Journal of Multidisciplinary Engineering Science and Technology (JMEST)
ISSN: 2458-9403
Vol. 3 Issue 7, July - 2016

Modeling the Temporal Variations of Municipal Solid Waste Generation for Future Projection in the Douala Municipality, Cameroon

Ndoh Mbue Innocent
Department of Industrial Quality, Hygiene, Safety and Environment
University of Douala
Douala-Cameroon
Email: dndoh2009@gmail.com

BITONDO Dieudonné
Department of Industrial Quality, Hygiene, Safety and Environment
University of Douala
Douala-Cameroon

Sammy Nana Jose
Department of Industrial Quality, Hygiene, Safety and Environment
University of Douala
Douala-Cameroon

***Abstract*—As Cameroon hurtles toward its urban future, the amount of municipal solid waste is growing even faster than the rate of urbanization. Though the current waste management policy represents an interesting solution to the current waste problem in the country, questions remain regarding the prediction of waste's quantity as a crucial tool for planning, designing and programming for sustainable municipal solid waste management. This descriptive and exploratory study uses the Autoregressive Integrated Moving Average (ARIMA) time series model to explore the dynamics of solid waste generation and to forecast monthly solid waste generation in the Douala municipality of Cameroon. The study uses monthly municipal solid waste generation data from January 2003 to December 2014 obtained from the database of "Hygiène et Salubrité du Cameroun", the nation's private waste management enterprise. Minitab 17.1 and SPSS 20 statistical softwares were used to build a class of ARIMA models following the Box-Jenkins method. Following the distribution of the autocorrelation and partial autocorrelation functions of the log transformed differenced(1) series and the principle of parsimony, ARIMA(1,1,1):X_t =-.29$+.549\alpha_{t-1}+\alpha_t$ was identified. The model was subjected to statistical diagnostic check using the Ljung-Box Q test statistic and the Schwarz Bayesian Information Criterion (BIC). The analysis proved that the model is statistically significant, appropriate and adequate. The forecasted values indicated that by January of 2017, the monthly generation will attain 46159 tons with a 95% confidence interval lying between 4540 and 51214 tons. The methods and findings of this study may assist experts, decision-makers, planners and scientists involved in waste management.**

Keywords: ARIMA, Hygiène et Salubrité du Cameroun, Douala municipality, BIC, Ljung-Box Q test statistic

Introduction

As the world hurtles toward its urban future, the amount of Municipal Solid Waste (MSW), one of the most important by-products of an urban lifestyle, is growing even faster than the rate of urbanization. Commonly called "trash" or "garbage," Municipal Solid Waste (MSW), includes wastes such as durable goods (e.g., tires, furniture), non-durable goods (e.g., newspapers, plastic plates/cups), containers and packaging (e.g., milk cartons, plastic wrap), and other wastes (e.g., yard waste, food), but excludes industrial, hazardous, and construction wastes [1]. World production has almost doubled in the past decade and is expected to reach 2.5 billion tons per year in 2025 [2], under the combined effect of urban development and changing consumption patterns. The same authors projected that, globally, solid waste management costs will increase from $205.4 billion per year in 2012 to about $375.5 billion per year in 2025. They further claim that this will be most severe in low income countries (more than 5-fold increases) and lower-middle income countries (more than 4-fold increases), managing waste effectively and efficiently remains one of the most intractable problems for local authorities in urban centers.

In the Douala municipality of Cameroon where this research was carried out, municipal solid waste generation has been steadily increasing from less than 2,000 tons per day in the year 2000 to more than 4,000 tons per day in 2014 [3]. Despite the modern techniques put in place by the municipality to ensure optimal management of municipal solid wastes, solid waste management is still constrained by a number of setbacks ranging from the absence of environmentally reliable disposal sites to inadequate solid waste transportation vehicles to the current disposal site. Consequently, the urban landscapes is characterized by such environmental problems such as open spaces and roadsides littered with refuse, drainage channels and gutters choked with waste, and beaches strewn with plastic solid wastes, leading to inundations and floods, air pollution, and public health impacts such as respiratory ailments, diarrhea and dengue fever. Accurate and detailed forecasts of solid waste generation can play an important role in strategic planning with respect to collection, personnel, and truck utilization, transportation to the landfill and final disposal.

A number of forecasting and prediction approaches have been used for MSW management including derived probability distributions [4], Interval-Parameter Fuzzy-Stochastic Programming Approach [5,6], and the Two-Stage Interval-Stochastic Programming Model [7]. In recent years, attention has been turned to the use of Artificial Neural Networks (ANNs) models [8]. In terms of dynamic modeling, a significant number of research deals with analysis and prediction in environmental applications based on univariate models, taking the information in the form of time series [9]. In this context [10] applied the grey prediction technique to deal with the forecasting of solid waste generation when the number of samples is very few.

