



Universiteit
Leiden
The Netherlands

Systems analysis of stock buffering: development of a dynamic substance flow-stock model for the identification and estimation of future resource, waste streams and emissions

Elshkaki, A.

Citation

Elshkaki, A. (2007, September 6). *Systems analysis of stock buffering: development of a dynamic substance flow-stock model for the identification and estimation of future resource, waste streams and emissions*. Retrieved from <https://hdl.handle.net/1887/12301>

Version: Not Applicable (or Unknown)

License: [Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)

Downloaded from: <https://hdl.handle.net/1887/12301>

Note: To cite this publication please use the final published version (if applicable).

Appendix

Appendix A - Additions to the regression analysis

1 Introduction

This appendix has a threefold aim. Firstly it gives an overview of more detailed statistical methods applied in regression analysis. Secondly it applies these methods to the results obtained in chapter 4. And thirdly, it performs a renewed more detailed analysis on the data of chapter 4.

Regression analysis is used in several scientific fields as a statistical tool to estimate and analyze the relation between a dependent variable and a number of independent, explanatory variables. Thus, regression analysis can for instance be used to describe the demand for electricity, metals or other commodities as a function of socio-economic variables such as GDP, population, price and further specific variables for any of these commodities (Liu et al. 1991; Labson and Crompton, 1993; Burney, 1995; Roberts, 1996; Moore et al. 1996; Ranjan and Jain, 1999; Guzman, et al. 2004; Mohamed and Bodger, 2005).

Regression analysis is used in this thesis for the analysis of the relative importance of different explanatory variables on the shape of the inflow of different metals; and subsequently on the stock-in-use over time. Regression analysis identifies the variables that are significant and contribute most to the dependent variable. It further examines the effect of separate significant variables, and also their combined effect. The fitting algorithm that determines the regression model parameters uses the ordinary least square (OLS) criterion (Gijbels and Rousson, 2001). The optimal regression model, the adequacy of the model and the significance of the variables are traditionally described in the following statistical parameters: the coefficient of determination (R^2), the adjusted coefficient of determination (R^2_{adj}), and the t- and F-statistics.

Although these statistical terms are normally used in the application of the regression model and in the determination of the significance of their individual variables, there are other statistical tests, which can be used to further examine the adequacy of the outcome of the regression analysis. This particularly pertains to the optimal number of explanatory variables, the degree of autocorrelation of the residues, the degree of multicollinearity and the assumption of homoscedasticity, that is, that the residuals do have a constant variance across observations. Below, these additional tests will be applied to the results of chapter 4.

In addition a new test round will be performed, using an adapted set of explanatory variables. The model presented in chapter 4 analyses the relation between the inflow of Cathode Ray Tubes (CRTs) into the stock-in-use as dependent variable and a number of socio-economic variables as explanatory variables. The socio-economic variables used are Gross Domestic Product (GDP), population size and time. These variables had been chosen on the basis of the above mentioned traditional statistical parameters: the coefficient of determination (R^2), the adjusted coefficient of determination (R^2_{adj}), and the t- and F-statistics.

Although in the analysis of Chapter 4, GDP and population size appeared to be the most influential variables and although this is also found in studies on other commodities (Liu et al. 1991 and Mohamed and Bodger, 2005), the inclusion of these variables in one model is questionable due to the high probability of multicollinearity between these two variables. This is the case because GDP is determined by the product of population size and welfare per capita. Therefore below a new analysis will be performed using population size and GDP per capita as explanatory variables.

In section 2 of this appendix a description will be given of the general aim and structure of regression analysis. In section 3 an overview will be given of the possibilities for more in depth testing the adequacy of the results of regression analysis. In section 4 these methods are applied to the results of chapter 4. In section 5, a renewed analysis will be performed, using other sets of explanatory variables. And in section 6 conclusions are drawn.

2 General aim and structure of regression analysis

2.1 Introduction

Regression analysis examines the strength of a relation between a dependent variable and a number of independent variables, also called explanatory variables. The mathematical model of the relation between the dependent variable and the explanatory variables is known as the regression model. The regression model contains one or more unknown parameters that are estimated using the given data on the explanatory variables. Eq. 1 describes the linear regression model that is used in the analysis:

$$Y(t) = \beta_0 + \sum_{i=1}^n \beta_i X_i(t) + \varepsilon(t) \quad (1)$$

where $Y(t)$ is the inflow of product into the product stock at time t , n is the number of explanatory variables, $X_i(t)$ is the explanatory variables at time t , β_i is the regression model parameter and $\varepsilon(t)$ is the residuals of the regression model.

The fitting algorithm that determines the regression model parameters (β 's) in Eq. (1) uses the ordinary least square (OLS) criterion (Gijbels and Rousson, 2001). OLS is a mathematical optimization technique that attempts to find the best function that minimizes the sum of the squares of the residuals.

