: The literature review on Time Series Forecasting Modeling has been provided in Section 2. Section 3 provides the detail of the Materials and Methods used in the study. Section 4 is sufficed with Results and Discussion. The conclusion has been given in Section 5. The implications of this work have been given in Section 6. The limitations and future research directions have been given in Section 7.

2 Literature review on time series forecasting modelling

Forecasting helps the organisations of different sectors to plan their operations in advance for an optimum output under static or dynamic (uncertain) environment (Mula et al., 2006; Rongfang et al., 2012; Wang et al., 2015). The studies conducted in these sectors include the issues such as forecasting of electricity pricing, agri-fresh supply chain, manufacturing or production planning, tourist arrivals and waste management (Chaâbane, 2014; Hassani et al., 2015; Martínez-Costa et al., 2014; Oribe-Garcia et al., 2015; Rasouli and Timmermans, 2012; Shukla and Jharkharia, 2011).

Majorly, the studies related to forecasting of waste generation have been carried out for addressing plastic waste, electronic waste and municipal solid waste (Al-Khatib et al., 2016; Denafas et al., 2014; Intharathirat et al., 2015; Peeters et al., 2015; Rimaityte et al., 2012; Tran et al., 2014). Doing so helps the waste disposal firms to formulate the strategies related to waste disposal planning that includes capacity building in terms of short-term (daily management) and long-term (proper design of facilities) decision-making (Chauhan et al., 2015; Oribe-Garcia et al., 2015). To work upon this, the forecasting models have been developed using single or multiple (hybrid) approach of regression analysis, ARIMA, simple autoregressive moving average, artificial neural network and grey theory, in a study (Intharathirat et al., 2015; Li et al., 2010; Shamshiry et al., 2014; Song and He, 2014; Weron, 2014; Xu et al., 2013). However, the applications of such methods including ARIMA model, for the forecasting of HCW generation, have received limited attention in the literature (Chauhan and Singh, 2016b).

Al-Khatib et al. (2016) suggested a regression model in their study to assess the hospital solid waste generation rate and its composition for developing a sustainable management of HCW. In addition, the authors considered the cross-sectional data to develop the regression model between dependent and independent variables such as daily total hospital waste and number of inpatients, respectively. Karpusenkaite et al. (2016) carried out a study to forecast HCW generation using three short and extra short data sets, i.e. 20, 10 and 6 observations; the best and most promising results of their study were obtained using artificial neural network, generalised additive and smooth splines models. However, the small data set considered in their study could be sufficient to forecast the generation of HCW in an illustration of comparing different methods, but it may not be appropriate to generalise the results of the studies due to high uncertainty (Chatfield and Prothero, 1973). In addition, the tests on seasonal and stationary behaviour of the data have not been performed, due to few data points, which may lead to the less reliable and poor results (Box and Jenkins, 1976; Granger and Newbold, 1974; Haykin, 2004; Navarro-Esbri et al., 2002).

In view of the reviewed literature, it could be stated that the majority of the forecasting models which have been developed in the context of HCWM are less reliable due to the consideration of very small data sets for model development. In the present study, the ARIMA model has been developed on the basis of kg day^{-1} waste generation data of 365 days, collected from different hospitals, to help the HCW disposal firm. Hence, with the help of ARIMA model developed for the forecasting of HCW generation, the waste disposal firm would be benefited in numerous ways including HCW disposal capacity planning, manpower planning and collection infrastructure.

3 Materials and methods

Time series forecasting models help in interpreting the relationship of the past observations of a variable to its future values (Cryer and Kellet, 1986; Shumway and Stoffer, 2011). Zhang (2003) stated that the application of a time series model becomes more vital in the absence of an explanatory model, which relates the prediction variable to other explanatory variables. In addition, the surge in the applications of forecasting models has also been highlighted by various researchers in their studies (Chang, 2014; Hassani et al., 2015; Zhang, 2003). Time series models include the modelling with ARIMA, neural network (NN) and support vector machines (SVM). However, the applications of such models are based on the dominant patterns that exist in the data sets, i.e. linear or non-linear pattern. According to Box and Jenkins (1965), the ARIMA modelling is most suitable for linear pattern data sets, whereas NN and SVMs are fit for non-linear data (Cristianini and Shawe-Taylor, 2000; Zhang et al., 1998). Because the data of the present study follow linear pattern; therefore, it has been modelled with the help of an ARIMA model.

3.1 The ARIMA model

The ARIMA model has the statistical properties that make it best fitted for the forecasting of linear data patterns (Box and Jenkins, 1976). This is assumed in an ARIMA model that the future values of a variable are linear function of numerous past observations and random errors. Therefore, the fundamental process of generating a time series can be given in the form of Eq. (1)

$$y_t = \alpha_0 + \alpha_1 y_{t-1} + \alpha_2 y_{t-2} + \cdots + \alpha_p y_{t-p} + \varepsilon_t - \beta_1 \varepsilon_{t-1} - \beta_2 \varepsilon_{t-2} - \cdots - \beta_q \varepsilon_{t-q} \tag{1}$$

where y_t and ε_t represents the actual value and random error at time period *t*, respectively; $\alpha_i \left(i = 1, 2, 3, \ldots\ldots, p\right)$ and $\beta_j \left(j = 0, 1, 2, 3, \ldots\ldots, q\right)$ are model parameters; *p* and *q* are integers that are usually referred as the order of the model. In addition, the random errors, ε_t, are assumed to be independently and identically distributed with a mean of zero and a constant variance of σ^2. The model converges into an autoregressive (AR) model of order *p*, if $q = 0$ in Eq. (1), and the model converges into a moving average (MA) model of order *q,* if $p = 0$ in Eq. (1). Likewise, it can be concluded that the main task of an ARIMA model formulation is to decide an appropriate model of order (p, q).