A number of other studies had employed the Autoregressive Integrated Moving Average models, ARIMA models [11]. It provides a convenient framework which allows an analyst to find an appropriate statistical model which could be used to answer relevant questions about the data. ARIMA models describe the current behaviour of variables in terms of linear relationships with their past values and have been proven to be relatively robust than more sophisticated structural models in terms of short-run forecasting ability [12]. ARIMA methodology of forecasting is different from most methods because it does not assume any particular pattern in the historical data of the series to be forecast. It uses an interactive approach of identifying a possible model from a general class of models.

Clearly there is no convergence in literature in terms of the method that can be best applied in the prediction of solid waste. However, because of the accuracy desired, the context of forecast, the relevance and availability of statistical data, the time period to be forecasted, easiness of interpretation and availability of guidelines from available literature, this study used the ARIMA time series model to explore the dynamics of solid waste generation and also forecast monthly solid waste generation in the Douala municipality of Cameroon. Despite the numerous successes registered by municipal authorities over the years, solid waste still remains one of the most pernicious local pollutants as a large quantity still remains uncollected. Accurate forecasting of municipal solid waste's quantity is seen by local experts as crucial for designing and programming municipal solid waste management system in the municipality. Specifically, the research aims to:

1. Explore the time series data of municipal solid waste generated in the Douala municipality to identify optimal parameters for accurate prediction

2. Build an Autoregressive Integrative Moving Average (ARIMA) model and use it to make predictions on the MSW generated in the municipality.

The study is expected to contribute to effective planning, cost effective and sustainable strategies for efficient solid waste collection, handling and disposal systems. Ultimately, the results of the study can be useful not only for future policy formulation and implementation but more importantly, for other cities that are experiencing similar solid waste management problems.

Materials & Methods

Study Area

The city of Douala, with a surface area of 210 km^2 /80 sq mi is the capital of the Littoral region of Cameroon. It is Cameroon's economic capital, the richest city in the whole CEMAC region of six countries, located on the banks of the Wouri River, at 4°02′53″ N Latitude 9°42′15″E Longitude, situated in the Wouri division at an average elevation of 13m above sea level[3]. Five urban municipalities (also known as districts) and one rural municipality form the urban community of Douala: the town districts of Douala I whose headquarters is at Bonanjo, Douala II whose headquarters is at New Bell, Douala III whose headquarters is at Logbaba, Douala IV whose headquarters is at Bonassama, Douala V whose headquarters is at Kotto, and Douala VI whose headquarters is at Manoka (Fig.1).

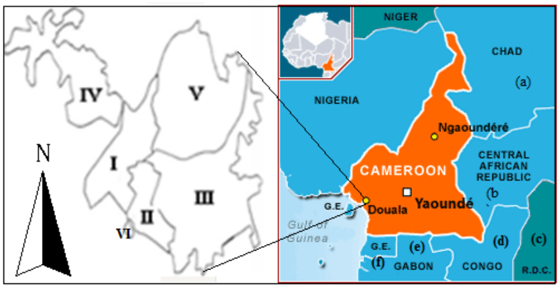


Figure 1. Location of the study area; "a-f": CEMAC countries

Together, the six local governments are commonly referred to as the Douala Urban Council (DUC). DUC is an industrial city and one of the fastest developing urban areas in Africa and ranks first at national level. Politically, the 2005 population census estimates the population to a controversial figure of 1.907 million [13]. However, according to current estimates from the Douala urban council, the population of the municipality is estimated to at 5,000,000 people with an average growth rate of 4.8%.

The city of Douala became the first city in Cameroon to outsource the management of its municipal solid waste to a private operator, "Hygiene et Salubrite du Cameroun" (HYSACAM), after having realized that the municipal solid waste management system was failing [14]. HYSACAM operates across the entire municipal solid waste management chain, from collection to processing.

Research design

Both descriptive and experimental research designs were explored to achieve the objectives of this study. These describe, examine relationships, and determine causality among variables, where possible. Statistical analysis was conducted to determine significant relationships and identify differences and/or similarities within and between different categories of data. According to [15], experimental research is often used where there is time priority in a causal relationship, consistency in a causal relationship, and also where the magnitude of the correlation is great.

Data Collection

Through their annual solid waste reports, HYSACAM office provided historic data on quantities of waste generated daily, weekly, monthly and yearly. The data represents the period from January 2003- December 2014, with a total of 144 samples. Additionally, an extensive review of the existing literature on solid waste management was carried out to understand the current state-of – the - art knowledge.

Model Development Procedure

The general non-seasonal model, ARIMA (*p, d, q*) was the starting point of the modeling process:

- A *p*th-order autoregressive model, or AR(p), takes the form:

$$Y_t = \phi_0 + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \ldots + \phi_p Y_{t-p} + \varepsilon_t \quad [1]$$

Where:

Y_t = response variable at time t

Y_{t-k} = observation (predictor variable) at time $t-k$

ϕ_i = regression coefficients to be estimated

ε_t = error term at time t

- *d* is the number of differences. ARIMA (p, 0, q) = ARMA (p, q). A model with autoregressive terms can be combined with a model having moving average terms to get an ARMA (p, q) model:

$$Y_t = \phi_0 + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \ldots + \phi_p Y_{t-p} + \varepsilon_t - \omega_1 \varepsilon_{t-1} - \omega_2 \varepsilon_{t-2} - \ldots - \omega_q \varepsilon_{t-q} \quad [2]$$

ARMA (p,q) models can describe a wide variety of behaviors for stationary time series.