2.2 The coefficient of determination (R^2)

The coefficient of determination R^2 is a non-dimensional measure of how well a regression model describes a set of data. R^2 is a measure of how much of the variance in the dependent variable is explained by the explanatory variables in the regression model. Large values of R^2 indicate better agreement between the model and the data. Eq. 2 gives the coefficient of determination (R^2):

$$R^2 = \frac{SS_{reg}}{SS_{tot}} = 1 - \frac{SS_{res}}{SS_{tot}} \quad (2)$$

where SS_{reg} is the regression sum of squares, SS_{tot} is the total sum of squares and SS_{res} is the residuals sum of squares.

R^2 does not restrict the number of explanatory variables. Adding another explanatory variable (thus removing a degree of freedom) will always lead to an increase in the regression sum of squares (SS_{reg}) (also called explained sum of squares) and a decrease in the residuals sum of squares (SS_{res}), consequently an increase in R^2 . Therefore R^2 by itself is not a good indication for the most optimal regression model.

This problem can be solved by replacing the coefficient of determination (R^2) by the adjusted coefficient of determination (R^2_{adj}).

2.3 The adjusted coefficient of determination (R^2_{adj})

The adjusted coefficient of determination (R^2_{adj}) also describes the total explained variance by the explanatory variables but it penalizes adding too many additional variables (Glanz and Slinker, 1990). Therefore R^2_{adj} is used to find the optimal regression. Eq. 3 gives the adjusted coefficient of determination (R^2_{adj}):

$$R^2_{adj} = 1 - \frac{MS_{res}}{MS_{tot}} = 1 - \frac{SS_{res} / (n - k - 1)}{SS_{tot} / (n - 1)} \quad (3)$$

where MS_{res} is the residuals mean squares, MS_{tot} is the total mean squares and k is the number of explanatory variables in the regression equation.

If another explanatory variable will be added, the number of explanatory variables (k) will increase and the residuals sum of squares (SS_{res}) will decrease. For R^2_{adj} to increase, the decrease in SS_{res} must be more than the decrease of $(n-k-1)$. The preferred model is the one with the highest R^2_{adj} .

2.4 t-statistics and F-statistics

The t-statistics indicates whether or not each regression parameters is significantly different from zero. The F-statistics indicates whether all explanatory variables taken together do significantly explain the dependent variable. Eq. 4 and Eq. 5 give the t-statistics and F-statistics. The critical values for t-statistics and F-statistics at different probability levels can be found in statistical books.

$$t = \frac{b_i}{Sb_i} \quad (4)$$

$$F = \frac{MS_{reg}}{MS_{res}} \quad (5)$$

where Sb_i is the standard error of b_i and MS_{reg} is the regression mean square

3 Possibilities for further analysis

3.1 Introduction

In addition to the above-mentioned statistical measures, there are other statistical tests that can be used for further examination of the adequacy of the regression model.

The first option for further improvement concerns the determination of the optimal regression model that is the optimal number of explanatory variables. For this reason Akaike's Information Criterion (AIC) and Schwartz Criterion (SIC) can be used. AIC is developed by Hirotugu Akaike in 1971. SIC, also called Bayesian information criterion (BIC) is developed by Schwartz in 1978. These two statistics have the same function as R^2_{adj} (Egelioglu et al. 2001 and Todeschini et al., 2004), but the SIC penalizes stronger for adding additional explanatory variables.

The second option concerns the determination of autocorrelation of the residuals from regression analysis. Autocorrelation of the residuals means that the model still can be improved and it leads to bias in the estimates of statistical significance of the parameters estimates. For this aim the Durbin – Watson Statistics, has been developed. Durbin – Watson Statistics is named after James Durbin and Geoffrey Watson.

Further points for improvement of regression analysis deal with the problem of multicollinearity and the problem of heteroscedasticity.

Multicollinearity refers to a high linear relationship between the explanatory variables used in the regression model. The best regression model is the model in which each of its explanatory variables correlates highly with the dependent variable but minimally with the other explanatory variables. Multicollinearity refers to a correlation between explanatory variables (R^2) above 0.80. The problem associated with multicollinearity is the resulting overfitting in the regression analysis model. Multicollinearity however, does not affect the usefulness of a regression equation for purely empirical description of data or prediction of new observations if no interpretation is made based on the individual coefficient because multicollinearity problem does not result in biased coefficient estimates (Glanz and Slinker, 1990) but it increases the standard error of the estimates. Multicollinearity is a computational problem, that is, with existing perfect multicollinearity, the least square method can not be carried out (Makridakis et al., 1998). Therefore, most computer statistical software programs warn against serious multicollinearity (Glanz and Slinker, 1990).

Heteroscedasticity is a violation of the assumption that the residuals of the regression has a constant variance across observations (homoscedasticity). Also the violation of the assumption of homoscedasticity does not invalidate the regression model because the estimators remain unbiased and strongly consistent, but heteroscedasticity leads to an underestimation of the standard error thus deriving too narrow confidence intervals or small P-value (White, 1980 and Long and Ervin, 2000).

