Bulletin of the Transilvania University of Braşov
Series: Economic Sciences • Vol. 8 (57) No. 1 – 2015

The influence of macroeconomic indicators on the emission of greenhouse gases. Treatment of outliers Case study - România

Evelina GRĂDINARU [1], Lorant BUCS[2], Gabriel BRĂTUCU[3]

Abstract: *This paper implements the multiple linear regression method in order to determine the correlation between a number of independent variables and a dependent variable. It begins with a brief introduction explaining the purpose of this analysis, and continues with the implementation of the econometric model in order to calculate the coefficient of determination that the four significant macroeconomic indicators, namely the amount of energy produced from renewable sources, gross domestic product (GDP), the price of Brent oil barrel on the European market and the energy intensity of the economy have on total emissions of greenhouse gases in Romania. The final part will expose the conclusions of the present analysis.*

Key-words: *econometric model, regression, renewable energy, greenhouse gases.*

1. Introduction

This paper aims to achieve a detailed analysis of the significant impact of several macroeconomic factors on greenhouse gas emissions. The need for such an analysis comes from a major concern that is currently observed both at a global, national and multinational level, for achieving the aspirations of sustainable ecological and economic development. Therefore, it will be demonstrated further along that apparently independent factors have a significant influence on each other with a decisive impact on environmental quality.

2. Objectives of the analysis

It is well known that global warming is a phenomenon whose magnitude has taken alarming proportions, both nationally and globally. Ultimately, the quality of

[1] Transilvania University of Braşov, evelina.gradinaru@gmail.com
[2] Transilvania University of Braşov, lorand.bucs@unitbv.ro
[3] Transilvania University of Braşov, gabriel.bratucu@unitbv.ro

life of each individual depends on the quality of the environment and human influence was noted through its rather negative than positive effects. In order to correct this injustice brought to the planet, a series of measures are being implemented in order to improve the current adverse effects of human lifestyle on the environment.

One of the most important challenges is represented by global warming caused by emissions of greenhouse gases in the atmosphere and measures such as limiting emissions from industry and transport or looking for alternative sources of energy became mandatory in most countries.

The European Union is one of the strongest supporters of reducing greenhouse gas emissions and Romania, as a member of this organization, aligns with European environmental policies. If in some areas, such as GDP and energy consumption, there are still noticeable differences between our country and other member states, in other sectors, such as renewable energy, the results are more than satisfactory. To support this assertion, there is a relatively easy to calculate indicator: the amount of energy produced from renewable sources per GDP. Thus, while the EU average is around 14 tons of oil equivalent energy to 1 million euro GDP, in Romania this value reaches approximately 38 tons.

Given these considerations, the objective of this analysis is to calculate the coefficient of determination that the four significant macroeconomic indicators, namely the amount of energy produced from renewable sources, gross domestic product (GDP), the price of Brent oil barrel on the European market and the energy intensity of the economy have on total emissions of greenhouse gases in Romania.

3. The econometric model

The study considers representative data for the period 2004 -2013, in order to determine the extent to which exogenous variables considered here the production of energy from renewable sources (x1), GDP (x2), oil prices (x3) and the energy intensity of the economy (x4) explain variations of the endogenous variable, which is the production of renewable energy (y).

The table below shows the values of the indicators referred to Romania:

Year	Greenhouse gas emissions [1000 tons CO2 equivalent]	Production of energy from renewable sources [1000 tons of oil equivalent]	GDP[mil. Euro]	European Oil Price FOB [US $ / barrel]	Energy intensity of the economy [kg of oil equivalent per 1000 euro GDP]
	y	x1	x2	x3	x4
2004	141.221	4.593,60	60.800,00	38,26	515,9
2005	141.314	4.984,20	80.225,60	54,57	491,3
2006	144.777	4.831,00	98.418,60	65,16	471,4
2007	142.804	4.717,70	125.403,40	72,44	441,5
2008	139.812	5.336,10	142.396,30	96,94	409,9
2009	119.917	5.274,60	120.409,20	61,74	387,4
2010	115.799	5.708,40	126.746,40	79,61	394,6
2011	121.514	5.027,50	133.305,90	111,26	393,7
2012	118.764	5.242,20	133.806,10	111,63	378,8
2013	116.626	5.500,00	144.282,20	108,56	371,7