Yule (1926) and Wold (1939) did the seminal work on time series data that became an important contribution for development of ARIMA model by Box and Jenkins (1976). Yule (1926) studied the correlations in detail and revealed that the blind follow-up of the magnitude of correlation coefficient may be misleading at times. In addition, the insights on the causes of absurd correlations, with the analysis of sample and the nature of time series, i.e. stationary or non-stationary, have been provided in their work. Wold (1939) conducted an in-depth statistical analysis on stationary time series by extending the theorems, proposed by Khintchine (1924), for developing other theorems of relative to discrete time series. Granger and Newbold (1974) described that a time series data set should be stationary for obtaining a good forecasting model, whereas the results obtained from a non-stationary data set may be spurious for forecasting. Therefore, with the help of the seminal works carried out in the past, Box and Jenkins (1976) developed the methodology of formulating ARIMA model. Box and Jenkins' three steps methodology of ARIMA model building include identification, parameter estimation and diagnostic checking.

1 Step 1: To check the time series is stationary or not; if, the time series is not stationary and it shows some trend and heteroscedecity, then the difference and power transformation are applied on it. With this, the time series converts into a stationary time series with uniform variance for fitting the ARIMA model.

2 Step 2: In view of the target of minimisation of overall errors, the model parameters are estimated out-rightly after the formulation of a tentative model.

3 Step 3: To check and validate, the model assumptions about error terms, ε_t , are accomplished. The goodness–of-fit tests have been performed using statistical information such as the value of Akaike Information criteria (AIC). Box and Jenkins (1962) stated the importance of AIC and SC, in terms of more the better; therefore, the values of these criteria help in choosing this model. Therefore, the diagnostic information helps in achieving the appropriate ARIMA model.

Finally, the above-mentioned three steps of the model formulation process are repeated to achieve a desired objective in the form of an ARIMA model. The model obtained after the required iterations may be used for the forecasting purpose.

3.2 Evaluation criteria for forecasting performance assessment

Mean absolute percentage error (MAPE) or mean absolute error (MAE) has been coined as a term for the measurement of variation of dependent series data from its model-predicted (forecasted) level (McCleary et al., 1980). Because this measure is independent of the units of a series, therefore, it can be used to compare the series with different units. The formula for the calculation of MAPE is as follows:

$$M = \frac{100\%}{n} \sum_{i=1}^{n} \frac{A_i - F_i}{A_i}$$

where A_i and F_i denote the actual and forecasted values, respectively. If the value of MAPE (M) is zero, then it would be considered as a perfect fit for an ARIMA model.

However, there is no specification about the upper limit of MAPE values (Box et al., 2008; DeLurgio, 1998).

3.3 Case study area and source of data

The data to meet the objective of the paper are collected from central bio-medical waste disposal firm (CBWDF)-Roorkee, Garhwal region, which records the total amount of HCW generated in the region. The CBWDF keeps the record of bio-medical waste generation in terms of yellow bag waste, red bag waste and blue bag waste, which are collected every day from different healthcare facilities. For this study, the CBWDF could provide the daily (per day) data of past 12 months. The CBWDF-Roorkee covers more than 400 primary (1 < 10 beds), secondary (10 < 50 beds) and tertiary (10 < beds) bio-medical waste generation facilities. The CBWDF-Roorkee is well equipped with incinerators, autoclaving machines and other equipment such as computers. The CBWDF has to abide by the guidelines of HCW disposal; therefore, it has to plan for the proper disposal of HCW in advance. The planning includes some long-term and short-term decisions such as purchasing an incinerator and vehicles for waste collection, respectively. In practice, these decisions depend on the basis of total quantities of waste collected that is generated by healthcare facilities; therefore, this study has been carried out to assist the CBWDF in the better planning of HCW disposal with the help of forecasting.

4 Results and discussion

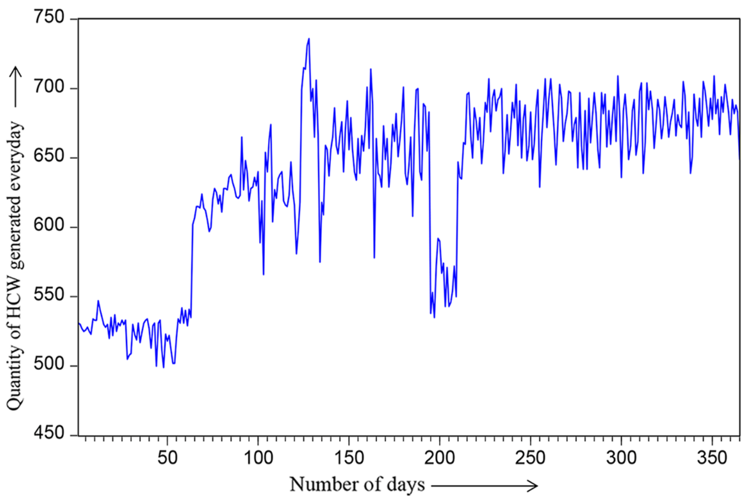
The e-views 8.1 version has been used for developing an ARIMA model for the forecasting of HCW in this study. The assumption related to the nature of data, such as stationarity, has been tested for the development of forecasting model. Therefore, the stationarity of the data has been checked with the help of a graphical representation and unit root test, as shown in Table 1 and Figure 1, respectively. For unit root test, the augmented Dickey-Fuller test statistic of the data is significant at the level of 5%, which confirms that the data of the present study are stationary. In addition, using *t*-statistic value also at 1, 5 and 10%, it can be stated that the unit-root test at 5% significance level is successful, which confirms the stationarity of data set.