There are several methods that can be used to reduce the effect of heteroscedasticity (Long and Ervin, 2000 and Mackinnon and White, 1985) and with the advent of robust standard errors, testing conditional homoscedasticity is not as important as it used to be.

3.2 The optimal regression model

Akaike's Information Criterion (AIC)

Akaike's Information Criterion (AIC) as given by Eq. 6, can be used to find the optimal regression model.

$$AIC = 2k - 2\ln(L) \quad (6)$$

where k is the number of explanatory variables in the regression equation and L is the likelihood function.

Assuming that the model errors are normally distributed, Eq. 6 can be written as given by Eq. 7:

$$AIC = 2k - 2\ln\left(\frac{SS_{res}}{n}\right) \quad (7)$$

where n is the number of observations.

In the present equation the regression model will only be improved as long as AIC is decreasing. Therefore, the preferred model is the one with the lowest AIC value.

Schwartz Criterion (SIC)

Schwartz Criterion (SIC) as given by Eq. 8, can also be used to find the optimal regression model.

$$SIC = -2\ln(L) + k \ln(n) \quad (8)$$

Assuming that the model errors are normally distributed, Eq. 8 can be written as given by Eq. 9:

$$SIC = -2\ln\left(\frac{SS_{res}}{n}\right) + k \ln(n) \quad (9)$$

Comparable to above, SIC is decreasing as long as the regression model is improving by adding additional explanatory variables. The preferred model is the one with the lowest value of SIC.

3.3 Autocorrelation of the residuals

Durbin – Watson Statistics

Eq. 10 gives the Durbin – Watson Statistics. The value of Durbin-Watson statistics always lies between 0 and 4. Ideally, when there is no autocorrelation, the value of Durbin-Watson statistics should be close to 2. The critical values of Durbin – Watson Statistics, which are determined by the number of explanatory variables and the number of observations, can be found in statistical books.

$$d = \frac{\sum_{t=2}^T (e_t - e_{t-1})^2}{\sum_{t=1}^T (e_t)^2} \quad (10)$$

where e_t is the residuals associated with the observations at time t .

3.4 Multicollinearity

To test for multicollinearity, the correlation between the explanatory variables has to be checked. If the correlation between two explanatory variables is high, that is R^2 higher than 0.80, the analysis is confounded by multicollinearity (Glanz and Slinker, 1990). This can be solved by a more careful selection or by redefinition of the explanatory variables, thus avoiding this problem.

3.5 Heteroscedasticity

Several tests can be used to examine heteroscedasticity such as Eyeball test, Breusch-Pagan test and others.

The **Eyeball test** is a graphic test used to test for heteroscedasticity by plotting the residuals of a regression model against one or more of the explanatory variables X 's or the predicted dependent variable Y . The assumption of homoscedasticity is not rejected if the dispersion of the residuals appears to be the same across all value of X or Y .

The **Breusch-Pagan test** is used to test for heteroscedasticity by testing whether the estimated residuals from a regression analysis is dependent on the values of the explanatory variables. The squared residuals is related to the explanatory variables. If the F-test confirms that the explanatory variables are jointly significant then the assumption of no heteroscedasticity is rejected.

4 Results of further elaboration of regression analysis of chapter 4

In this section, the results of the further analysis, using the above tests, of the CRT's inflow into the stock-in-use will be presented.

4.1 The optimal regression model

The explanatory variables used in the analysis are Gross Domestic Product (GDP), Population (Pop), and a Time variable (T). The time variable is used as a proxy for the combined influence of other variables on the inflow trend. The fitting algorithm that determines the regression model parameters ($\beta_0, \beta_1, \beta_2, \beta_3$) in Eq. (11) uses the ordinary least square (OLS) criterion (Gijbels and Rousson, 2001).

$$Y(t) = \beta_0 + \beta_1 \cdot GDP(t) + \beta_2 \cdot Pop(t) + \beta_3 \cdot T(t) + \varepsilon(t) \quad (11)$$

where $Y(t)$ is the inflow of goods at time t , β_0 is the overall mean response or regression intercept, $\beta_1, \beta_2, \beta_3$ are the regression model parameters, GDP, Pop, and T are the explanatory variables and $\varepsilon(t)$ is the residuals of the regression model.

Table 1 shows the result of the regression analysis, together with the results of the AIC and SIC statistics.