Table 1. *Data on analyzed indicators*
Source: Eurostat

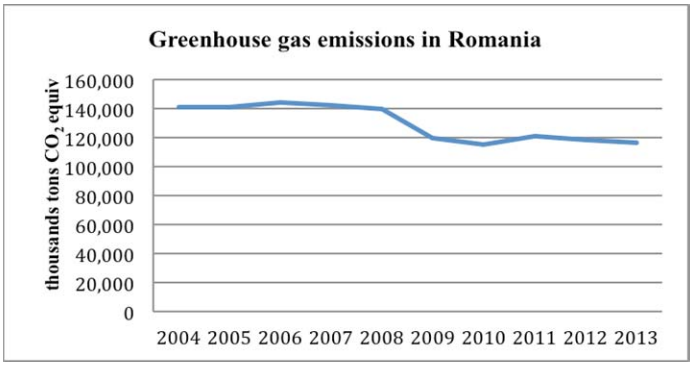

Fig. 1. *Greenhouse gas emissions between 2004-2013 in Romania*

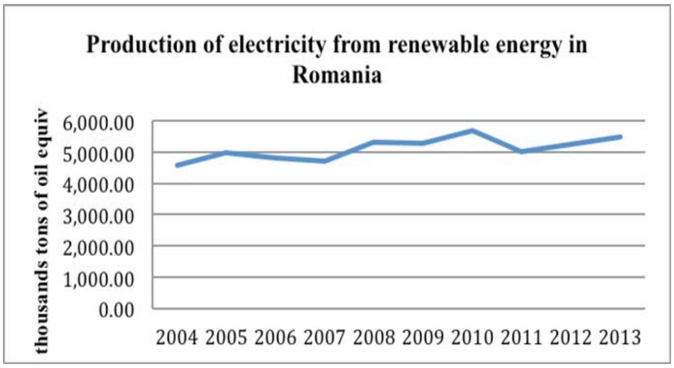

Fig.2. *The production of electricity from renewable energy in Romania between 2004-2013*

It can be seen that the positive trend is accompanied by frequent adjustments in production, so there can be observed both positive and negative variations from year to year in the considered period.

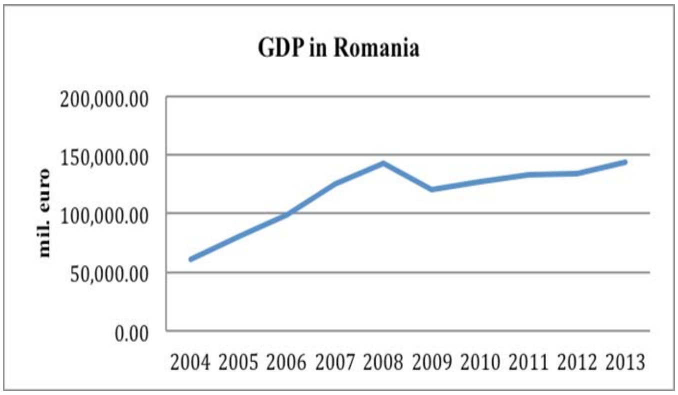

Fig. 3. *Evolution of GDP in Romania between 2004-2013*

Romania confirms its emerging country status with high potential for development. Its GDP recorded a strong growth in the considered period, 2004-2013. However, the economic crisis registered in the middle of the considered period produces larger corrections for Romania.

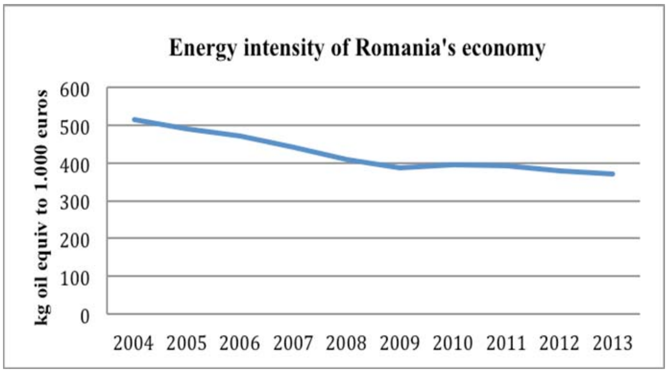

Fig. 4. *Energy intensity of Romania's economy*

In terms of the energy intensity of the economy, Romania is out of the EU average with values exceeding even 3 times the energy consumption per unit of GDP, confirming the highly energy-intensive economy status compared to EU average. Unfortunately, although the trend described by this indicator is similar to the European one, the degree of efficiency of energy consumption is much less evident, further widening the gap of efficiency of our country.