Table 1 Test statistics to check stationary nature of present time series

Null hypothesis: daily total HCW generation has a unit root			
Exogenous: constant		t-Statistic	Prob.*
Lag length: 2 (automatic - based on SIC, maxlag = 16)			
Augmented Dickey-Fuller test statistic		–2.857	0.049
Test critical values	1% level	–3.448	
	5% level	–2.869	
	10% level	–2.571	

*MacKinnon (1996) one-sided *p*-values.

Figure 1 Graphical representation of daily amount of HCW generation (Kg/day) (see online version for colours)



Correlogram for the data is inclusive of autocorrelation (for AR lags) and partial autocorrelation (for MA lags). From Figure 1, the recursive combination of models is tested, and then the best-fit model is concluded to be utilised for forecasting. Table 2 details the different ARMA models than have been tested and compared on the basis of *p*-value, Akaike information criteria (AIC) and Schwarz criteria (SC). The AR(1)MA(2) model is not significant on the basis of *p*-value; hence, it cannot be used as a forecasting model. In contrast to this, the models AR(1)MA(1), AR(2)MA(2) and AR(2)MA(1) are significant on the basis of *p*-value; therefore, they would be appropriate for the forecasting of HCW generation in this study.

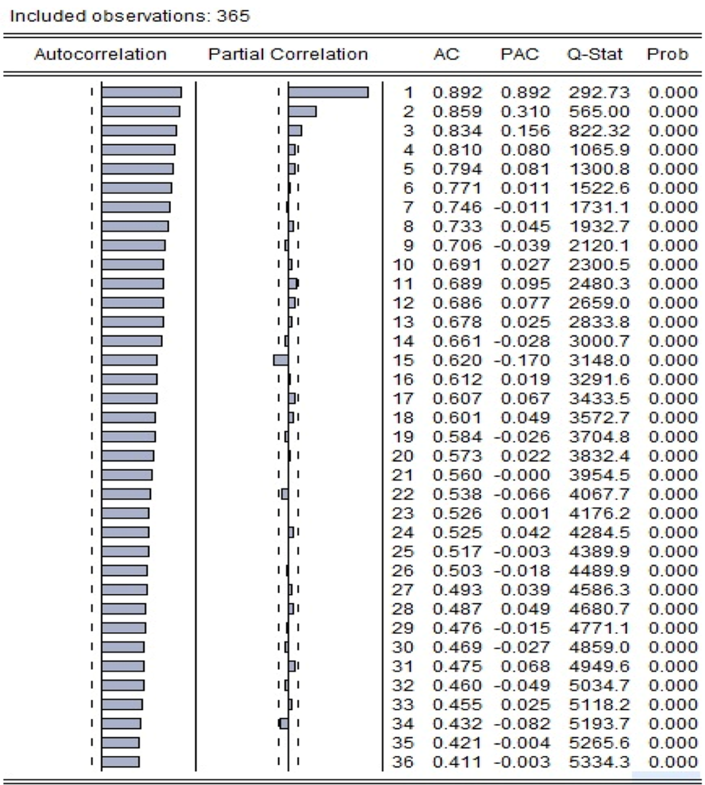
Table 2 Different ARIMA models tested for the study

	Coefficient	*p-Value*	*AIC*	*SC*
C	648.705	0.000	9.510	9.542
AR(1)	0.973	0.000		
MA(1)	–0.504	0.000		
C	653.715	0.000	9.263	9.295
AR(2)	0.957	0.000		
MA(2)	-0.474	0.000		
C	637.487	0.000	9.427	9.459
AR(1)	0.890	0.000		
MA(2)	0.011	0.838		
C	637.790	0.000	9.409	9.441
AR(2)	0.796	0.000		
MA(1)	0.811	0.000		

Because we are supposed to select the best ARMA model for the forecasting of HCW generation in the present case study; therefore, we have compared the models AR(1)MA(1), AR(2)MA(2) and AR(2)MA(1) using AIC and SC values, and it has been noticed from Table 2 that the AIC and SC values are highest for AR(1)MA(1). Thereby, AR(1)MA(1) model have been chosen over AR(2)MA(2) and AR(2)MA(1) models for the forecasting of HCW generation.

The correlogram shown in Figure 2 includes autocorrelation (AC), partial autocorrelation (PAC), Q-statistics and probabilities of these statistics. The correlogram helps in the identification of an appropriate autoregressive (AR) and moving average (MA) series for a forecasting estimation equation, i.e. Eqs. (1), (2a) and (2b), respectively. The equations, with detailed description, have been provided in Table 3.