Table 1: Statistical analysis related to the regression analysis, using GDP, population and T as explanatory variables

Estimation	Socio-economic variables	R^2	R^2_{adj}	AIC	SIC	D/W	t-value	t-value	t-value	F-value
1	GDP	0.48	0.43	91.05	91.62	0.74	3.18			10.15
2	POP	0.68	0.65	84.70	85.26	1.06		4.84		23.49
3	Time	0.77	0.75	80.61	81.17	1.35			9.97	36.23
4	GDP, POP	0.70	0.64	86.02	87.15	1.24	-0.73	2.67		11.51
5	POP, Time	0.82	0.79	78.94	80.07	1.71		-1.80	2.85	23.46
6	GDP, Time	0.78	0.73	82.13	83.26	1.48	-0.61		3.63	17.27
7	GDP, POP, Time	0.85	0.80	78.64	80.34	2.07	1.32	-2.17	3.08	17.38

R^2 is used in the analysis as indication of the model that gives maximum explained variance by the explanatory variables independent of the number of these variables. It is clear from Table 1 that the maximum R^2 ($R^2 = 0.85$) was obtained when the three explanatory variables are included in the regression model (estimation 7). R^2_{adj} is similar to R^2 but it reaches the maximum with the optimal number of explanatory variable. The maximum R^2_{adj} ($R^2_{adj} = 0.80$) was also obtained when the three variables are included in the regression model. The t-statistics for the individual coefficients shows that in this model the GDP and population are not significant. However, the F-statistics indicates that all the three explanatory variables taken together are significant and contribute to the inflow. Therefore the optimal model is the model that includes the three explanatory variables GDP, population and time.

4.2 Further tests for the optimal regression model

The optimal model in chapter 4 is determined using R^2_{adj} . Other tests comparable to R^2_{adj} can also be used to find the optimal model such as the AIC and SIC tests. The SIC test however, penalizes adding additional variable more than the AIC test. It is clear from estimation 7, when the three variables are included in the regression model, that the AIC has the lowest value. With respect to the SIC test, the results show that SIC has the lowest value when two variables are included in the model (estimation 5), however it is only slightly lower than the one associated with estimation 7. Therefore, with the given set of explanatory variables, the best model is the one that includes GDP, population and time as explanatory variables.

4.3 Autocorrelation of the residuals

The Durbin – Watson Statistics, also presented in table 1, shows no autocorrelation of the residuals from regression analysis. So the estimates of statistical significance of the parameters are not biased.

4.4 Multicollinearity

The explanatory variables in the chosen model are tested for multicollinearity. The correlation between the GDP and time is 0.69, and the correlation between the POP and GDP is 0.80. Both values are not above the limit value set for multicollinearity. Therefore, the multicollinearity problem does not exist in the model when GDP, population and time are used.

4.5 Heteroscedasticity

The chosen model is tested for heteroscedasticity using the Eyeball test and Breusch-Pagan test. The Eyeball test is used to examine the regression model, including the GDP, population and time, for heteroscedasticity. The residuals of the regression model are plotted against the explanatory variables and the predicted value of the inflow. The results are shown in figure 1.

The variability of the residuals appears to be the same across all values of the explanatory variables and of the predicted inflow. It is known that the dispersion of the residuals is increasing if the values of the explanatory variables and of the predicted dependent variable Y are increasing when heteroscedasticity

does exist; in the present model this is not the case. Therefore the assumption of homoscedasticity is not rejected.

The regression model is also tested for heteroscedasticity using the Breusch-Pagan test. The squared residuals are related to the explanatory variables (GDP, population and time). The F-test with a value of 0.26 indicates that all the explanatory variables taken together are not significant. The F-test confirms that the explanatory variables together is not significant, therefore the assumption of no heteroscedasticity is not rejected.

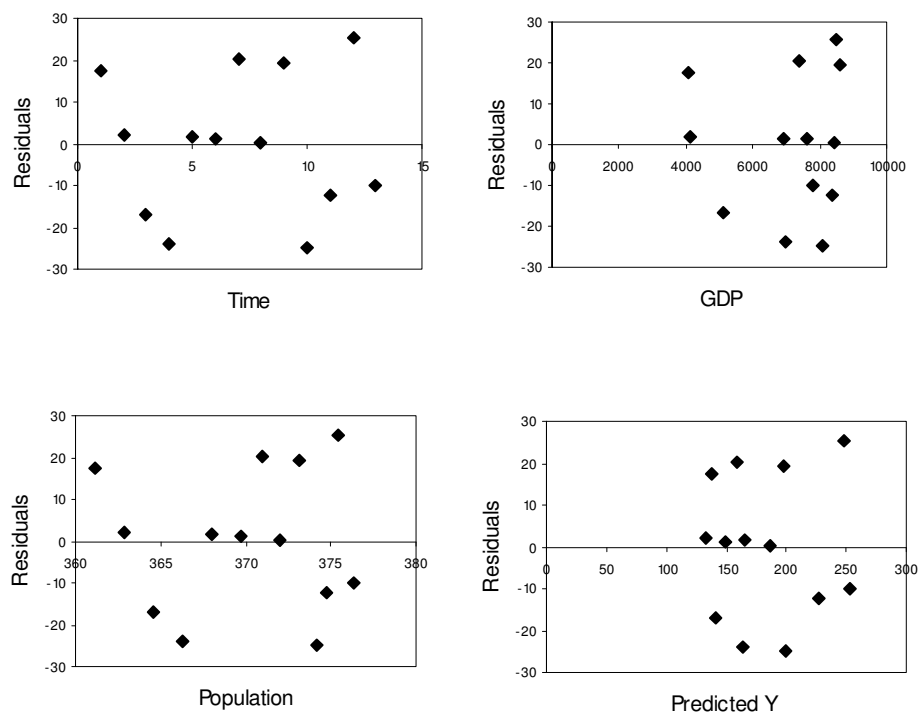


Fig. 1: Plot of the residuals of the regression model against the explanatory variables using the Eeyball test