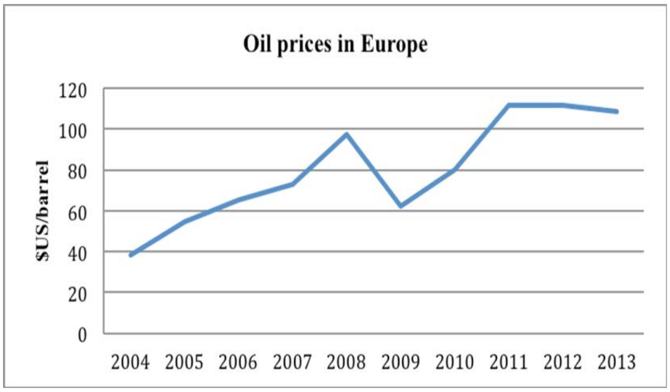

Fig. 5. *Oil prices in Europe between 2004-2013*

In terms of oil prices, for the purpose of this analysis we considered as reference oil price in Europe, given the strong degree of Romania energy interconnection to the European market for oil.

Considering the indicators included in the data analysis, graphics to highlight correlations between each exogenous variable and the studied endogenous variable were done.

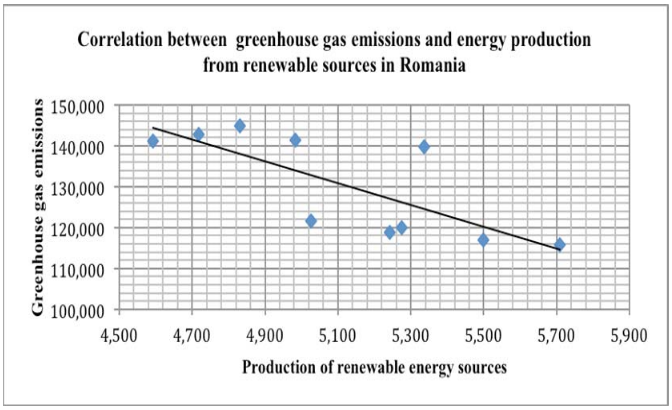

Fig. 6. *Correlation between greenhouse gas emissions and energy production from renewable sources in Romania*