- A *qth*-order moving average model, or MA (q), takes the form:

$$Y_t = \mu + \varepsilon_t - \omega_1 \varepsilon_{t-1} - \omega_2 \varepsilon_{t-2} - \ldots - \omega_q \varepsilon_{t-q} \quad [3]$$

Where:

Y_t = response variable at time t

μ = constant mean of the process

ω_i = regression coefficients to be estimated

ε_{t-k} = error in time period t - k

In addition to the non-seasonal ARIMA (p, d, q) model, is the seasonal ARIMA (P, D, Q) parameters for the data were also identified:

- Seasonal autoregressive (P),
- Seasonal Differencing (D) and
- Seasonal moving average (Q).

The general form of the above model describing the current value X*t* of a time series by its own past is:

$$(1-\emptyset_1\beta)(1-\propto_1\beta^{12})(1-\beta^{12})X_t = (1-\theta_1\beta)(1-\gamma_1\beta^{12})\epsilon_t \quad [4]$$

Where:

$1-\emptyset_1\beta$ is the non seasonal autoregressive of order 1
$1-\propto_1\beta^{12}$ is the seasonal autoregressive of order 1
X_t is the current value of the time series examined
$1-\theta_1\beta$, the non-seasonal moving average of order 1
β is the backward shift operator:

$$\beta X_t = X_{t-1} \text{ and } \beta^{12} X_t = X_{t-2}$$

1- β = First order non-seasonal difference
$1-\beta^{12}$ = Seasonal difference of order 1
$1-\gamma_1\beta^{12}$ = Seasonal moving average of order 1

ARIMA modeling was developed using the Statistical Package for the Social Sciences (SPSS) version 20 software and Minitab version 17.1 according to the following five –step algorithm:

In the preliminary stage, the data was explored to identify and eliminate possible cyclical and seasonal behavioral patterns in the municipal solid waste data, as they frequently exhibit such behaviors. The objective was to find the integer values of p, d, and key questions answered at this stage included:

a) Is there a trend, or on average, do the measurements tend to increase (or decrease) over time?
b) Is there seasonality- a regularly repeating pattern of highs and lows related to calendar time such as seasons, quarters, months, days of the week, and so on?
c) Is there a long-run cycle or period unrelated to seasonality factors?
d) Is there constant variance over time, or is the variance non-constant?, and
e) Are there any abrupt changes to either the level of the series or the variance?

Using descriptive statistics, correlograms plots (autocorrelation and partial autocorrelation functions), multiplicative time series decomposition and autocorrelation plots [16] such worries were thoroughly explored for stationery behaviours. An autocorrelation function (ACF) shows the serial correlation coefficients for consecutive lags while partial autocorrelation function (PACF) partials out the immediate autocorrelations and estimates the autocorrelation at a specific lag.

In the second stage (model identification) stationery behaviours were eliminated. The following techniques were used to stationarise the time series:

a) Detrending: This is simply the removal of the trend component from the time series:

$$X_t = (mean + trend * t) + \varepsilon \qquad [5]$$

Where,
X_t= MSW generated at time t, and
$t = time$
ε = error term: $\varepsilon \sim N(0, \sigma^2)$,

The part in the parentheses was removed and the rest was used to build an appropriate model for the data.

b) Seasonality: Because data is likely to exhibit a kind of seasonal pattern that is not stable over time, considerations were made for the possible addition of sAR term to the model in case the autocorrelation of the appropriately differenced series becomes positive at lag s, where s is the number of periods in a season. It is worth mentioning that if the autocorrelation of the differenced series is negative at lag s, an sMA term should be added to the model. This situation is likely to occur if a seasonal difference has been used, which should have been done in this analysis if the data had a stable and logical seasonal pattern.

c) Differencing: This technique was used to remove non-stationarity in the data. Non-stationary stochastic process is indicated by the failure of the estimated autocorrelation functions to die out rapidly. To achieve stationarity, a certain degree of differencing (d) is required. In this paper, this was done by fitting the first order AR model to the raw MSW data to test whether the coefficients φ is less than one. The objective was to identify an appropriate sub-class of model from the general ARIMA family (Eq. 1).

$$\emptyset(B)\nabla^d Z_t = \theta(B)\alpha_t \qquad [6]$$

The degree of differencing (d), necessary to achieve stationarity is attained when the autocorrelation functions (Eq. 4) die out fairly quickly.

$$X_t = (1 - \mathrm{B})^d = \nabla^d Z_t \qquad [7]$$

The autocorrelation function of an AR (p) process tails off, while its partial autocorrelation function has a cut off after lag p. Conversely, the ACF of a MA (q) process has a cut off after lag q, while its partial autocorrelation function tails off. However, if both the ACF and PACF tail off, a mixed ARMA (p,q) process is suggested. The ACF of a mixed ARMA (p,q) process is a mixture of exponentials and damped sine waves after the q-p lags. Conversely, the PACF of a mixed ARMA (p,q) process is dominated by a mixture of exponentials and damped sine waves after the first p-q lags. In this paper, the time series was differentiated until a rapidly decaying Autocorrelation Function (ACF) compatible with that of an ARIMA process was obtained [16].