Figure 2 Correlogram of daily HCW generation data (see online version for colours)



Included observations: 365

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.892	0.892	292.73	0.000
		2	0.859	0.310	565.00	0.000
		3	0.834	0.156	822.32	0.000
		4	0.810	0.080	1065.9	0.000
		5	0.794	0.081	1300.8	0.000
		6	0.771	0.011	1522.6	0.000
		7	0.746	-0.011	1731.1	0.000
		8	0.733	0.045	1932.7	0.000
		9	0.706	-0.039	2120.1	0.000
		10	0.691	0.027	2300.5	0.000
		11	0.689	0.095	2480.3	0.000
		12	0.686	0.077	2659.0	0.000
		13	0.678	0.025	2833.8	0.000
		14	0.661	-0.028	3000.7	0.000
		15	0.620	-0.170	3148.0	0.000
		16	0.612	0.019	3291.6	0.000
		17	0.607	0.067	3433.5	0.000
		18	0.601	0.049	3572.7	0.000
		19	0.584	-0.026	3704.8	0.000
		20	0.573	0.022	3832.4	0.000
		21	0.560	-0.000	3954.5	0.000
		22	0.538	-0.066	4067.7	0.000
		23	0.526	0.001	4176.2	0.000
		24	0.525	0.042	4284.5	0.000
		25	0.517	-0.003	4389.9	0.000
		26	0.503	-0.018	4489.9	0.000
		27	0.493	0.039	4586.3	0.000
		28	0.487	0.049	4680.7	0.000
		29	0.476	-0.015	4771.1	0.000
		30	0.469	-0.027	4859.0	0.000
		31	0.475	0.068	4949.6	0.000
		32	0.460	-0.049	5034.7	0.000
		33	0.455	0.025	5118.2	0.000
		34	0.432	-0.082	5193.7	0.000
		35	0.421	-0.004	5265.6	0.000
		36	0.411	-0.003	5334.3	0.000

Table 3 Equations for forecasting with AR(1)MA(1) model

Equation number	*Equation name*	*Detailed description of an equation*	*The detailed equations with notations**
Equation 1	Estimation	This equation helps in achieving the suitable ARMA model, in this case it is AR(1)MA(1)	LS (DERIV=AA)_365_days HCW C AR(1) MA(1)
Equation 2a	Forecasting	On the basis of this equation, the forecasting for the desired period could be conducted for the usage of CBWDF.	_365_days HCW = C(1) +[AR(1)=C(2),MA(1)=C(3),BACKCAST=2,ESTSMPL= "2365"]
Equation 2b	Equation after substituting the coefficients with numerical values	The linguistics such as coefficient (C) is replaced with numeric values to obtain the forecasted values for a defined period for the usage of CBWDF	_365_days HCW=648.705 +[AR(1)=0.973,MA(1)=-0.504,BACKCAST=2,ESTSMPL= "2 365"]

*Notations used in the above equations are as follows:

365 stands for the data set, i.e. number of days

HCW stands for healthcare waste management

AR stands for autoregressive

MA stands for moving averages (error terms)

C(1), C(2) and C(3) stands for coefficients related to intercept, AR and MA, respectively.

BACKCAST stands for backcasting which helps in obtaining two initial error terms

ESTSMPL stands for estimated sample.

As discussed above, the ARIMA model obtained with the help of estimation equation is of the order of AR(1)MA(1). The AR(1)MA(1) model can also be confirmed with the help of correlogram using AC, PAC and *Q*-Stat, as given in Figure 2.

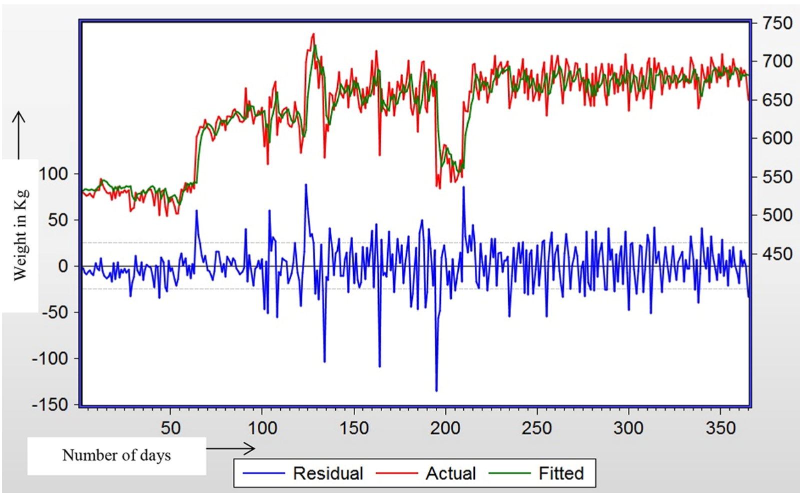
Table 4 shows that AR(1)MA(1) and co-efficient (c) are significant on 5% confidence level; therefore, it has been chosen as an ARIMA model for forecasting of HCW generation. In addition to this, Table 4 represents the other important statistics of a prediction model which are the Akaike information criteria (AIC), Schwarz criterion (SC), adjusted *R*-square and mean square error (MSE). Moreover, the adjusted *R*-square value shows the explanatory power of the model, which is 0.83 or 83% in the AR(1)MA(1) series of present study. In addition, the evaluation of the forecasting performance has also been carried out with the help of mean absolute percentage error (MAPE) value. In the present analysis, the MAPE value for the forecasted series can be observed from Table 3, i.e. 5.96% which shows the model is a good fit, as it is near to 0.