5 Results when using per capita GDP (GDP/C) as explanatory variable

5.1 The optimal regression model

In this section a new analysis will be performed, based on a more strict preselection of explanatory variables aiming at avoidance of correlation between explanatory variables. More specifically, GDP is replaced by GDP per capita (GDP/C), thus avoiding overlap with POP. The results of the statistical tests when GDP/C is used in the analysis are presented in table 2.

Table 2: Statistical analysis using GDP/C, population and T as explanatory variables

Estim- -ation	Socio-economic variables	R^2	R^2_{adj}	AIC	SIC	D/W	t- value	t- value	t- value	F- value
1	GDP/C	0.46	0.41	91.54	92.11	0.71	3.06			09.37
2	POP	0.68	0.65	84.70	85.26	1.06		4.85		23.49
3	Time	0.77	0.75	80.61	81.17	1.35			6.02	36.24
4	GDP/C, POP	0.70	0.64	86.01	87.14	1.24	-0.74	2.80		11.53
5	POP, Time	0.82	0.79	78.94	80.07	1.71		-1.80	2.86	23.47
6	GDP/C, Time	0.78	0.73	82.17	83.30	1.47	-0.59		3.74	17.21
7	GDP/C, POP, Time	0.85	0.81	78.53	80.22	2.09	1.36	-2.21	3.11	17.56

It is clear from the estimation 7 in Table 2 that when all the three variables are included in the regression model, not only R^2 but also R^2_{adj} has the highest value ($R^2 = 0.85$ and $R^2_{adj} = 0.81$). Comparably, AIC has the lowest value in estimation 7. SIC still has its lowest value in estimation 5 with two explanatory variables. The t-statistics for the individual coefficients shows that in this model the GDP/C and population are not significant. However, the F-statistics indicates that all the three explanatory variables taken together are significant and contribute to the inflow. Therefore the model with three explanatory variables is still the optimal model.

If the results of the model that includes GDP/C are compared with the results of the model, which includes GDP as explanatory variable, we can observe that the use of GDP/C gives slightly better results. However there is no big difference between the two models.

5.2 Autocorrelation of the residuals

The Durbin – Watson Statistics in table 2 shows that there is no autocorrelation in the residuals of regression analysis. So the estimates of statistical significance of the parameters are not biased.

5.3 Multicollinearity

The explanatory variables in the chosen model are tested for multicollinearity. The correlation between the GDP/C and time is 0.66, and the correlation between the POP and GDP/C is 0.77, therefore, the multicollinearity problem does not exist in the model.

5.4 Heteroscedasticity

The chosen model is tested for heteroscedasticity using the Eyeball test and Breusch-Pagan test.

The results of the Eyeball test are shown in figure 2. The variability of the residuals appears to be the same across all values of the explanatory variables (GDP/C, population and time) and the predicted dependent variable (inflow of CRTs). So also here the assumption of homoscedasticity is not rejected.

When using the Breusch-Pagan test, the F-statistics shows a value of 0.26, which indicates that all the explanatory variables taken together are not significant. The F-statistics confirms that the explanatory variables are jointly not significant, therefore the assumption of no heteroscedasticity is not rejected.