Once a stationary series had been obtained, an optimal model was built in stage 3 (Model Estimation) by comparing the sample ACF and PACF plots to the theoretical ACF and PACF for the various ARIMA models. The principle of parsimony was employed for the final model selection.

The fourth stage was a diagnostic stage, which is, assessing to see how well your model fits your data. The objective was to find optimal parameters for the model. We assess this through ACFs and PACFs. Several (p,d,q) combinations were explored to ensure that values found in previous section were not just approximate estimates. The Bayesian Information Criterion, BIC [17], (Eq. 2), was employed for model selection among the finite set of models under test.

$$BIC = n\left[ln\left\{\frac{1}{n-1}\sum_{i-1}^{n}(x_i - \bar{x})^2\right\}\right] + k[\{\ln(n)\}] \qquad [8]$$

Where,

n = The number of data points/observations, or the sample size;

k = The numbers of free parameters to be estimated,

$\hat{\sigma}^2{}_e = \frac{1}{n-1}\sum_{i-1}^{n}(x_i - \bar{x})^2$ = The error variance, and

x = The observed data (here, MSW generated)

To test the overall randomness of the model (or independence of residuals), the Ljung-Pierce Q-statistics [18] was employed (Eq 3).

$$Q = T(T+2)\sum_{k=1}^{s}\frac{r^2{}_k}{T-K} \qquad [9]$$

Where
T= number of observations
s = length of coefficients to test autocorrelation
r_k = Autocorrelation coefficient (for lag k)

If the sample value of Q exceeds the critical value of a χ^2 distribution with *s* degrees of freedom, then at least one value of *r* is statistically different from zero at the specified significance level.

As a general rule, given any set of estimated models, the model with the lower value of BIC is the one to be preferred. In addition to the residual plots and Ljung-Pierce Q-statistics, the R-squared, Root Mean Square Error (RMSE), Mean Absolute Percentage Error (MAPE) were used to check on the efficiency of the model. Usually the fitting process is guided by the principal of parsimony, by which the best model is the simplest possible model – the model with the fewest parameters that adequately describes the data.

Once the final ARIMA model was found, it was then used in a final stage (stage 5) to make predictions on the future time points for MSW. Prediction intervals based on the forecasts were also constructed.

Results and Discussions

Exploring the MSW data for a pattern

From January 2003 to December 2013, the minimum quantity of MSW generated in the Douala municipality was 14554 tons and this occurred in the month of January, 2003. The maximum amount of generated was approximately 43417 tons which occurred in August, 2013(table 1).

Table 1: Summary Statistics of Solid Waste Generated in the DM

Minimum	1st Quartile	Median	Mean	Std. deviation	3rd Quartile	Maximum	Skewness	Kurtosis
14554	21756.50	31183.00	29792.83	8109.883	36980.25	43417	.005	-1.451

The median > mean signifying a right skewed distribution (Skewness = .005). The kurtosis= -1.451< 3 signifying a near platykurtic distribution (a bit flatter than a normal distribution with a wider peak). This implies that the probability for extreme values is less than for a normal distribution, and the values are wider spread around the mean. Clearly, this indicates a strict deviation from normality (Fig. 2).

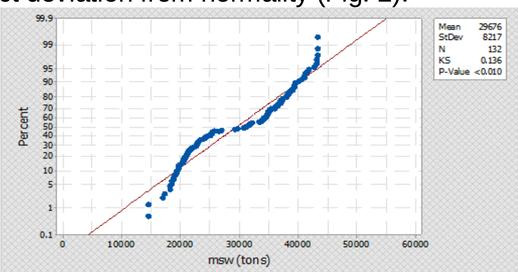


Figure 2: Normal probability distribution plot for MSW data, X_t

A visual observation of the MSW correlogram (Fig 3) shows an increasing trend with fluctuations across time, indicating that data is non-stationary.