Table 4 Test statistics of forecasting time series

Variable	*Coefficient*	*STD. error*	*t-Statistic*	*Prob.*
C	648.705	24.417	26.568	0.000
AR(1)	0.973	0.012	80.757	0.000
MA(1)	−0.504	0.049	−10.301	0.000
R-squared	0.832	Mean dependent var		634.882
Adjusted *R*-squared	0.831	S.D. dependent var		60.214
MAPE	5.962			
S.E. of regression	24.739	Akaike info criterion		9.263
Sum squared resid	220933.500	Schwarz criterion		9.295
Log likelihood	−1682.834	Hannan-Quinn criterion.		9.276
F-statistic	894.769	Durbin-Watson stat		1.926
Prob(*F*-statistic)	0.000			
Inverted AR roots	0.970			
Inverted MA roots	0.500			

Figure 3 shows the graph between actual and estimated quantities of HCW generation, and the graph of residual is also shown in the figure.

Figure 3 Actual values versus ARIMA model fitted values (see online version for colours)



The adjusted *R*-square of the obtained ARMA model is 0.83 or 83%, which shows its fitness for the forecasting of HCW generation. In addition to this, the small value of mean square error and mean average percentage error shows the strength of the model. Furthermore, the positive behaviour of autoregressive (AR) and negative behaviour of moving averages (MA) (errors), as shown in Table 2, predict the strength of the model in terms of precise forecasting. The positive coefficient of AR indicates positive dependency on past data on waste generation, whereas the case reverses for errors lagged values recorded under MA.

The forecasting of HCW generation with an AR(1) MA(1) model would help the central bio-medical waste disposal firm in better planning of its infrastructure. The HCW disposal firm's infrastructure includes capacity planning-related decisions such as determining the waste disposal rate, hiring trained and educated manpower and hiring of specially modified HCW collection vehicles.

5 Conclusion

On the basis of the results obtained in the analysis of data, it can be concluded that AR(1)MA(1) model would provide the better forecasting of the generation of HCW, for the considered data set, in comparison with other models, i.e. AR(1)MA(2), AR(2)MA(1) and AR(2) MA(2), which have been tested for the study. In addition, with the help of AR(1) MA(1) forecasting model, the HCW disposal firm would plan in advance for the collection, treatment and disposal-related activities. Because the cost of purchase and operations of HCW collection vehicles and incinerators are very high, therefore, the pre-planning using the forecasting model may contribute significantly in diminishing expenses and increasing profits of the waste disposal firm. Moreover, the usage of forecasting model for pre-planning could result into an achievement of a win-win situation in terms of economic, social and environmental health for HCW disposal firm and various environmental bodies including pollution control board.

6 Implications of the study

The development of a forecasting model of HCW generation has a few direct and indirect impacts on the planning of central bio-medical waste disposal firm and the ministry of environment and forest, respectively. With the help of forecasting model, the ministry of environment and forests can assess, review, implement and form the HCW management rules for HCW disposal firms. Along with this, the environmental bodies can investigate in advance the sustainability of future strategy of HCW disposal firms, including the infrastructure and technology, which would directly benefit the society and environment.

7 Limitations and future research directions

This study has been conducted in a setting where the waste collection and waste disposal both operations have been completed by a single firm, and it may have left the void for the addressal of the situation of different parties working together for collection and disposal of HCW. Furthermore, the ARIMA model developed in the present study explains the linear part of the data very efficiently; therefore, this model may have limitations in dealing with non-linear data sets. The future studies may include the working upon the development of a hybrid model which could consider both the linear and non-linear components of the data sets more effectively. In addition, an ARIMA model can be developed in future for the separate forecasting of the generation of hazardous and infectious HCW. Along with this, the studies can be conducted in future using multi-criteria decision-making and statistical methods for the identification of more variables and their relationships with the quantities of HCW generation.

Acknowledgements

The authors of the present study are thankful to reviewers for devoting their valuable time and efforts in reviewing the article. With the help of reviewer's comments, a significant improvement has been noticed in the present work. Therefore, we are grateful to reviewers in enhancing the work and making it more lucid for the understanding of readers.

References

Al-Khatib, I.A., Abu Fkhidah, I., Khatib, J.I. and Kontogianni, S. (2016) 'Implementation of a multi-variable regression analysis in the assessment of the generation rate and composition of hospital solid waste for the design of a sustainable management system in developing countries', *Waste Management & Research*, Vol. 34, No. 3, pp.225–234. doi:10.1177/0734242X15622813

Alvim-Ferraz, M.C.M. and Afonso, S.A.V. (2005) 'Incineration of healthcare wastes: management of atmospheric emissions through waste segregation', *Waste Management (New York, N.Y.)*, Vol. 25, No. 6, pp.638–648. doi:10.1016/j.wasman.2004.07.017

Ananth, A.P., Prashanthini, V. and Visvanathan, C. (2010) 'Healthcare waste management in Asia', *Waste Management (New York, N.Y.)*, Vol. 30, No. 1, pp.154–161. doi:10.1016/j.wasman.2009.07.018

Box, G. and Jenkins, G. (1962) 'Some statistical aspects of adaptive optimization and control', *Journal of the Royal Statistical Society. Series B (Methodological)*, Vol. 24, No. 2, pp.297–343. Retrieved from http://www.jstor.org/stable/2984225

Box, G.E. and Jenkins, G.M. (1976) *Series Analysis: Forecasting and Control*, 4th ed., Holden-Day, San Francisco.

