Forecasting and Cross-Correlation of Series on the Waste Generation, Population, GDP and Household Consumption Expenditures in Jordan

Omar Alhanaqtah[1]

[1] College of Business, Tafila Technical University, Tafila, Jordan

Correspondence: Omar Alhanaqtah, Department of Economics, College of Business, Tafila Technical University, Tafila, PO Box 66110, Jordan. Tel: 962-77-211-0181. E-mail: omarhanaqtah@yahoo.com

Received: December 18, 2023 Accepted: January 5, 2024 Online Published: January 12, 2024

doi:10.5430/ijfr.v15n1p1 URL: https://doi.org/10.5430/ijfr.v15n1p1

Abstract

The household waste is the biggest contributor to the total municipal solid waste generation. The main objective of the research is to discover the Jordanian economy in four dimensions: the household waste generation, the number of population, the gross domestic product (GDP) at purchasers' prices, and the household final consumption expenditures. In the focus of interest are (1) finding adequate time series models for forecasting of the number of population, the GDP at purchasers' prices, the household final consumption expenditures, and the household waste generation; (2) the cross-correlation between the household waste generation variable and other variables – population, GDP and household final consumption expenditures. The analysis of cross-correlation functions has revealed that pairs of series (waste/gdp, waste/consumption, waste/population) move in one direction with no shift in time of one series to the other. The relationship in every pair of series is significant. It is expected that in the short-run, ceteris paribus, with the growth of the population numbers, GDP and/or household consumption expenditures, the level of household waste generation in the country will also grow in a close relationship. Appropriate ARIMA models have been proposed to use for the short-run forecastsing of time series. The research outcomes are useful for policy makers to realize the scale of the household waste problem and to optimize capital expenditures into the waste management system of Jordan.

Keywords: ARIMA models, consumption, cross-correlation, GDP, log-linear regression, population, waste

1. Introduction

Million tons of the household waste are produced in Jordan every year. Along with other factors this situation has been dramatically worsened by the alerting refugee problem. Jordan has become a safety island in the economically and politically unstable Middle East. The household waste is the biggest contributor to the total municipal solid waste generation.

There is a number of influencing factors on waste generation, revealed by different researches. Many authors indicate the individual income, economic growth, population, the household's size and consumption patterns among the typical factors affecting the waste generation in the country (Abdoli, Falahnehzad, & Behbouldian, 2011; Afroz, Hanaki, & Tudin, 2010; Cheng, Shi, Yi, & Fu, 2020; Chung, 2010; Hage, Sandberg, Söderholm, & Berglund, 2018; Herianto, Maryono, & Budihardjo, 2019; Hockett, Lober, & Pilgrim, 1995; Liu, Li, Gu, & Wang, 2019; Lu, Zhang, Hai, & Lei, 2017; Samson, Akinlabi, Muzenda, & Aboyade, 2017).

Liu & Wu (2010) mentioned the following influencing factors for urban households: population, annual per capita disposable income of urban households, annual per capita consumption expenditure, total energy consumption, per capita GDP, the number of the large cities. They focused on urban development as an important factor influencing municipal solid waste generation.

Han et al. (2018) considered influencing factors on waste generation in rural areas: economic factors (types of industry in villages, energy and fuel structure, household income and expenditure); social factors (population, education (environmental education, training and demonstration projects), culture (traditional and national cultures, consumption and living habits)); natural factors (geography and climate (including humidity, rainfall, temperature, harvest seasons); other factors (the administrative levels of communities and survey methods).

The results of study of Zhao, Diunugala, & Mombeuli (2021) showed that the personal attitude, a household's income, age and assets are the most important factors influencing waste generation. Liu, Li, Gu, & Wang (2019) focused on the unique consumption patterns impacting solid waste generation. The authors pointed out that the housing rent, being a large part of household expenditures, impacts greatly the actual purchasing power parity of households that in turn leads to different waste generation patterns in different regions. In other words, the relationship between the financial expenditure on housing and the waste generation is negative, while, for example, the relationship between the household expenditures on food and the waste generation is positive. This study confirms the importance of the household expenditure and consumption factors for the waste generation.

In the research for Jordan, Alhanaqtah (2020) found out that the consumption is the most informative factor for the waste generation. This factor is highly correlated with the GNI and the population numbers. In order to avoid multi-collinearity, the author selects consumption as an input variable for the model estimation. In the study for Jordan, Aldayyat, Saidan, Abu Saleh, & Hamdan (2019) pointed out that the population number is a significant factor influencing waste generation in the country. The authors admitted a great refugee problem in Jordan, and the adverse effect of growing numbers of Syrian refugees on the solid waste generation in the country.

Demographic and socio-economic data to model solid waste generation are used in the works of Atieno, Oindo, & Bosire, (2017), Chhay, Reyad, Suy, Islam, & Mian (2018), Khan, Kumar, & Samadder (2016), Prades, Gallardo, & Ibàñez (2015) and Trang, Dong, Toan, Hanh, & Thu (2017). Abir, Datta, & Saha (2023) indicated institutional, cultural, socio-economic, technical factors as having strong relationship with the waste generation. Socio-economic factors are also essential in the work of Adeleke, Akinlabi, Jen, & Dunmade (2021).

The main purpose of the article is to discover the Jordanian economy in four dimensions: the waste generation, the number of population, the gross domestic product (GDP) at purchasers' prices, and the household final consumption expenditures. The particular tasks are: (1) finding adequate time series models for forecasting of the number of population, the gross domestic product (GDP) at purchasers' prices, the household final consumption expenditures, and the household waste generation; (2) to study cross-correlation between the waste generation variable and three other variables.

The research outcomes are useful for policy makers to realize the scale of the household waste problem and to optimize capital expenditures into the waste management system of Jordan.

The paper is arranged as follows. The second part describes the methodology of the research, the research plan and the variables used in the analysis. Analysis of the data trends on the population, GDP, private consumption and the waste generation is conducted in the first paragraph of the third part. Time series modelling of four mentioned variables is conducted in the second paragraph of the third part. The cross-correlation of time series is analyzed in the end of the third part. The forth part reports the results and conclusion of the study.

2. Methodology

Time series and regression techniques are used for forecasting of the waste, total population, GDP and household final consumption expenditures. ARIMA models as a subset of linear regression models are more suitable for the short-run forecasting while regression is more suitable for explaining the effect of input variables on an outcome. In order to measure the similarity between the series we analyze cross-correlation functions (CCF).

2.1 Research Plan and Methods Employed

First, we analyze the population numbers, GDP, household final consumption expenditures and waste generation trends with the help of the log-linear regression.