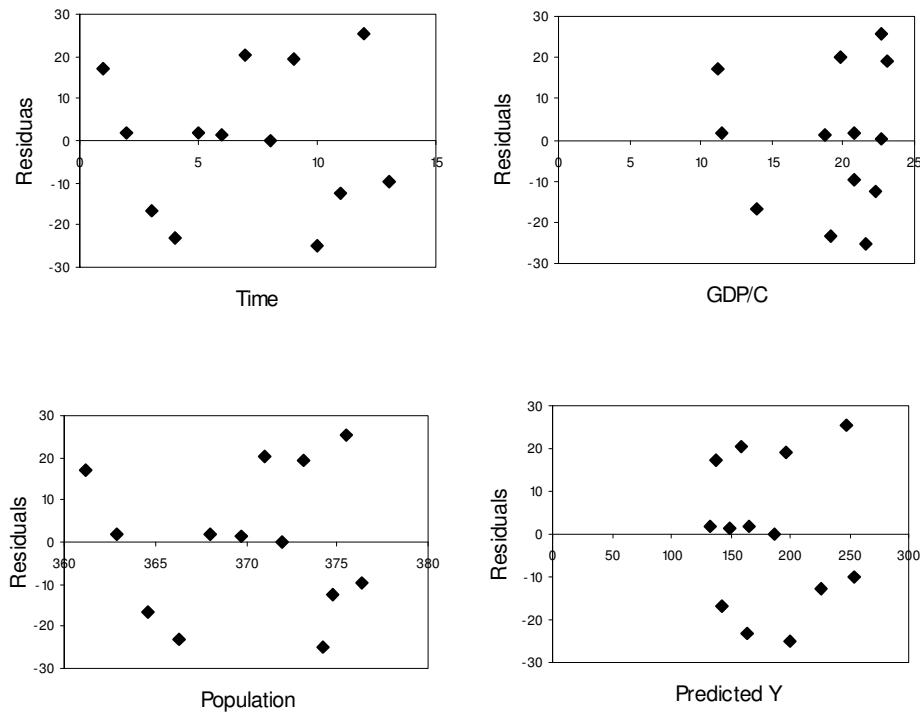


Fig. 2: Plot of the residuals of the regression model against the explanatory variables, when using GDP/C instead of GDP in the regression model

6 Conclusions

The statistical methods presented in this appendix are more elaborate than those used in chapter 4. However, there appears to be no change in the general conclusions drawn from the regression analysis. The best model is the one that includes GDP (or GDP/C), population and time as explanatory variables.

The results of the additional statistical tests in this appendix did validate the results of this model. The multicollinearity problem does not exist in the model. There appears to be no autocorrelation of the residuals from the regression analysis. And the assumption of homoscedasticity is not rejected. Thus the model including the three explanatory variables can be used.

Although the results obtained when GDP/C is used in combination with other explanatory variables in the regression model appeared to be slightly better than those when GDP is used, there is no big difference between the outcomes of the two models. Therefore the results of chapter 4 still hold.

References

- Burney N., (1995). Socioeconomic development and electricity consumption. *Energy Economics*, 17 (3), 85-195.
- Egelioglu, F., Mohamad, A. and Guven, H. (2001). Economic variables and electricity consumption in Northern Cyprus. *Energy* 26, 355-362.
- Gijbels, I and Rousson, V. (2001). A non parametric least square test for checking a polynomial relationship. *Statistics and Probability Letters*, 51 (3), 253-261.
- Glantz, S.A. and Slinker, B.K. (1990). Primer of applied regression and analysis of variance. McGraw-Hill INC., New York, USA.
- Guzman, J., Nishiyama, T., Tilton J. (2004). Trends in the intensity of copper use in Japan since 1960. *Resource Policy*. In press.
- Labson, B.S. and Crompton, P.L. (1993). Common trend in economic activity and metal demand; cointegration and the intensity of use debate. *Journal of environmental Economics Management* 25, 147-161.
- Liu, X.A., Ang, B.W. and Goh, T.N. (1991). Forecasting of electricity consumption: a comparison between an economic model and neural network model. In IEEE International Conference on Neural Networks, Vol. 2, 1254-1259.
- Long, J.S and Ervin, L.H. (2000). Using Heteroscedasticity Consistent Standard Errors in the Linear Regression Model. *The American Statistician* 54: 217-224.
- Mackinnon, J. G. and White, H. (1985). Some heteroscedasticity consistent covariance matrix estimators with improved finite sample properties. *Journal of Econometrics*, 29, 305-325.
- Makridakis, S., Wheelwright, S.C., and Hyndman, R. J. (1998). Forecasting methods and applications. Wiley, New York, USA.
- Mohamed, Z., Bodger, P., 2005. Forecasting electricity consumption in New Zealand using economic and demographic variables. *Energy*, 30 (10), 1833-1843.
- Moore, D. J., Tilton, J. E., Deborah J. S. (1996). Economic growth and the demand for construction materials. *Resources Policy*, 22 (3), 197-205.
- Ranjan, M., Jain, V.K., (1999). Modelling of electrical energy consumption in Delhi. *Energy*, 24, 351-361.
- Roberts, M. C., (1996). Metal use and the world economy. *Resources Policy*, 22 (3), 183-196.
- Todeschini, R., Consonni, V., Mauri, A. and Pavan, M. (2004). Detecting “bad” regression model: multicriteria fitness function in regression analysis. *Analytica Chimica Acta* 515, 1, 199-208.
- White, H. (1980). A heteroscedasticity consistent covariance matrix estimators and a direct test for heteroscedasticity. *Econometrica*, 48, 817-838.