Box, G.E.P., Jenkins, G.M. and Reinsel, G.C. (2008) *Time Series Analysis*, 4th ed., John Wiley & Sons, Inc, Hoboken, NJ, pp.551–595.

Box, G.E.P. and Jenkins, G.M. (1965) 'Mathematical models for adaptive control and optimization', *In AI Chem. E Symposium Series*.

Campion, N., Thiel, C.L., Woods, N.C., Swanzy, L., Landis, A.E. and Bilec, M.M. (2015) 'Sustainable healthcare and environmental life-cycle impacts of disposable supplies: a focus on disposable custom packs', *Journal of Cleaner Production*, Vol. 94, pp.46–55. doi:10.1016/j.jclepro.2015.01.076

Caniato, M., Tudor, T. and Vaccari, M. (2015) 'International governance structures for health-care waste management: a systematic review of scientific literature', *Journal of Environmental Management*, Vol. 153, pp.93–107. doi:10.1016/j.jenvman.2015.01.039

Chaâbane, N. (2014) 'A hybrid ARFIMA and neural network model for electricity price prediction', *International Journal of Electrical Power & Energy Systems*, Vol. 55, pp.187–194. doi:10.1016/j.ijepes.2013.09.004

Chang, W-Y. (2014) 'A literature review of wind forecasting methods', *Journal of Power and Energy Engineering*, Vol. 2, No. 2, pp.161–168. doi:10.4236/jpee.2014.24023

Chatfield, C. and Prothero, D. (1973) 'Box-Jenkins seasonal forecasting: problems in a case-study', *Journal of the Royal Statistical Society. Series A (General)*, Vol. 136, No. 3, pp.295–336. Retrieved from http://www.jstor.org/stable/2344994

Chauhan, A. and Singh, A. (2016a) 'A hybrid multi-criteria decision making method approach for selecting a sustainable location of healthcare waste disposal facility', *Journal of Cleaner Production*, Vol. 139, pp.1001–1010. doi:10.1016/j.jclepro.2016.08.098

Chauhan, A. and Singh, A. (2016b) 'Healthcare waste management: a state-of-the-art literature review', *International Journal of Environment and Waste Management*, Vol. 18, No. 2, pp.120–144.

Chauhan, A., Singh, A. and Jharkaria, S. (2015) 'Healthcare waste management practices' identification and evaluation to rank hospitals', *International Journal of Operational Research*.

Chauhan, A., Singh, A. and Jharkharia, S. (2016) 'An ISM and DEMATEL method approach for the analysis of barriers of waste recycling in India', *Journal of the Air & Waste Management Association*, doi:10.1080/10962247.2016.1249441

Cristianini, N. and Shawe-Taylor, J. (2000) An Introduction to Support Vector Machines and Other Kernel-Based Learning Methods. Retrieved from https://books.google.co.in/books?hl=en&lr=&id=_PXJn_cxv0AC&oi=fnd&pg=PR9&dq=support+vector+machines&ots=xRSk9F3wZd&sig=ttLJbNt3aBmPqswJ1XbCOHsBkbc (access date:15-07-2016).

Cryer, J. and Kellet, N. (1986) *Time Series Analysis*, Duxbury Press, Boston. Retrieved from http://link.springer.com/content/pdf/10.1007/978-0-387-75959-3.pdf (access date: 12-07-2016).

DeLurgio, S.A. (1998) *Forecasting Principles and Applications*, McGraw-Hill, Boston.

Denafas, G., Ruzgas, T., Martuzevičius, D., Shmarin, S., Hoffmann, M., Mykhaylenko, V., Ogorodnik, S., Romanov, M., Neguliaeva, E., Chusov, A., Turkadze, T., Bochoidze, I. and Ludwig, C. (2014) Seasonal variation of municipal solid waste generation and composition in four East European cities. *Resources, Conservation and Recycling*, Vol. 89, pp.22–30. doi:10.1016/j.resconrec.2014.06.001

Dyson, B. and Chang, N. (2005) 'Forecasting municipal solid waste generation in a fast-growing urban region with system dynamics modeling', *Waste Management*, Vol. 25, No. 7, pp.669–679. Retrieved from http://www.sciencedirect.com/science/article/pii/S0956053X04001850

Granger, C.W. and Newbold, P. (1974) 'Spurious regressions in econometrics', *Journal of Econometrics*, Vol. 2, No. 2, pp.111–120.