Semi-log models are employed when it is necessary to determine the growth rate or the rate of increment of a particular economic variable. Log-linear analysis is often used for a model building. The fitting criterion of the log-linear analysis is a chi-squared distribution statistic. The assumptions of log-linear analysis are as follows: observations must be randomly selected; observations are independent; normal distribution of observed frequencies; linearity.

Second, we simulate the most adequate autoregressive integrated moving average (ARIMA) models to forecast levels of the total population, GDP, household final consumption expenditures and waste generation for the 2025, 2030 and 2035 years. We also perform a Leung-Box test to make sure that there is no autocorrelation of residues. Then we check whether simulated ARIMA models predict well.

ARIMA models are good for making short-term forecasts in conditions of unpredictability of the behavior of the studied process. It employs historical time series. We apply ARIMA modelling when the data show the evidence of non-stationarity $y_t \sim ARIMA(p, d, q)$, where y_t is a value of a certain economic parameter in the current moment of time t; p is the order of integration (lag order); q is the order of the moving average; d is the degree of differences. The main obstacle with the non-stationary data is variance which is not constant in time, i.e. the further the process goes in time the less an observer is confident that the process is in the vicinity of a zero point. When we see such a notation as $y_t \sim ARIMA(p, 1, q)$, it is equivalent to the notation $\Delta y_t \sim ARMA(p, q)$. It means that in order to transform a non-stationary series into a stationary series, there is the need to take first differences one time, i.e. to shift from y_t to Δy_t. So Δy_t-series are simulated instead of simulating y_t-series. By that the stationary ARMA-process is simulated by the non-stationary ARIMA-process. In the non-stationary process of random walk (drift) a mean is also not constant. Models based on time series analysis, in particular ARIMA, are difficult to configure for a complete analysis, but they almost always give a good result where a high-quality forecast for the medium and short term is required.

The Leung-Box test is a statistical test that checks for the presence of autocorrelation in a time series. The problem of autocorrelation is difficulties in identifying meaningful relationship between two time series. The test is applied to the residuals of the ARIMA model, and not to the original data. The Leung-Box test has the following hypotheses: H_0 – the residuals of the model are distributed independently; H_A – the residuals are not distributed independently; they show a serial correlation. Ideally, the null hypothesis is not rejected, i.e. the p-value of the test must be greater than 0.05, because this means that the residuals for the studied time series model are independent, which is often the assumption made when creating the model. The Leung-Box criterion remains consistent even if the process does not have a normal distribution.

Third, we analyze the cross-correlation between the time series on the waste and the series on the household final consumption expenditures, GDP and population numbers, separately.

The cross-correlation function is employed to analyze the relationship between two time series, i.e. whether the relationship is significant or not. It is important to distinguish between the cross-correlation and autocorrelation. From one hand, they are both measures of similarity between series. From the other hand, the way of application of these two econometric concepts are different. The cross-correlation is applied to analyze the relationship between two series. And the CCF helps to identify the delay in time (shift) of one series relative to the other, i.e. the degree of similarity between them. The autocorrelation studies the relationship within a single time series: the degree of similarity between independent variables and its lagged copies. Its objective is to detect periodicity within a series, to identify trends in data. On balance, autocorrelation is a specific case of the cross-correlation.

Finally, there are conclusions and recommendations. Computations are supplemented by R-scripts in the Appendix.

2.2 Data Set

The data set is an annual time series for Jordan in the period 2010–2022 (Table 1).

Table 1. Data set

year	population	gdp	consumption	waste
2010	6.93	27 133.8	17 892.96	2 579 709.263
2011	7.11	29 524.15	21 784.51	2 726 015.848
2012	7.21	31 634.56	24 440.85	2 881 766.557
2013	7.69	34 454.44	28 939.44	3 038 268.862
2014	8.66	36 847.64	30 083.10	3 184 189.887
2015	9.49	38 587.02	30 719.72	3 308 928.323
2016	9.96	39 892.55	31 622.54	3 411 704.06
2017	10.22	41 608.44	32 867.61	3 494 070.989
2018	10.46	43 370.86	33 223.94	3 558 266.393
2019	10.7	44 503.01	33 240.85	3 448 903.0
2020	10.93	43 579.92	NA	3 561 852.0
2021	11.15	45 116.32	NA	NA
2022	11.29	47 451.5	NA	NA

Note: 'NA' means there is no data.

Data source: constructed by the author based on the World Development Indicators (2023), Ritchie & Mathieu

(2023).

Description of variables:

waste – total household waste generation (kg);

population – total population (million people);

gdp – GDP at purchaser's prices (current US$, million);

consumption – household final consumption expenditure (formerly private consumption) (current US$, million).

3. Results

3.1 Data Trends Patterns

There are graphical visualizations of the trends for the population, GDP, private consumption and waste generation (Figure 1).

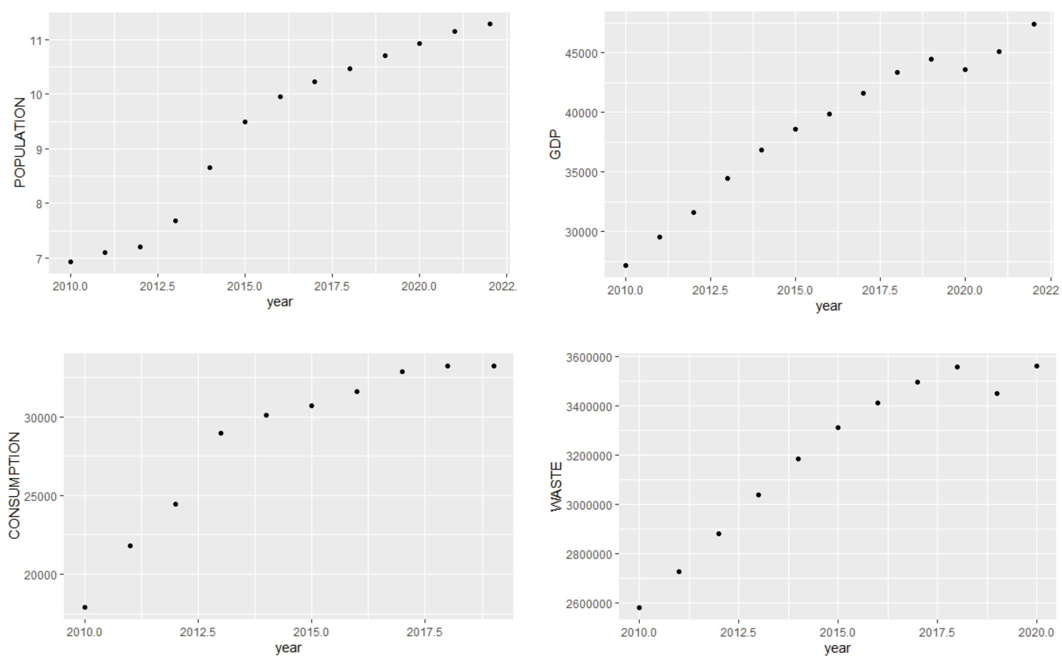

Figure 1. Data trends for population, GDP, consumption and waste

Data source: constructed by the author in R/R-Studio.