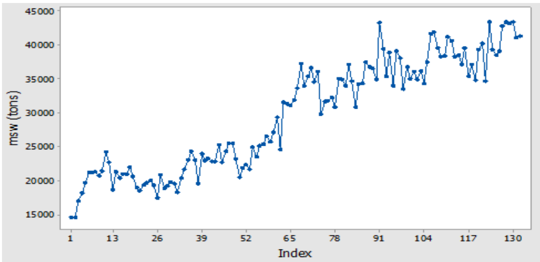
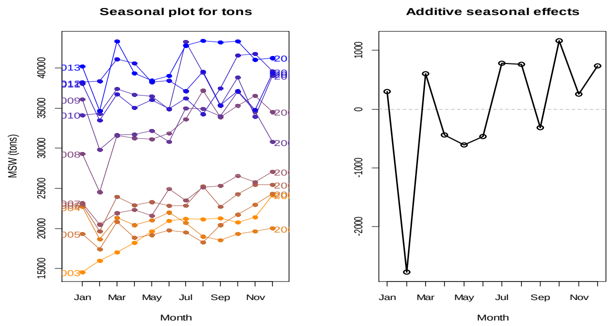


Figure 3: Graph of original series X, of MSW (2003-2014)

It can be inferred from the correlogram that: (1)the month to month trend clearly shows that the MSW generations have been increasing without fail, (2) there seem to be the presence of trend in the mean since the left hand side of the plot is lower than the right hand side. The fluctuation differences also suggest trend in variance, and (3) there is no evidence of seasonal components since no regular peaks and troughs are observed.

The stacked annual plot with the accompanying additive seasonal effects (Fig. 4) shows that there is a fairly consistent month on month variation with the months from July to October as the peak months for MSW generation..



(a) (b)

Figure 4: Seasonal plot for the waste data, X_t

It is observed that the seasonal pattern is quite similar from year to year (Fig. 3a), and that the MSW generated over the decade are gradually increasing (monotonic increasing). If they weren't, the lines would be on top of each other (all mushed together).

The data was further decomposed to (multiplicative time series decomposition) to decipher the underlying patterns (cyclical, exponential, seasonal) in MSW generation data using the following model (Eq.4).

$$Y_t = Trend_t * Seasonality_t * Remainder_t \qquad [4]$$

Figure 5 clearly shows that the data exhibits a non-systematic linear trend but the existence of seasonality is suggested.

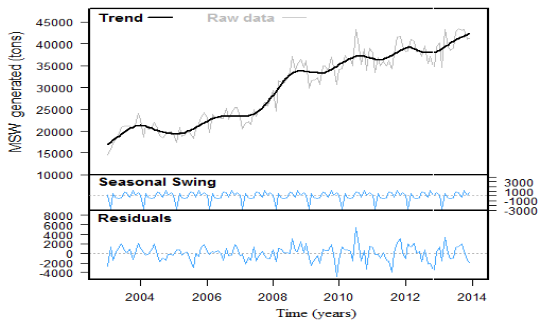


Figure 5: Decomposition of multiplicative MSW generated time series.

This behaviour could be as a result of irregular components in the time series. An alternative to this model is the additive decomposition time series model. However, it would have made very little difference in terms of conclusion drawn from this time series decomposition data. Additionally, plain vanilla decomposition models like these are rarely used for forecasting. Their primary purpose is to understand underlying patterns in temporal data to use in more sophisticated analysis like Holt-Winters seasonal method or ARIMA.

Model Identification

The sample autocorrelations of the original series (Fig 6) failed to die out at high lags, with significant Ljung-Box Q(18)($X^2{}_{(d=18)}$ =1731.568, p=.000), confirming the non-stationarity behaviour of the series, and a weak ARIMA model.

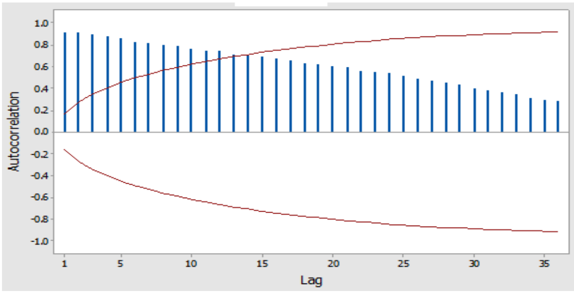


Figure 6: Residual plot of ACF for MSW generated, ARIMA (0, 0, 0) with a constant

Seasonality usually causes the series to be nonstationary because the average values at some particular times within the seasonal span (months, for example) may be different than the average values at other times. This syndrome can be treated by either the difference or logarithmic methods of transformations or both. Firstly, the series was transformed using the first order difference method (d = 1) and stationary was attained (Figure 7).

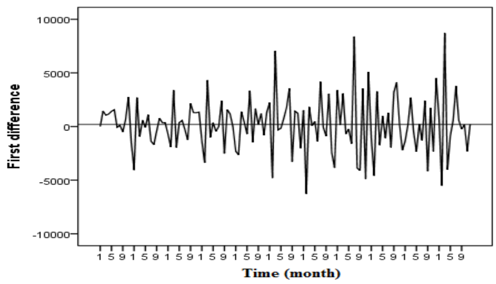


Figure 7 First Difference natural log of Solid waste Data, X_t

From the figure, it can be seen that the differenced series looks stationary, as the observations somehow beat about a constant mean. However, the series seem not to be stationary on variance i.e. variation in the plot is increasing as we move towards the right of the chart. Consequently, the series was transformed by taking the second differences of the natural logarithms of the values in the series so as to attain stationarity in the second moment (Eq. 5).

$$X_t^{new} = log_e(X_t) = log_e(X_t) - log_e(X_{t-1}) \quad [5]$$

Where,

$$X_t = \phi_0 + \phi_1 X_{t-1} + \phi_2 X_{t-2} + ... + \phi_p X_{t-p} + \varepsilon_t$$

Following this transformation, the data became stationary on both mean and variance (Fig 8).