Hassani, H., Webster, A., Silva, E. and Heravi, S. (2015) 'Forecasting US tourist arrivals using optimal singular spectrum analysis', *Tourism Management*, Vol. 46, pp.322–335. Retrieved from http://www.sciencedirect.com/science/article/pii/S0261517714001368

Haykin, S. (2004) *A comprehensive foundation*. *Neural Networks*, 2nd ed. NEURAL NETWORKS A Comprehensive Foundation Second Edition Simon Haykin McMaster University Hamilton, Ontario, Canada Prentice Hall Prentice Hall Upper Saddle River, New Jersey, 07458. Retrieved from http://ieeexplore.ieee.org/iel4/91/8807/x0153119.pdf (access date: 10-07-2016).

IndiaStat (2013) http://www.indiastat.com/table/environmentandpollution/11/solidwaste/261/910950/data.aspx (access 06 02 2016).

Intharathirat, R., Abdul Salam, P., Kumar, S. and Untong, A. (2015) 'Forecasting of municipal solid waste quantity in a developing country using multivariate grey models', *Waste Management*, Vol. 39, pp.3–14. doi:10.1016/j.wasman.2015.01.026

Karpusenkaite, A., Ruzgas, T. and Denafas, G. (2016) 'Forecasting medical waste generation using short and extra short datasets: case study of Lithuania', *Waste Management & Research*, Vol. 34, No. 4, pp.378–387. doi:10.1177/0734242X16628977

Khintchine, A. (1924) Einige Sätze über Kettenbrüche, mit Anwendungen auf die Theorie der diophantischen Approximationen. *Mathematische Annalen*. Retrieved from http://www.springerlink.com/index/pp854j162255g473.pdf (accessed date: 14-07-2016).

Komilis, D., Fouki, A. and Papadopoulos, D. (2012) 'Hazardous medical waste generation rates of different categories of health-care facilities', *Waste Management (New York, N.Y.)*, Vol. 32, No. 7, pp.1434–1441. doi:10.1016/j.wasman.2012.02.015

Li, Z., Rose, J.M. and Hensher, D.A. (2010) 'Forecasting automobile petrol demand in Australia: an evaluation of empirical models', *Transportation Research Part A: Policy and Practice*, Vol. 44, No. 1, pp.16–38. doi:10.1016/j.tra.2009.09.003

MacKinnon, J.G. (1996) 'Numerical distribution functions for unit root and cointegration tests', *Journal of applied econometrics*, pp.601–618.

Martínez-Costa, C., Mas-Machuca, M., Benedito, E. and Corominas, A. (2014) 'A review of mathematical programming models for strategic capacity planning in manufacturing', *International Journal of Production Economics*, Vol. 153, pp.66–85. doi:10.1016/j.ijpe.2014.03.011

McCleary, R., Hay, R.A., Meidinger, E.E. and McDowall, D. (1980) *Applied Time Series Analysis for the Social Sciences*, Sage Publications Ltd, Thousand Oaks, CA. Retrieved from https://www.ncjrs.gov/App/abstractdb/AbstractDBDetails.aspx?id=74039 (access date: 14-07-2016).

Mula, J., Poler, R., Garcia-Sabater, J. and Lario, F. (2006) 'Models for production planning under uncertainty: a review', *International Journal of Production Economics*, Vol. 103, No. 1, pp.271–285. Retrieved from http://www.sciencedirect.com/science/article/pii/S0925527306000041

Navarro-Esbri, J., Diamadopoulos, E. and Ginestar, D. (2002) 'Time series analysis and forecasting techniques for municipal solid waste management', *Resources Conservation and Recycling*, Vol. 35, No. 3, pp.201–214.

Nema, A., Pathak, A., Bajaj, P., Singh, H. and Kumar, S. (2011) 'A case study: biomedical waste management practices at city hospital in Himachal Pradesh', *Waste Management & Research : The Journal of the International Solid Wastes and Public Cleansing Association, ISWA*, Vol. 29, No. 6, pp.669–673. doi:10.1177/0734242X10396753

Oribe-Garcia, I., Kamara-Esteban, O. and Martin, C. (2015) 'Identification of influencing municipal characteristics regarding household waste generation and their forecasting ability in Biscay', *Waste Management*, Vol. 39, pp.26–34. Retrieved from http://www.sciencedirect.com/science/article/pii/S0956053X15001002

Oroei, M., Momeni, M., Palenik, C.J., Danaei, M. and Askarian, M. (2014) 'A qualitative study of the causes of improper segregation of infectious waste at Nemazee Hospital, Shiraz, Iran', *Journal of Infection and Public Health*, Vol. 7, No. 3, pp.192–198. doi:10.1016/j.jiph.2014.01.005

Peeters, J., Vanegas, P., Kellens, K. and Wang, F. (2015) 'Forecasting waste compositions: a case study on plastic waste of electronic display housings', *Waste Management*, Vol. 46, pp.28–39. Retrieved from http://www.sciencedirect.com/science/article/pii/S0956053X15301288

Pruss, A. (2014) 'Safe management of wastes from health care activities', in Chartier, Y., Emmanuel, J., Pieper, U., Prüss, A., Rushbrook, P., Stringer, R. and Zghondi, R. (Eds.), *2nd ed., World Health Organization*, Geneva, pp.1–242.

Pruss, A., Giroult, E. and Rushbrook, P. (1999) *Safe Management of Wastes from Healthcare Activities*, World Health Organisation, Geneva, Switzerland.