3.1.1 Population

First, we analyze the total population trend. From the Figure 1 we can see that there exists an exponential growth pattern. Thus, we'll use a log-linear function to fit an exponential regression model. The regression formula for population is as follows:

$$\ln(population) = -89.55 + 0.046 \cdot year \quad (1)$$

Interpretation: every year the population of Jordan increases on average by 4.6 %.

We may use the regression (1) to make a prediction of population numbers. Making computations in R/R-Studio, numbers of the total population in Jordan for 2025 is 13.058 million, for 2030 – 15.110 million, for 2035 – 17.157 million people.

3.1.2 GDP

Second, we analyze the GDP trend. From the Figure 1 we can see the linear trend with a shift in 2020 (the covid-19 fall in the economy). Thus, we may use a linear function to fit a regression model. However, to make estimates of the model more representative for interpretation, it is better to use a log-linear function to fit a regression. The regression formula for GDP is as follows:

$$\ln(gdp) = -77.2 + 0.044 \cdot year \tag{2}$$

Interpretation: every year the GDP of Jordan increases on average by 4.4 %.

We may use the regression (2) to make a prediction for the GDP. Making computations in R/R-Studio, the GDP in Jordan for 2025 is 53 278.12 million, for 2030 – 61 351.25 million, for 2035 – 69 424.39 million (current US$).

3.1.3 Household Final Consumption Expenditures

Third, we analyze the consumption trend. From the Figure 1 we can see the exponential growth pattern. Thus, we'll use a log-linear function to fit an exponential regression model. The regression formula for consumption is as follows:

$$\ln(consumption) = -115.48 + 0.062 \cdot year \tag{3}$$

Interpretation: every year the level of the household private consumption expenditures (private consumption) in Jordan increases on average by 6.2 %.

We may use the regression (3) to make a prediction for consumption levels. Making computations in R/R-Studio, the private consumption in Jordan for 2025 is 45 601.44 million, for 2030 – 53 753.76 million, for 2035 – 61 906.09 million (current US$).

3.1.4 Waste Generation

Finally, we analyze the waste generation trend. From the Figure 1 we can see a linear trend. In order to move on to the percentage terms, we'll use a log-linear function to fit the model. The regression formula for the waste generation is as follows:

$$\ln(waste) = -49.783 + 0.032 \cdot year \tag{4}$$

Interpretation: every year the waste generation in Jordan increases on average by 3.2 %.

We may use the regression (4) to make a prediction for the waste generation level. Making computations in R/R-Studio, the waste generation in Jordan for 2025 is 4 196 778 kg, for 2030 – 4 695 454 kg, for 2035 – 5 194 131 kg.

In the following paragraph we will compare these estimations with estimations computed with the help of time series modelling.

3.2 Time Series Modelling

Based on the data trends we forecast values for the population number, GDP, private consumption and waste generation for 2025, 2030 and 2035 years. For this purpose, we use time series technique – ARIMA models. A key feature of ARIMA is that these models don't consider exogenous variables, in their basic forms, time is used as a predictor variable. ARIMA models as a subset of linear regression models are more suitable for the short-term forecasting while a regression is more suitable for explaining the effect of input variables on an outcome.

3.2.1 Population

First, we forecast total population numbers. Auto simulation in R suggests ARIMA (2,1,1) with drift as the best model (Figure 2). It means that we have two autoregressive terms, one lag and one moving average term in series.

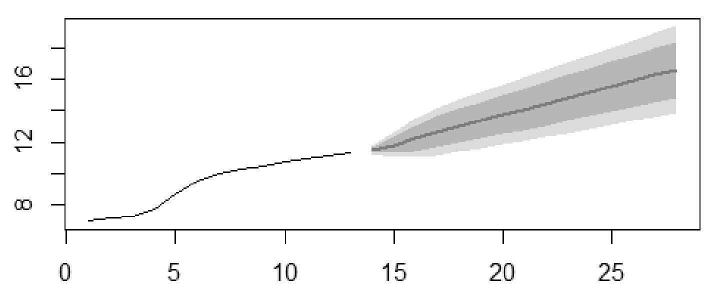

Figure 2. Population forecasting from ARIMA(2,1,1) with drift

Data source: constructed by the author in R/R-Studio.

ARIMA(2,1,1) with drift is a non-stationary process. Analyses of residuals indicates that the forecast is shifted because the distribution of residuals is right-skewed (Figure 3).

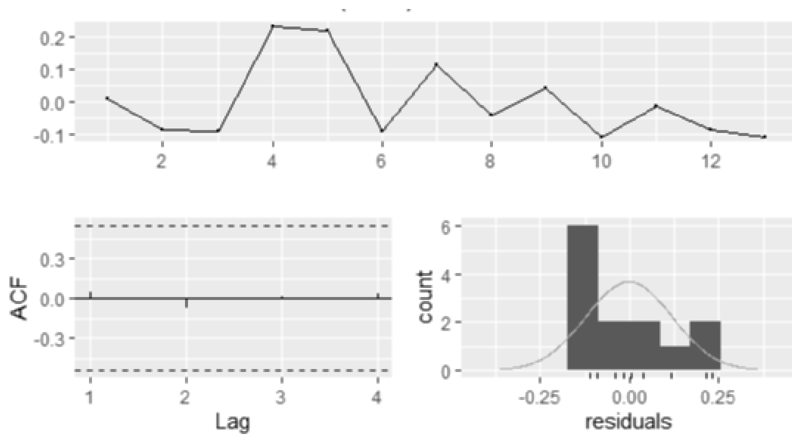

Figure 3. Residuals from ARIMA(2,1,1) with drift

Data source: constructed by the author in R/R-Studio.