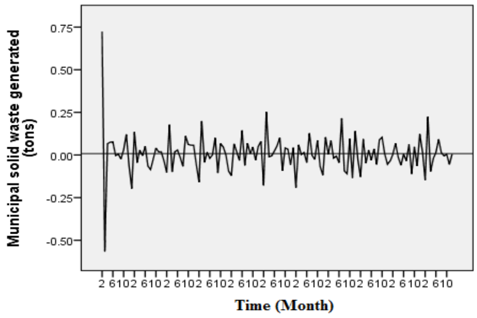
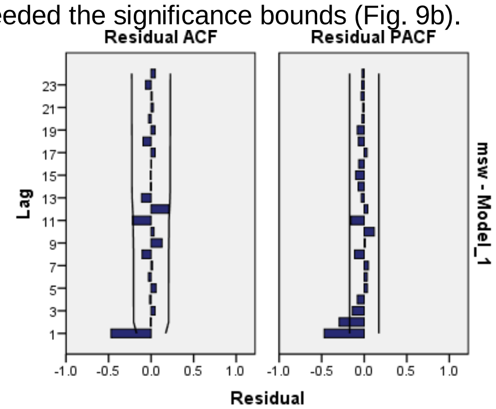


Figure 8: Transforms:Natural log, Difference (1)

Clearly, the series looks stationary on both mean and variance. This also gives us the clue that I or integrated part of our ARIMA model will be equal to 1 as first difference is making the series stationary. The autocorrelation and partial autocorrelation functions of the log, differenced series indicated no need for further differencing as they tend to be tailing off rapidly. They also indicated no sign of seasonality since they do not repeat themselves at lags that are multiples of the number of periods per season.

Model Identification

For the first log differenced series, ARIMA (p, 1, q) are considered where d=1 is the order of differencing. Comparing the resulting autocorrelation functions with their error limits, significant autocorrelation was observed at lag 1(Fig 9a). Partial autocorrelations were observed at lags 1 and 2 as the spikes exceeded the significance bounds (Fig. 9b).



(a) (b)

Figure 9. Sampled ACF (a) and PACF (b) for differenced (1) Solid Waste Data

Clearly, there is no need for sARIMA as patterns in the log difference (1) data suggests a variety of ARIMA models including ARIMA(2,1,2), ARIMA (2, 1, 1), ARIMA (2, 1, 0), ARIMA(1,1,1) and ARIMA(1,1,2) as potential fits for data.

Model Diagnosis and Selection

The models efficiencies were evaluated using the R-squared, Root Mean Squared Errors (RMSE), Mean Absolute Percentage Error (MAPE), the normalized BIC and the Ljung-BoxQ statistics (table 2).

Table 2. Evaluation of various ARIMA models

Model	Model statistics				Ljung-BoxQ(18)			Outliers
	R-squared	RMSE	MAPE	Normalized BIC	Statistics	DF	Sig.	
ARIMA(2,1,2)	.890	2715.80	6.36	16.037	13.469	14	.490	0
ARIMA(2,1,1)	.890	2704.20	6.38	15.991	13.813	15	.540	0
ARIMA(2,1,0)	.880	2714.74	6.53	15.962	17.711	16	.341	0
ARIMA(1,1,2)	.890	2705.83	6.36	15.992	13.632	15	.554	0
ARIMA(1,1,1)	.890	2698.59	6.38	15.950	15.447	16	.492	0

All ACF and PACF residuals of fitted well in all models. All the estimated coefficients are significantly different from zero and the root mean square errors (RMSE), MAPE, R-squared and BIC are similar for all models. In addition, there were no white noises at 5% significance limit for all models as there were no spikes outside the insignificant zone for both ACF and PACF plots. Furthermore, all residuals were independent, identically distributed and were therefore adequate for the observed data. Therefore, either model is adequate and provides nearly the same three-step-ahead forecasts.

Since the AR (1) model has two parameters (including the constant term) and the MA(2) model has three parameters, applying the principle of parsimony the simpler AR(1) model was selected to forecast future readings. Moreso, ARIMA (1, 1, 1) outperformed other potential models in terms of MAPE, and RMSE and BIC measures. The coefficient of both the AR and the MA were not significantly different from zero with values of -.298 and .549 respectively (Table 3).