Rasouli, S. and Timmermans, H. (2012) 'Uncertainty in travel demand forecasting models: literature review and research agenda', *Transportation Letters*, Vol. 4, No. 1, pp.55–73. Retrieved from http://www.maneyonline.com/doi/abs/10.3328/TL.2012.04.01.55-73

Rimaityte, I., Ruzgas, T., Denafas, G., Racys, V. and Martuzevicius, D. (2012) 'Application and evaluation of forecasting methods for municipal solid waste generation in an Eastern-European city', *Waste Management & Research*, Vol. 30, No. 1, pp.89–98. doi:10.1177/0734242X10396754

Rongfang, Y., Kangzhou, W., Guowei, L. and Na, L. (2012) 'Analytical characterisations of a seasonal compound poisson demand process with growth trend', *International Journal of Services and Operations Informatics*, Vol. 7, No. 2, pp.167–181. doi:10.1504/IJSOI.2012.051395

SBC, U. (2013) *Technical Guidelines on the Environmentally Sound Management of Biomedical and Healthcare Wastes, Chatelaine*, Secretariat of the Basel Convention, Switzerland. Retrieved from http://www.basel.int/Portals/4/BaselConvention/docs/pub/techguid/tech-biomedical (access 10 March 2016).

Schuh, G., Kampker, A. and Wesch-Potente, C. (2011) 'Condition based factory planning', *Production Engineering*, Vol. 5, No. 1, pp.89–94. doi:10.1007/s11740-010-0281-y

Shamshiry, E., Mokhtar, M.B. and Abdulai, A. (2014) 'Comparison of Artificial Neural Network (ANN) and multiple regression analysis for predicting the amount of solid waste generation in a tourist and tropical area - Langkawi Island', *International Conference on Biological, Civil and Environmental Engineering (BCEE-2014)*, 17–18 March, Dubai (UAE), pp.161–166. Retrieved from http://iicbe.org/upload/2201C0314099.pdf

Shukla, M. and Jharkharia, S. (2011) 'ARIMA models to forecast demand in fresh supply chains', *International Journal of Operational Research*, Vol. 11, No. 1, pp.1–18. doi:10.1504/IJOR.2011.040325

Shumway, R.H. and Stoffer, D.S. (2011) 'ARIMA models', *Time Series Analysis and Its Applications*, Springer, New York, pp.83–171.

Song, J. and He, J. (2014) 'A multistep chaotic model for municipal solid waste generation prediction', *Environmental Engineering Science*, Vol. 31, No. 8, pp.461–468. Retrieved from http://online.liebertpub.com/doi/abs/10.1089/ees.2014.0031

Tran, Q., Sharp, A. and Nakatani, J. (2014) 'Forecasting the E-waste generation and its future trends in Vietnam', *(ENRIC2014) The 1st Environment and Natural Resources International Conference*, 6th–7th Nov. 2014, Bangkok, Thailand, pp.43–49, Retrieved from https://www.researchgate.net/profile/Thao_Tran49/publication/291766181_Forecasting_of_the_E-Waste_Generation_and_its_Future_Trends_in_Vietnam/links/56a5d9e608aeef24c58d9d1b.pdf

Tudor, T.L., Barr, S.W. and Gilg, A.W. (2007) 'Linking intended behaviour and actions: a case study of healthcare waste management in the Cornwall NHS', *Resources, Conservation and Recycling*, Vol. 51, No. 1, pp.1–23. doi:10.1016/j.resconrec.2006.06.009

USEPA (2016) United States Environmental Protection Agency. Retrieved from https://www3.epa.gov/ (access 28 May 2016).

Wang, W., Rothschild, D., Goel, S. and Gelman, A. (2015) 'Forecasting elections with non-representative polls', *International Journal of Forecasting*, Vol. 31, pp.980–991. doi:10.1016/j.ijforecast.2014.06.001

Weron, R. (2014) 'Electricity price forecasting: a review of the state-of-the-art with a look into the future', *International Journal of Forecasting*, Vol. 30, pp.1030–1081. doi:10.1016/j.ijforecast.2014.08.008

WHO (2013) WHO, Healthcare Waste Management: Documents. Available at: http://www.healthcarewaste.org/resources/documents/ (access 10 May 15).

Wold, H. (1939) 'A study in the analysis of stationary time series', *Journal of the Institute of Actuaries*, Vol. 70, No. 2, pp.119–269. Retrieved from http://www.jstor.org/stable/41137811

Xu, L., Gao, P., Cui, S. and Liu, C. (2013) 'A hybrid procedure for MSW generation forecasting at multiple time scales in Xiamen City, China', *Waste Management*, Vol. 33, No. 6, pp.1324–1331. doi:10.1016/j.wasman.2013.02.012

Yule, G.U. (1926) 'Why do we sometimes get nonsense-correlations between time-series? - a study in sampling and the nature of time-series', *Journal of the Royal Statistical Society*, Vol. 89, No. 1, pp.1–63.

Zhang, G.P. (2003) 'Time series forecasting using a hybrid ARIMA and neural network model', *Neurocomputing*, Vol. 50, pp.159–175. doi:10.1016/S0925-2312(01)00702-0

Zhang, G., Patuwo, B.E. and Hu, M.Y. (1998) 'Forecasting with artificial neural networks: the state of the art', *International Journal of Forecasting*, Vol. 14, No. 1, pp.35–62.