Even in the case of a non-normal distribution, the forecasts can be good. However, the prediction intervals that are calculated taking into account the normal distribution may be inaccurate. Thus, the forecast is shifted. Taking first differences gives us a stationary process ARIMA (2,0,0) with a non-zero mean (at a 5 % level of significance) with a more accurate predictive trajectory of the population (Figure 4).

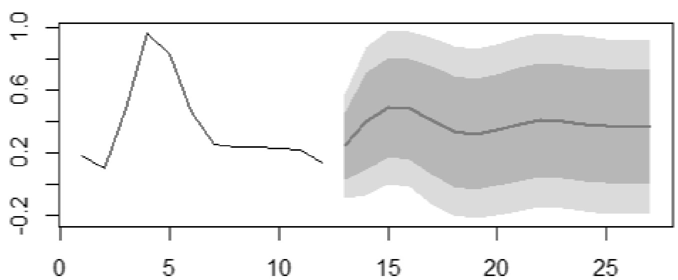

Figure 4. Population forecasting from ARIMA(2,0,0) with non-zero mean

Data source: constructed by the author in R/R-Studio.

The model represented in the Figure 4 shows that population numbers will fluctuate around the non-zero mean.

We perform a Leung-Box test for ARIMA(2,1,1) with drift to make sure that there is no autocorrelation of residues due to the white noise. p-value is 0.6311 so the null-hypothesis is not rejected, and the data are independently distributed, i.e. data don’t exhibit serial correlation. The residuals obtained in the forecast look like the white noise. This allows to use the resulting model ARIMA(2,1,1) with drift for the short-run forecast.

The Figure 2 shows that the level of population in Jordan will increase in the nearest years with the average annual growth rate 4.6 %. Predicted values are in the Table 2.

Table 2. Population forecast

Year	Point forecast	Confidence interval	
		Lower 95 %	Higher 95 %
2025	12.43	11.34	13.51
2030	14.33	12.77	15.89
2035	16.29	14.33	18.26

Data source: computed by the author in R/R-Studio.

Point forecasts of the number of population in 2025 is 12.4 million, for 2030 is 14.3 million, for 2035 is 16.3 million people. Forecast made with the help of ARIMA highly correlates with predicted numbers for population from the regression formula (1) above.

Then we check whether the model ARIMA(2,1,1) with drift predicts well. Making forecast for 2021 gives us 11.31 million versus a real number of 11.15 million people. Forecast for 2022 gives us 11.86 million versus a real number of 11.29 million people. The obtained values indicate that the model ARIMA(2,1,1) with drift predicts well.

We explain the population growth by the natural growth and the refugee problem, alarming for Jordan.

3.2.2 GDP

Second, we forecast the GDP. Auto simulation in R suggests ARIMA (0,1,0) with drift as the best fit model (Figure 5).

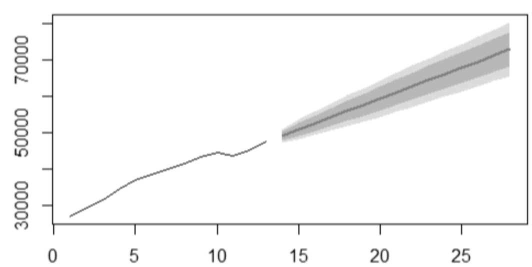

Figure 5. GDP forecasting from ARIMA(0,1,0) with drift

Data source: constructed by the author in R/R-Studio.

ARIMA(0,1,0) with drift is a non-stationary process. This is the series with infinitely slow mean reversion. Analyses of residuals indicates that the forecast is shifted because the distribution of residuals is left-skewed (Figure 6).

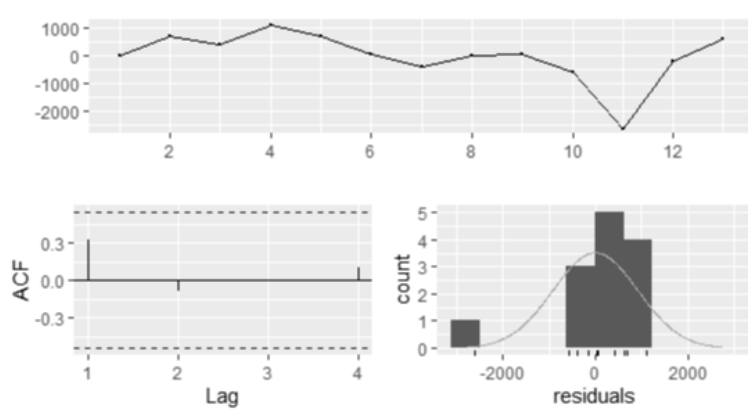

Figure 6. Residuals from ARIMA(0,1,0) with drift

Data source: constructed by the author in R/R-Studio.

We know that even in the case of a non-normal distribution, the forecasts can be good but the prediction intervals may be inaccurate. We perform a Leung-Box test to make sure that there is no autocorrelation of residues. p-value is 0.5554 so the null-hypothesis is not rejected, and the data don't exhibit serial correlation. This allows to use the resulting model for the short-run forecast.

The Figure 5 shows that the GDP in Jordan will increase in the nearest years with the average annual growth rate 4.4 %. Predicted values are in the Table 3.

Table 3. GDP forecast

Year	Point forecast	Confidence interval	
		Lower 95 %	Higher 95 %
2025	52 530.93	49 266.06	55 795.79
2030	60 996.63	55 665.13	66 328.14
2035	69 462.34	62 665.98	76 258.70

Data source: computed by the author in R/R-Studio.

Point forecasts for the GDP in 2025 is 52 531 million, for 2030 is 60 997 million, for 2035 is 69 462 million (current US$). Forecasts made with the help of ARIMA model highly correlate with predicted numbers for the GDP from the regression formula (2) above.

Then we check whether the model ARIMA(0,1,0) with drift predicts well. Making forecast for 2022 gives us 46 027.7 million versus a real number of 47 451.5 million (current US$). The obtained values indicate that the model ARIMA(0,1,0) with drift predicts rather well.

3.2.3 Household Final Consumption Expenditures

Third, we forecast the household final consumption expenditures (private consumption) in Jordan. Auto simulation in R suggests ARIMA (0,2,0) as the best model (Figure 7).

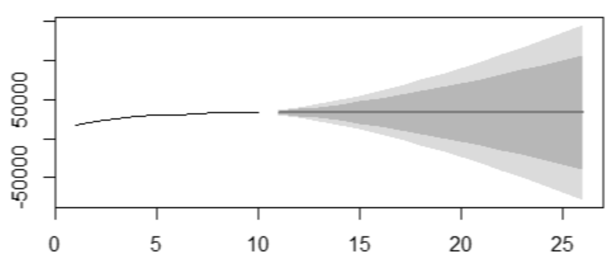

Figure 7. Household final consumption expenditures forecasting from ARIMA(0,2,0)

Data source: constructed by the author in R/R-Studio.