Appendix B – List of symbols in the thesis

Appendix B 1 – List of symbols in chapter 3

$F_{PC,i}^{in}$ is the inflow of a product into the i th product stock.
 $F_{C,i}^{in}$ is the inflow of the substance into the i th product stock.
 SC is the substance content.
 F_C^{in} is the total inflow of the substance into the consumption phase.
 $F_{C,E,i}^{out}$ is the outflow of the substance in the i th product due to emissions.
 $F_{C,D,i}^{out}$ is the outflow of the substance in the i th product due to the delay mechanism.
 $S_{C,i}$ is the stock-in-use of the substance in the i th product.
 F_H^{in} is the inflow of the substance into the hibernating stock.
 F_H^{out} is the outflow of the substance from the hibernating stock.
 F_P^{in} is the input of the substance into primary production processes (the extracted flow).
 F_P^{out} is the outflow of the substance from primary production processes (refined primary substances).
 F_X^{in} is the inflow of the co-produced substances into the economy.
 F_{PP}^{in} is the total inflow of the substance into the production processes of all applications.
 $F_{PP,i}^{in}$ is the inflow of the substance into the production process of the i th product.
 $F_{PP,i,P}^{out}$ is the outflow of the substance in the produced i th product.
 $F_{PP,i,A}^{out}$ is the outflow of the substance to the air during the production process of the i th product.
 $F_{PP,i,W}^{out}$ is the outflow of the substance to the water during the production process of the i th product.
 $F_{PP,i,L}^{out}$ is the outflow of the substance to landfill sites from the production process of the i th product.
 $F_{SC,i}^{in}$ is the amount of waste collected for recycling purposes.
 $F_{inc,land,i}^{in}$ is the inflow of the substance into incineration plants and landfill sites from the discarded i th product.
 $F_{inc,DC,i}^{in}$ is the inflow of the substance into incineration plants from the discarded i th product.
 $F_{land,DC,i}^{in}$ is the inflow of the substance into landfill sites from the discarded i th product.
 $F_{SC,i}^{out}$ is the outflow of the substance from the collection phase.
 $F_{SC,i,E}^{out}$ is the emitted outflow from the collection phase.
 $F_{SC,i,R}^{out}$ is the outflow which goes to the recycling processes.
 S_{sc} is the stock of the substance in the collection phase.
 F_R^{in} is the inflow of the substance into recycling processes
 F_R^{out} is the total outflow of the substance from recycling processes.
 $F_{R,E}^{out}$ is the emitted outflow of the substance.
 $F_{R,W}^{out}$ is the landfilled outflow of the substance.
 $F_{R,CNI}^{out}$ is the waste outflow used in other applications (non-intentional applications).
 $F_{R,R}^{out}$ is the outflow of the refined substance from recycling processes.
 $F_{inc,t}^{in}$ is the total inflow of the substance into incineration plants
 $F_{inc,others}^{in}$ is the inflow of the substance into incineration plants from sources other than the discarded products.
 $F_{inc,t}^{out}$ is the total outflow of the substance from incineration plants.
 $F_{inc,B}^{out}$ is the outflow of the substance from incineration plants in bottom ash.
 $F_{inc,F}^{out}$ is the outflow of the substance from incineration plants in fly ash.
 $F_{inc,A}^{out}$ is the emissions of the substance from incineration plants.
 $F_{land,t}^{in}$ is the total inflow of the substance into landfill sites from all sources.
 $F_{land,others}^{in}$ is the inflow of the substance into landfill sites from sources other than the discarded products.
 $F_{land,t}^{out}$ is the total outflow from landfill sites.
 F_{ST}^{in} is the inflow of the substance into sewage treatment plants.
 $F_{C,E,ST}^{out}$ is the flow of the substance originating from the emissions during the use phase.
 $F_{pp,E}^{out}$ is the flow of the substance originating from the production processes.
 $F_{C,FA,ST}^{out}$ is the flow of the substance originating from the consumption of food and animal products.
 $F_{ST,SS}^{out}$ is the outflow of the substance from sewage treatment processes in sewage sludge.
 $F_{ST,W}^{out}$ is the outflow of the substance from sewage treatment processes in water.
 $F_{SS,S}^{out}$ is the flow of the substance with sewage sludge applied to soil.
 $F_{SS,INC}^{out}$ is the flow of the substance with incinerated sewage sludge.
 $F_{SS,L}^{out}$ is the flow of the substance with landfilled sewage sludge.