Table 3: ARIMA (1, 1, 1) model parameters

Variable	Estimate	SE	t	Sig.
Constant	-.003	.009	-.343	.732
AR1	-.298	.118	-2.533	.013
MA1	.549	.104	5.26	.000

Hence, the forecasting model for X_t , should be of the form:

$$\hat{X}_t = \mu + (\emptyset X_{t-1} + \cdots + \emptyset_p X_{t-1}) - (\theta_1 e_{t-1} + \cdots + \theta_q e_{t-q})$$

Where:

μ *is a constant term*,
$(\emptyset X_{t-1} + \cdots + \emptyset_p X_{t-1})$ = AR terms (lagged value of x), and
$-(\theta_1 e_{t-1} + \cdots + \theta_q e_{t-q})$ = MA terms (lagged errors).

Usually p + q ≤ 2 and either p = 0 or q = 0 (pure AR or pure MA model). By convention, the AR terms are (+) while the MA terms are (-). This model enables us to write the model equation as:

$$X_t = -.298t_{t-1} + .549\alpha_{t-1} + \epsilon_t \quad [7]$$

The low value of RMSE (2698.550) indicates a good fit for the model. Also, the high value of the R-squared (.890) and MAPE (6.382) indicate a perfect prediction over the mean.

A further look at the plots of residuals, ACF and PACF (Fig 10) reveal a random variation from the origin (0). The points below and above are all uneven, hence, the model fitted is adequate.

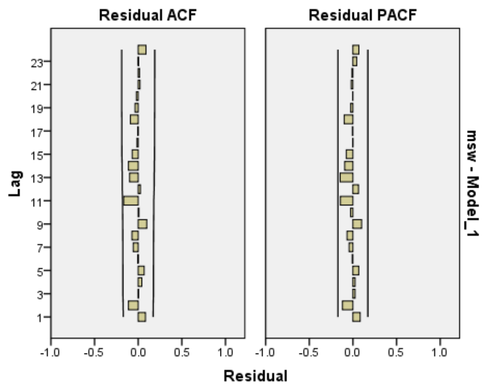


FIGURE 10. AUTOCORRELATION & PARTIAL AUTOCORRELATION FUNCTIONS OF THE RESIDUALS

The adequacy and significant appropriateness of the model was confirmed by exploring the normalized Bayesian Information Criterion (BIC). As Indicated, amongst the set of statistically significant ARIMA (P,D,Q) models fitted to the series, ARIMA (1,1,1) model had the least BIC value of 15.950. Hence, the model seems adequate and appropriate for the MSW data. The plot that follows (Fig 12) represents the actual series and the fitted ARIMA (1, 1, 1).

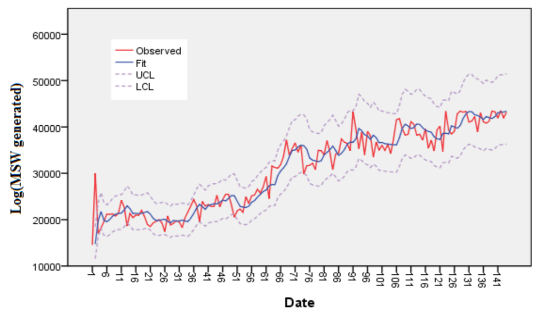


Figure 12: Plot of observed and fitted ARIMA (1,1,1) model

Our model agrees with results from recent studies [19, 20, 21] who also suggested that ARIMA (1, 1, 1) model is the best among all parametric time series models for forecasting solid waste generation.

FORECAST OF MSW GENERATED USING THE BEST FIT ARIMA MODEL

The forecasted values indicate that by January of 2017, the monthly generation according to the model will attain 46159 tons with a 95% confidence interval lying between 4540 – 51214 tons (Fig. 11; Table 4)

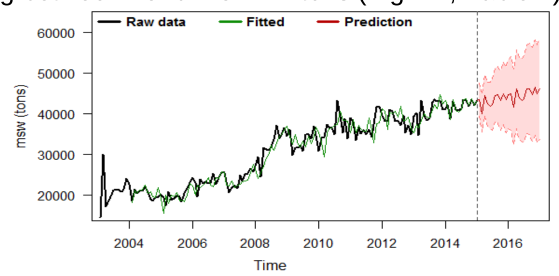


Figure 11: Holt-Winters Additive prediction for tons of MSW generated

Table 4. Three-step forecast of the ARIMA (1, 1, 1) Model

	Jan	Feb	Mars	Apr.	May	Jun	Jul	Aug.	Sep.	Oct.	Nov.	Dec.
2015	40958	43550	40123	44626	42384	41920	42255	44597	44696	43342	45001	43437
2016	44685	45025	41598	46101	43859	43394	46071	46071	46178	44817	46475	44911
2017	**46159**	47791	43156	48322	45587	45853	46118	46384	46649	43542	43808	44073

From figure 11, the forecasted solid waste values (in tonnage) are shown by the thicker pink sinusoidal curve, whilst the bounded light pink shaded region areas show 80% and 95% prediction intervals respectively.