ARIMA(0,2,0) is a non-stationary process. Analyses of residuals indicates that the distribution of residuals is symmetric, close to normal (Figure 8).

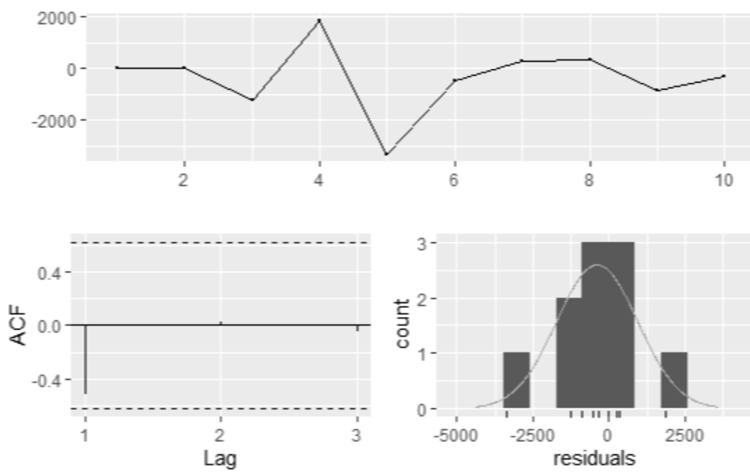

Figure 8. Residuals from ARIMA(0,2,0)

Data source: constructed by the author in R/R-Studio.

The point forecasts and forecasts for prediction intervals are rather accurate because residuals follow nearly normal distribution. We perform a Leung-Box test to make sure that there is no autocorrelation of residues. p-value is 0.304 so the null-hypothesis is not rejected, and the data are independently distributed. This allows to use the resulting model for the forecast.

The Figure 7 shows that the final household consumption expenditure in Jordan will very slowly decline in the nearest years. Predicted values are in the Table 4.

Table 4. Forecast for the household final consumption expenditures

Year	Point forecast	Confidence interval	
		Lower 95 %	Higher 95 %
2025	33 342.31	5 669.03	61 015.59
2030	33 426.86	-31 828.36	98 682.08
2035	33 511.41	-78 691.92	145 714.74

Data source: computed by the author in R/R-Studio.

Point forecasts for the private consumption in 2025 is 33 342 million, for 2030 is 33 427 million, for 2035 is 33 511 million (current US$). Then we check whether the model ARIMA(0,2,0) predicts well. Making forecast for 2019 gives us 33 580.27 million current US$ versus a real number of 33 240.85 million (current US$). The obtained values indicate that the model ARIMA(0,2,0) predicts well.

We explain the decline in the private consumption by the high level of unemployment in the country in cope with the high share of young population and the refugee problem that dramatically worsens this trend, as well as by the personal income decrease.

3.2.4 Waste Generation

Finally, we forecast the waste generation in Jordan. Auto simulation in R suggests ARIMA (0,2,1) as the best fit model (Figure 9).

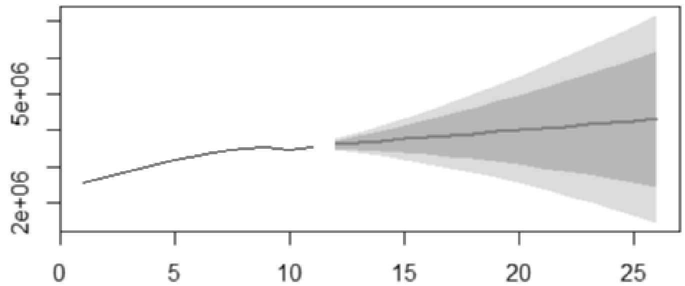

Figure 9. Waste generation forecasting from ARIMA(0,2,1)

Data source: constructed by the author in R/R-Studio.

ARIMA(0,2,1) is a non-stationary process. Analyses of residuals indicates that they follow nearly normal distribution (Figure 10).

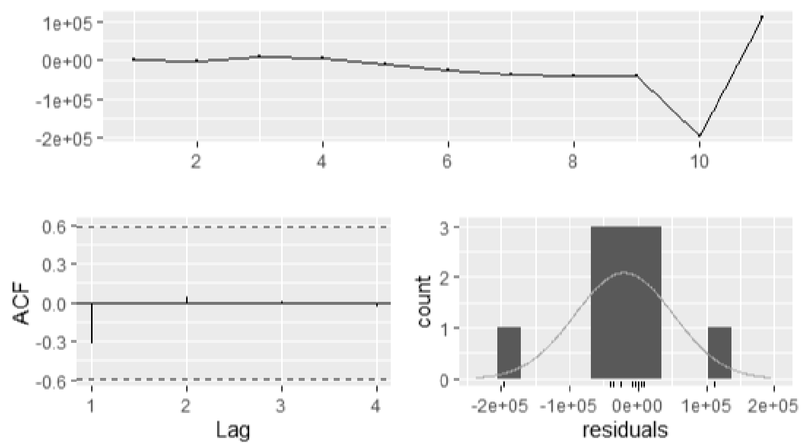

Figure 10. Residuals from ARIMA(0,2,1)

Data source: constructed by the author in R/R-Studio.

The point forecasts and forecasts for prediction intervals are rather accurate because residuals follow nearly normal distribution. We perform a Leung-Box test to make sure that there is no autocorrelation of residues. p-value is 0.6636 so the null-hypothesis is not rejected, and the data don't exhibit serial correlation. This allows to use the resulting model for forecasting.

The Figure 9 shows that the waste generation in Jordan will steadily increase in the nearest years. Predicted values are in the Table 5.

Table 5. Forecast for the waste generation

Year	Point forecast	Confidence interval	
		Lower 95 %	Higher 95 %
2025	3 812 863.0	3 087 809.0	4 537 918.0
2030	4 063 875.0	2 392 102.0	5 735 647.0
2035	4 314 886.0	1 470 807.0	7 158 965.0

Data source: computed by the author in R/R-Studio.

Point forecasts for the waste generation in 2025 is 3.81 million, for 2030 is 4.06 million, for 2035 is 4.31 million (kg). Forecast made with the help of ARIMA correlates with predicted numbers for the waste generation from the regression formula (4) above.