$F_{S,Y}^{in}$ is the inflow of the substance through mixed primary resources.
 $F_{S,x}^{out}$ is the outflow of the substance to a specific destination.
 F_S^{in} is the inflow of the substance into agricultural soil.
 $F_{AD,S}^{out}$ is the deposited flow from air into agricultural soil.
 $F_{F,S}^{in}$ is the flow of the substance through fertilizers into agricultural soil.
 $F_{M,S}^{in}$ is flow of the substance through manure into agricultural soil.
 $F_{S,FOOD}^{out}$ is the uptake outflow of the substance from agricultural soil to food
 $F_{S,FODDER}^{out}$ is the uptake outflow of the substance from agricultural soil to fodder
 $F_{S,L}^{out}$ is the leaching outflow of the substance from agricultural soil to water.
 F_{An}^{in} is the inflow of the substance into animal production.
 F_{An}^{out} is the outflow of the substance from animal production.
 $F_{CFA,An}^{in}$ is the inflow of the substance into the consumption of food and animal products from animal production.
 F_{CFA}^{in} is the total inflow of the substance into the consumption of food and animal products.
 F_{CFA}^{out} is the total outflow of the substance from the consumption of food and animal products.
 $F_{CFA,ST}^{out}$ is the outflow of the substance from the consumption of food and animal products to sewage treatment.
 $F_{PP,I,A}^{out}$ is emissions from the production of different applications.
 $F_{inc,A}^{out}$ is emissions from the incineration process.
 $F_{R,A}^{out}$ is emissions from the recycling processes.
 $F_{C,E}^{out}$ is emissions from the use of different applications.
 $F_{NIU,A}^{out}$ is emissions from non-intentional use.
 F_A^{out} is the total outflow of the substance from air.
 $F_{A,DS}^{out}$ is the deposited flow of the substance in soil.
 $F_{A,DW}^{out}$ is the deposited flow of the substance in water.
 $F_{A,TF}^{out}$ is the deposited flow of the substance outside the modelled system.
 $F_{PP,i,W}^{out}$ is emissions from the production of different applications.
 $F_{R,W}^{out}$ is emissions from the recycling processes to water.
 $F_{NIU,W}^{out}$ is emissions from non-intentional use to water.
 $F_{A,DW}^{out}$ is the deposition from air to water.
 $F_{land,W}^{out}$ is the leaching from landfill sites to water.
 $F_{ST,W}^{out}$ is the effluent from sewage to water.
 F_S^{in} is the inflow into the stock in non-agricultural soil.
 F_S^{out} is the outflow from the stock in non-agricultural soil.
 $F_{C,E}^{out}$ is emissions from the use of different applications to non-agricultural soil.
 $F_{NIU,A}^{out}$ is emissions from non-intentional use to non-agricultural soil.
 $F_{A,DS}^{out}$ is the deposition from air to non-agricultural soil.
 $F_{land,S}^{out}$ is the leaching from landfill sites to non-agricultural soil.
 F_{Res}^{in} is the addition to the stock of resources.
 F_{Res}^{out} is the outflow of the substance from the stock of resources.
 F_P^{in} is the inflow of the substance to its primary production processes.

Appendix B 2 – List of symbols in chapter 6

AEEPC is air emissions from electricity production from coal
AEEPO is air emissions from electricity production from oil
AEIN is air emissions from incineration plants
AEOHM is air emissions from other heavy metals production
AEOP is air emissions from oil production
BAC is bottom ash from coal fired power plants
BAIN is bottom ash from incineration
FAC is fly ash from coal fired power plants
FAIN is fly ash from incineration
LOHM is the landfilled stream from the production of other heavy metals
LSM is the landfilled stream of secondary materials (fly ash and bottom ash)
LSS is the landfilled stream of sewage sludge
MPR is mixed primary resources
RZP is refined zinc production
SEOP is soil emissions from oil production
SPIS is iron and steel slag
SS is sewage sludge
WEOHM is water emissions from other heavy metals production
WEOP is water emissions from oil production
WEST is water emissions from sewage treatment plants
WSIA is waste stream from intentional applications
WWOHM is waste water from other heavy metals production

Appendix B 3 – List of symbols in chapter 7

BAC is coal fired power plants bottom ash
BAIN is incineration bottom ash
FAC is coal fired power plants fly ash
FACAG is coal fired power plants fly ash applied in aggregates
FACAS is coal fired power plants fly ash applied in asphalt
FACB is coal fired power plants fly ash applied in building materials
FAIN is incineration fly ash
LFAIN is landfilled fly ash from the incineration plants
OHM is other heavy metals
SPIS is iron and steel slag
SS is sewage sludge
UFAIN is utilized fly ash from the incineration plants

Publications in this thesis

Chapter	Reference
4	Elshkaki, A., Voet, E. van der, Holderbeke, M. van, and Timmermans, V. Dynamic Stock modelling: A method for the identification and estimation of future waste streams and emissions based on past production and product stock characteristics. <i>Energy</i> 2005; 30, 8, 1353-1363.
5	Elshkaki, A., Voet, E. van der, Holderbeke, M. van, and Timmermans, V. The environmental and economic consequences of the developments of lead stocks in the Dutch economic system. <i>Resources, Conservation and Recycling</i> 2004; 42: 133-154.
6	Elshkaki, A., Voet, E. van der, Holderbeke, M. Van, Timmermans, V. Long term consequences of non-intentional flows of substances: Modeling non-intentional flows of lead in the Dutch economic system and evaluating their environmental consequences. (submitted).
7	Elshkaki, A. and Voet, E. van der. Holderbeke, M. Van, Timmermans, V. Long term consequences of non-intentional flows of substances: Long-term consequences of substances presence in utilized secondary materials. (submitted).
8	Elshkaki, A., Voet, E. van der. The consequences of the use of platinum in new technologies on its availability and on other metal cycles. In Loeffe, C. V., (Editor). <i>Conservation and Recycling of Resources: New Research</i> . Nova Science Publisher, ISBN 1-60021-125-9. 2006.