Conclusion

One of the major challenges in many developing economies is to develop and promote sustainable solid waste management systems that could anticipate and take into account the quantity of solid waste that could be generated in the foreseeable future. This study analyzed, compared and selected the best time series model for forecasting amount of solid waste generated in the Douala municipality of Cameroon among ARIMA models. The past data used are monthly amount of solid waste collected by the city waste management enterprise (HYSACAM) from year January 2003 to January 2014.The result indicated that ARIMA (1, 1, 1) outperformed other potential models in terms of MAPE, R-squared, RMSE, BIC and AIC measures and hence used to forecast the amount of the solid waste generation for the next years. The forecasted values indicate that by January of 2017, the monthly generation according to the model will attain 46159 tons with a 95% confidence interval lying between 4540 – 51214 tons. The model is validated and could be adequate for forecasting solid waste generation in the foreseeable future. The paper recommends that since solid waste generation is a function of population growth, planning and implementing a comprehensive program, such as an integrated solid waste management system for waste planning, collection, transport, and disposal along with activities to prevent or recycle waste can eliminate the numerous problems which these wastes do cause. An integrated solid waste management (ISWM) program can help reduce greenhouse gas emissions and slow the effects of climate change.

References

[1]. U.S. Environmental Protection Agency (EPA). Municipal solid waste generation, recycling, and disposal in the United States: Facts and figures. 2015. Accessed June 2016. Available: https://www.epa.gov/

[2]. Hoornweg, D. and Bhada-Tata, P. What a Waste - Global review of Solid Waste Management. World Bank; 2012

[3]. Mbue NI, Bitondo D, Balgah RA. Municipal Solid Waste Generation, Composition, and Management in the Douala Municipality, Cameroon. Journal of Environment and Waste Management. 2015;2(3): 091-101.

[4]. Wilson, B.G.; Baetz, B.W. Modeling Municipal Solid Waste Collection Systems Using Derived Probability Distributions I: Model Development; ASCE J. Environ. Eng. 2001; 127, 1031-1038.

[5]. Huang, G.H.; Sae-Lim, N.; Liu, L.; Chen, Z. An Interval-Parameter Fuzzy-Stochastic Programming Approach for Municipal Solid Waste Management and Planning; Environ. Model. Assess. 2001; 6, 271-283.

[6]. Li, Y.P. and Huang, G.H. Minimax regret analysis for municipal solid waste management: An interval-stochastic programming approach, J. Air Waste Manage. Assoc. 2006a;56, 931-944

[7]. Maqsood, I.; Huang, G.H. A Two-Stage Interval-Stochastic Programming Model for Waste Management under Uncertainty; J. Air & Waste Manage. Assoc. 2003; 53, 540-552.

[8]. Karaca, F. and Özkaya, B. NN-LEAP: A neural network-based model for controlling leachate flow-rate in a municipal solid waste landfill site. Environ. Model. Software. 2006; **21,** 1190-1197.

[9]. Katsamaki A. Willems S. Diamadopoulos E. Time series analysis of municipal solid waste generation rates. J Environ Eng 1998; 124: 178-183.

[10]. Chen, H.W. and Chang, N.B. Prediction Analysis of Solid Waste Generation Based on Grey Fuzzy Dynamic Modelling, Resources, Conservation and Recycling. 2000; 29, 1–18.

[11]. Box- Jenkins. Time Series Analysis: Forecasting and Control. Wiley Series in Probability and Statistics. 1976

[12]. Litterman, R. B. Forecasting with Bayesian Vector Autoregressions-five Years of experience. Journal of Business & Economic Statistics, 1986; 25-38.

[13]. Parrot, L.,Sotamenou, J., Kamgnia Dia, B. Municipal solid waste management in Africa: strategies and livelihoods in Yaoundé, Cameroon, Waste Management 2009; 29(2), 986-995. USA.

[14]. UNdata. UNSD Demographic Statistics. 2013. Accessed 15/05/2015 ; Available: http://data.un.org/

[15]. Leedy, P. D. Practical research planning and design (6th ed.). Upper Saddle River, NJ: Prentice-Hall, Inc., 1997; 232-233.

[16]. Brockwell P.J. and R.A. Davis. Introduction to Time Series and Forecasting. Springer. 2002

[17]. Schwarz, G. E. Estimating the dimension of a Model. Annals of Statistics. 1978:6(2): 461-464.

[18]. Ljung, G., and Box, G. E. P. On a Measure of lack of fit in Time Series Models. Biometrika. 1978; 65:553-564

[19]. Navarro-Esbrí J, Diamadopoulos E, Ginestar D. Time series analysis and forecastingtechniques for municipal solid waste management. Resour Conserv Recy 2002; 35:201-214.

[20]. Ebenezer O-S, Emmanuel H, Ebenezer B. Forecasting and Planning for Solid Waste Generation in the Kumasi Metropolitan Area of Ghana: An ARIMA Time Series Approach. International Journal of Sciences. 2013. 2; PP 1-8

[21]. Amon M, Dmitry K, Silas M. Time Series Forecasting of Solid Waste Generation in Arusha City- Tanzania. Mathematical Theory and Modeling. 2014;4(8); 29-39