Then we check whether the model ARIMA(0,2,1) predicts well. Making forecast for 2020 gives us 3.34 million versus a real number of 3.56 million (kg). The obtained values indicate that the model ARIMA(0,2,1) predicts rather well.

3.3 Cross-Correlation of Time Series

In this part we model in pairs the cross-correlation between the time series on waste and the series on the population, the GDP and the private consumption. The cross-correlation is a measure of similarity between two series. Respectively, we use the CCF to analyze whether there is a relationship between them, i.e. whether and how one series is shifted in time relative to the other.

Usually (the rule of thumb), a cross-correlation is significant (at the level of significance 5 %) when the absolute value is greater than the following expression:

$$\frac{2}{\sqrt{n-|k|}} \tag{5}$$

where k is a lag, n is a number of observations (Interpret…, 2023). In our case, the number of observations is 10 (2010 – 2019 years).

3.3.1 Cross-Correlation Between Series on 'Waste' and "Consumption'

First, we measure the cross-correlation between the waste generation and the household private consumption series (Figure 11).

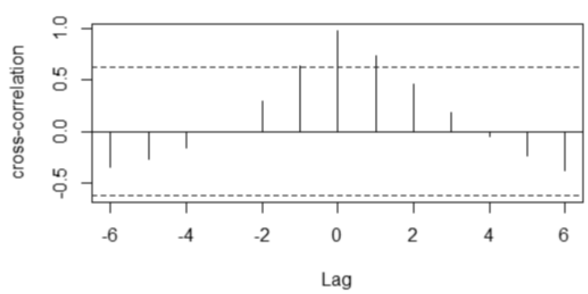

Figure 11. CCF between series on 'waste' and 'consumption'

Data source: constructed by the author in R/R-Studio.

The Figure 11 shows that the value of CCF is 0.975 at the lag 0. Computations of the expression (5) give us the value 0.633. Using the rule above, since 0.975 is greater than 0.633, there is the evidence of significant relationship between these two series. The cross-correlation value is close to 1, so series tend to move in one direction. There is the absence of shift in time of one series relative to the other (CCF is the highest at the lag 0).

3.3.2 Cross-Correlation Between Series on 'Waste' and "GDP'

Second, we measure the cross-correlation between the waste generation and the GDP series (Figure 12).

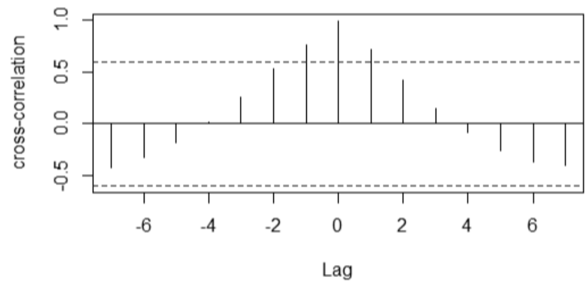

Figure 12. CCF between series on ‘waste’ and ‘gdp’

Data source: constructed by the author in R/R-Studio.

The Figure 12 shows that the value of CCF is 0.983 at the lag 0. The value is greater than the value from the computations of the expression (5), which is 0.633. So there is the evidence of significant relationship between these two series. There is the absence of shift in time of one series relative to the other (CCF is the highest at the lag 0). The cross-correlation value is close to 1, so series tend to move in one direction.

3.3.3 Cross-Correlation Between Series on ‘Waste’ and “Population’

Finally, the cross-correlation between the waste and population is measured (Figure 13).

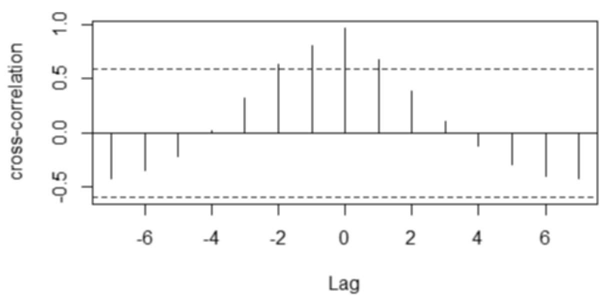

Figure 13. CCF between series on ‘waste’ and ‘population’

Data source: constructed by the author in R/R-Studio.

The value of CCF is 0.965 at the lag 0. Using the rule above (the expression (5)), since 0.965 is greater than 0.633, there is the evidence of the significant relationship between these two series. The cross-correlation value is close to 1, so series are identical. CCF is the highest at the lag 0, there is the absence of shift in time of one series relative to the other.

The analysis of cross-correlation function has revealed that pairs of series (waste/consumption, waste/gdp, waste/population) are identical, moving in one direction with no shift in time of one series to the other. The relationship in every pair of series is significant.

4. Results and Conclusion

It is difficult to decouple the waste generation from the economic growth of a country. Although, the household waste is the biggest contributor to the total municipal solid waste generation. In the research we discovered the Jordanian economy in four dimensions: the waste generation, the number of population, the gross domestic product (GDP) at purchasers’ prices, and household final consumption expenditures. The following tasks have been fulfilled: (1) adequate time series models for the short-run forecasting of the number of population, the gross domestic product (GDP) at purchasers’ prices, and household final consumption expenditures have been simulated; (2) the cross-correlation between the waste generation series and three other series have been analyzed.

We analyzed whether time series models for population, GDP, private consumption and waste generation predict adequately. We have observed that the oscillation directions coincide with the initial data. The actual and forecast values at the end of the period are close in magnitude. The trend is reflected adequately. Thus, ARIMA(2,1,1) with drift for the population, ARIMA(0,1,0) with drift for the GDP, ARIMA(0,2,0) for the household final consumption expenditures, and ARIMA(0,2,1) for the waste generation may be used for the short-run forecasts.

The analysis of cross-correlation functions has revealed that pairs of series (waste/gdp, waste/consumption, waste/population) are identical, moving in one direction with no shift in time of one series to the other. The relationship in every pair of series is significant. We expect that in the short-run, *ceteris paribus*, with the growth of population, GDP and/or private consumption, the level of household waste generation in the country will also grow in a close relationship.

The most sensitive limitation of the econometric analysis is the quality of data. The accuracy of its collection largely depends on the efficiency of statistical bodies. In general, the data is less trustworthy in countries with the lack of financial resources and/or its collection in rural and remote areas where it is highly likely remains unrecorded. Nevertheless, the latest official statistical data has been used in the research.

ARIMA models as a subset of linear regression models are more suitable for the short-run forecasting while regression is more suitable for explaining the effect of input variables on an outcome. For the future research, it is supposed to make forecasting of the household waste generation with the help of the Regression with ARIMA Errors combining two powerful instruments – ARIMA and linear regression. To apply this technique, high-quality, disaggregated and regularly collected data is required.

The research outcomes are useful for policy makers to realize the scale of the household waste problem and to optimize capital expenditures into the waste management system of Jordan. The research contributes to the solution of the waste management problem in terms of forecasting, necessary to put forward appropriate plans.

References

Abdoli, M., Falahnehzad, M., & Behboudian, S. (2011). Multivariate econometric approach for solid waste generation modelling: a case study of Mashhad, Iran. *Environmental Engineering Science*, *28*(9). https://doi.org/10.1089/ees.2010.0234

Abir, T., Datta, M., & Saha, S. (2023). Assessing the factors influencing effective municipal solid waste management system in Barishal metropolitan areas. *Journal of Geoscience and Environmental Protection*, *11*(1), 49-66. https://doi.org/10.4236/gep.2023.111004

Adeleke, O., Akinlabi, S., Jen, T., & Dunmade, I. (2021). An overview of factors affecting the rate of generation and physical composition of municipal solid waste. *IOP Conference Series Materials Science and Engineering*, *1107*(1). https://doi.org/10.1088/1757-899X/1107/1/012096

Afroz, R., Hanaki, K., & Tudin, R. (2010). Factors affecting waste generation: a study in a waste management program in Dhaka city, Bangladesh. *Environmental Monitoring and Assessment*, *179*(1-4), 509-519. https://doi.org/10.1007/s10661-010-1753-4

Aldayyat, E., Saidan, M., Abu Saleh, M., & Hamdan, S. (2019). Solid waste management in Jordan: impacts and analysis. *Journal of Chemical Technology and Metallurgy*, *54*(2), 454-462. Retrieved from https://www.researchgate.net/publication/330925093_Solid_Waste_Management_in_Jordan_Impacts_and_Analysis

Alhanaqtah, V. (2020). Econometric modelling of solid waste generation in Jordan. *Journal of Economics and Sustainable Development*, *11*(16), 204-220. https://doi.org/10.7176/JESD/11-16-22

Atieno, J., Oindo, B., & Bosire, E. (2017). Modeling household solid waste generation in urban estates using socioeconomic and demographic data, Kisumu City, Kenya. *Scholars Journal of Arts, Humanities and Social Sciences*, *5*(5), 468-477. Retrieved from https://repository.maseno.ac.ke/handle/123456789/441

Cheng, J., Shi, F., Yi, J., & Fu, H. (2020). Analysis of the factors that affect the production of municipal solid waste in China. *Journal of Cleaner Production*, *259*(4), 120808. https://doi.org/10.1016/j.jclepro.2020.120808

Chhay, L., Reyad, M., Suy, R., Islam, M., & Mian, M. (2018). Municipal solid waste generation in China: influencing factor analysis and multi-model forecasting. *Journal of Material Cycles and Waste Management*, *20*(3), 1761-1770. https://doi.org/10.1007/s10163-018-0743-4

Chung, S. (2010). Projecting municipal solid waste: the case of domestic waste in Hong Kong special administration region. *Resources Conservation and Recycling*, *54*(11), 759-768. https://doi.org/10.1016/j.resconrec.2009.11.012

Hage, O., Sandberg, K., Söderholm, P., & Berglund, C. (2018). The regional heterogeneity of household recycling: a spatial-econometric analysis of Swedish plastic packing waste. *Letters in Spatial and Resource Sciences*, *11*(2). https://doi.org/10.1007/s12076-017-0200-3

Han, Z., Liu, Y., Zhong, M., Shi, G., Li, Q., Zeng, D., Zhang, Y., Fei, Y., & Xie, Y. (2018). Influencing factors of domestic waste characteristics in rural areas of developing countries. *Waste Management*, *72*, 45-54. https://doi.org/10.1016/j.wasman.2017.11.039

Herianto, H., Maryono, M., & Budihardjo, M. (2019). Factors affecting waste generation: household study in Palangka Raya City, Central Kalimantan. *The 4th International Conference on Energy, Environment, Epidemiology and Information System (ICENIS 2019)*, *125*, 07007. https://doi.org/10.1051/e3sconf/201912507007

Hockett, D., Lober, D., & Pilgrim, K. (1995). Determinants of per capita municipal solid waste generation in the Southeastern United States. *Journal of Environmental Management*, *45*(3), 205-217. https://doi.org/10.1006/jema.1995.0069

Khan, D., Kumar, A., & Samadder, S. (2016). Impact of socioeconomic status on municipal solid waste generation rate. *Waste Management*, *49*, 15-25. https://doi.org/10.1016/j.wasman.2016.01.019

Liu, C., & Wu, X. (2011). Factors influencing municipal solid waste generation in China: A multiple statistical analysis study. *Waste Management and research: The Journal of a Sustainable Circular Economy*, *29*(4), 371-378. https://doi.org/10.1089/cpb.2008.0321

Liu, J., Li, Q., Gu, W., & Wang, C. (2019). The impact of consumption patterns on the generation of municipal solid waste in China: evidences from provincial data. *International Journal of Environmental Research and Public Health*, *16*(10), 1717. https://doi.org/10.3390/ijerph16101717

Lu, J., Zhang, S., Hai, J., & Lei, M. (2017). Status and perspectives of municipal solid waste incineration in China: a comparison with developed regions. *Waste Management*, *69*, 170-186. https://doi.org/10.1016/j.wasman.2017.04.014

Prades, M., Gallardo, A., & Ibàñez, M. (2015). Factors determining waste generation in Spanish towns and cities. *Environmental Monitoring and Assessment*, *187*, 4098. https://doi.org/10.1007/s10661-014-4098-6

Ritchie, H., & Mathieu, E. (2023). Waste management. *Our World in Data*. Retrieved from https://ourworldindata.org/waste-management

Samson, M., Akinlabi, E., Muzenda, E., & Aboyade, A. (2017). A review on factors affecting municipal solid waste generation. *2nd International Engineering Conference*. Federal University of Technology, Minna, Niger State (October 17-19, 2017). Retrieved from https://www.academia.edu/35943085/A_Review_on_Factors_affecting_Municipal_Solid_Waste_Generation

Trang, P., Dong, H., Toan, D., Hanh, N., & Thu, N. (2017). The effects of socio-economic factors on household solid waste generation and composition: a case study in Thu Dau Mot, Vietnam. *Energy Procedia*, *107*, 253-258. https://doi.org/10.1016/j.egypro.2016.12.144

World Development Indicators. (2023). *The world Bank*. Retrieved from https://databank.worldbank.org/source/world-development-indicators

Zhao, Y., Diunugala, H., & Mombeuli, C. (2021). Factors affecting household solid waste generation and management in Sri Lanka: an empirical study. *Environmental Monitoring and Assessment*, *193*, 838. https://doi.org/10.1007/s10661-021-09633-7

Appendix

R-scripts for calculations and figures

```
library("ggplot2")
library("forecast")
library("tseries")
#variable 1 "population"
year<-c(2010,2011,2012,2013,2014,2015,2016,2017,2018,2019,2020,2021,2022)
population<-c(6.93,7.11,7.21,7.69,8.66,9.49,9.96,10.22,10.46,10.7,10.93,11.15,11.29)
d<-data.frame(year,population)
qplot(year,population,data=d,ylab = "POPULATION")
model <- lm(log(population) ~ year, data = d)
summary(model)
# prediction
model_0<-lm(population~year,data=d)
nd_pop<-data.frame(year=c(2025,2030,2035))
nd_pop
predict(model_0,nd_pop)
# Time series
Y<-d$population
tsdisplay(Y)
adf.test(Y)
mod_a <- auto.arima(Y,trace=TRUE,ic="aic")
summary(mod_a)
checkresiduals(mod_a)
dy<-diff(Y)
mod_a1<-auto.arima(dy)
mod_a1 # ARIMA(2,0,0) with non-zero mean
adf.test(dy)
# Prediction
prediction_a <- forecast(mod_a, h=15)
prediction_a
plot(prediction_a)
prediction_b <- forecast(mod_a1, h=15)
prediction_b
plot(prediction_b)
#variable 2 "gdp"
year<-c(2010,2011,2012,2013,2014,2015,2016,2017,2018,2019,2020,2021,2022)
gdp<-c(27133.8,29524.15,31634.56,34454.44,36847.64,38587.02,39892.55,41608.44,43370.86,44503.01,43579.92,
45116.32,47451.5)
g<-data.frame(year,gdp)
qplot(data=g,year,gdp,ylab="GDP")
model <- lm(log(gdp) ~ year, data = g)
```

```
summary(model)
model_0<-lm(gdp~year,data=g)
nd_gdp<-data.frame(year=c(2025,2030,2035))
nd_gdp
predict(model_0,nd_gdp)
# Time series
Y<-g$gdp
tsdisplay(Y)
adf.test(Y)
mod_a <- auto.arima(Y,trace=TRUE,ic="aic")
summary(mod_a)
checkresiduals(mod_a)
# Prediction
prediction_a <- forecast(mod_a, h=15)
prediction_a
plot(prediction_a)
#variable 3 "consumption"
year<-c(2010,2011,2012,2013,2014,2015,2016,2017,2018,2019)
consumption<-c(17892.96,21784.51,24440.85,28939.44,30083.1,30719.72,31622.54,32867.61,33223.94,33240.85)
h<-data.frame(year,consumption)
qplot(data=h,year,consumption,ylab="CONSUMPTION")
model <- lm(log(consumption) ~ year, data = h)
summary(model)
model_0<-lm(consumption~year,data=h)
nd_cons<-data.frame(year=c(2025,2030,2035))
nd_cons
predict(model_0,nd_cons)
# Time series
Y<-h$consumption
tsdisplay(Y)
adf.test(Y)
mod_a <- auto.arima(Y,trace=TRUE,ic="aic")
summary(mod_a)
checkresiduals(mod_a)
# Prediction
prediction_a <- forecast(mod_a, h=16)
prediction_a
plot(prediction_a)
#variable 4 "waste" (linear function)
year<-c(2010,2011,2012,2013,2014,2015,2016,2017,2018,2019,2020)
waste<-c(2579709.263,2726015.848,2881766.557,3038268.862,3184189.887,3308928.323,3411704.06,3494070.98
9,3558266.393,3448903.0,3561852.0)
```

```
w<-data.frame(year,waste)
qplot(data=w,year,waste,ylab="WASTE")
model <- lm(log(waste) ~ year, data = w)
summary(model)
model_0<-lm(waste~year,data=w)
nd_waste<-data.frame(year=c(2025,2030,2035))
nd_waste
predict(model_0,nd_waste)
# Time series
Y<-w$waste
tsdisplay(Y)
adf.test(Y)
mod_a <- auto.arima(Y,trace=TRUE,ic="aic")
summary(mod_a)
checkresiduals(mod_a)
mod_a <- auto.arima(Y,trace=TRUE,ic="aic")
summary(mod_a)
# Prediction
prediction_a <- forecast(mod_a, h=15)
prediction_a
plot(prediction_a)
# Back-forecast with the help of ARIMA
#variable 1 "population"
year<-c(2010,2011,2012,2013,2014,2015,2016,2017,2018,2019,2020)
population<-c(6.93,7.11,7.21,7.69,8.66,9.49,9.96,10.22,10.46,10.7,10.93)
d<-data.frame(year,population)
Y<-d$population
#variable 2 "gdp"
year<-c(2010,2011,2012,2013,2014,2015,2016,2017,2018,2019,2020,2021)
gdp<-c(27133.8,29524.15,31634.56,34454.44,36847.64,38587.02,39892.55,41608.44,43370.86,44503.01,43579.92,
45116.32)
g<-data.frame(year,gdp)
Y<-g$gdp
#variable 3 "consumption"
year<-c(2010,2011,2012,2013,2014,2015,2016,2017,2018)
consumption<-c(17892.96,21784.51,24440.85,28939.44,30083.1,30719.72,31622.54,32867.61,33223.94)
h<-data.frame(year,consumption)
Y<-h$consumption
#variable 4 "waste"
year<-c(2010,2011,2012,2013,2014,2015,2016,2017,2018,2019)
waste<-c(2579709.263,2726015.848,2881766.557,3038268.862,3184189.887,3308928.323,3411704.06,3494070.98
9,3558266.393,3448903.0)
```

```
w<-data.frame(year,waste)
Y<-w$waste
mod_a <- auto.arima(Y,trace=TRUE,ic="aic") # for all variables separately
summary(mod_a)
prediction_a <- forecast(mod_a, h=15)
prediction_a
plot(prediction_a)
# Cross-correlation
w<-w$waste
c<-h$consumption
g<-g$gdp
p<-d$population
cor <- ccf(w,c, ylab = "cross-correlation")
cor <- ccf(w,g, ylab = "cross-correlation")
cor <- ccf(w,p, ylab = "cross-correlation")
```