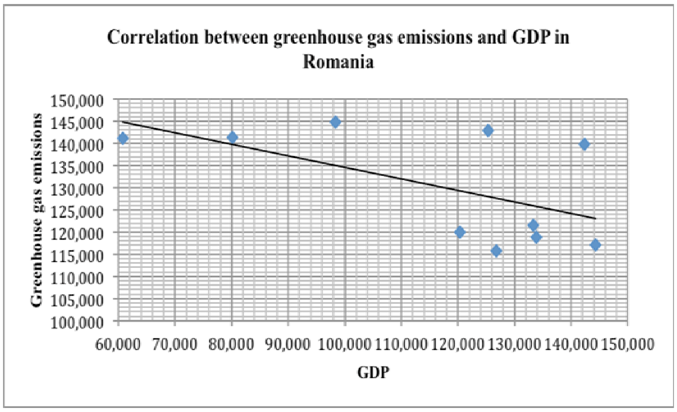

Fig. 7. *Correlation between greenhouse gas emissions and GDP in Romania*

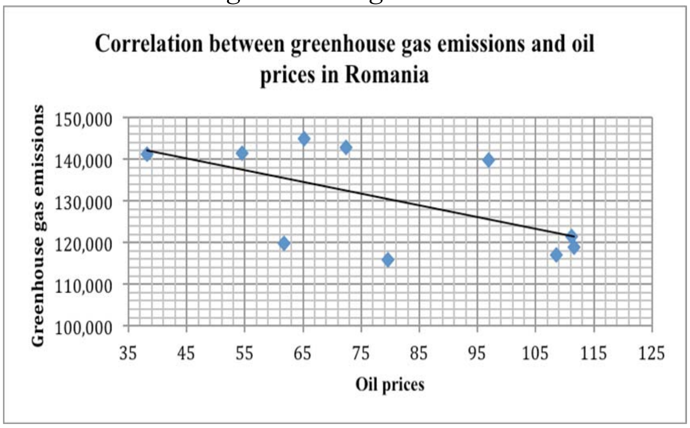

Fig. 8. *Correlation between greenhouse gas emissions and oil prices in Romania*

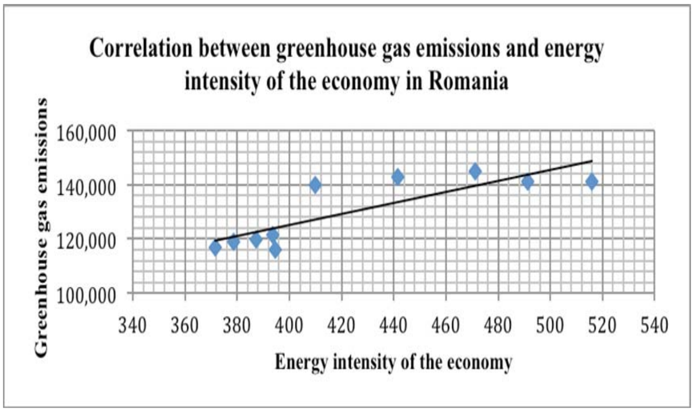

Fig. 9. *Correlation between greenhouse gas emissions and energy intensity of the economy in Romania*

As expected, the strongest influence on the evolution of green house gas emissions is made by the production of energy from renewable sources, and by the energy intensity of the economy. The correlation between the endogenous variable studied and GDP is not to be neglected as well, knowing that value is generated through the consumption of production factors and energy is one of the most important.

4. Results of the research

Using the above data, the regression analysis was made resulted in the following regression table.

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0,9753
R Square	0,9513
Adjusted R Square	0,9123
Standard Error	3697
Observations	10

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression, SSE	4	1333887896	333471974	24,3950	0,0018
Residual, SSR	5	68348404	13669681		
Total, SST	9	1402236300			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*
Intercept	-87957	63969	-1,3750	0,2275	-252393	76480
Production of energy from renewable sources	-5,7820	5,7580	-1,0042	0,3614	-20,5833	9,0193
GDP	0,6171	0,1345	4,5884	0,0059	0,2714	0,9628
Oil prices	-112,9329	98,8792	-1,1421	0,3051	-367,11	141,2442
Energy intensity of the economy	434,5546	75,9150	5,7242	0,0023	239,4088	629,7005

Table 2. *Regression model with 4 variables*

The first part of the table reveals that the regression model is well chosen, explaining in a proportion of about 95% the variation in greenhouse gas emissions - the dependent variable y (R^2 = 0,9813). The value of R indicates a very strong correlation between the endogenous variable y and exogenous variables $x_1 ... x_4$.

The table ANOVA exposes the variation of the x variables and evaluates the overall significance of the regression, reflecting in a globally significant regression with 4 variables, as follows:

SSE (sum of squared regression factors) is greater than SSR (sum of squared residual factors)

F theoretical (for α = 5%) = 5,19 is less than the calculated value for F* = 24,3950 for a significance threshold of 0,18% (*Significance F* = 0,0018)

The third part of the table highlights our regression model identified based on the data used, as follows:

$$\hat{y}_t = -87957-5,7820\, x_1+0,6171x_2-112,9329x_3+434,5546x_4$$

Also, in addition to the global significance test, the regression table does a test of importance for each of the individual exogenous variables.

The Student Ratio (t Stat) for each regression coefficient (in absolute values) is compared with the theoretical t Stat, which for this model has a value of 2,57 (for α = 5% and 5 degrees of freedom).

	t Stat	*Theoretical t Stat*	*P-value*	α
Intercept	-1,3750	2,57	0,2275	0,05
Production of energy from renewable sources	-1,0042	2,57	0,3614	0,05
GDP	4,5884	2,57	0,0059	0,05
Oil prices	-1,1421	2,57	0,3051	0,05
Energy intensity of the economy	5,7242	2,57	0,0023	0,05

Table 3. *Significance test for each individual exogenous variable*

It can be seen that the constant Intercept and the variables "renewable energy production" and "oil prices" fail the test of individual significance. Firstly, student ratio values are below those of the theoretical value of this indicator, and secondly, the indicator P-value for the constant and the two variables indicates a too high significance threshold, respectively 22,75%, 36,14% and 30,51% (this threshold should be less than 5% for the value of the estimator to be significantly different from zero with a probability of 95%).

The confidence intervals for the five estimators for the coefficients of the explanatory variables are shown in the table below:

	Lower 95%	*Upper 95%*
Intercept	-252393	76480
Production of energy from renewable sources	-20,5833	9,0193
GDP	0,2714	0,9628
Oil prices	-367,1100	141,2442
Energy intensity of the economy	239,4088	629,7005

Table 4. *Confidence intervals for the estimators of the 4 variables*

It is obvious that for the variables GDP and energy intensity of the economy the sign of the coefficient is maintained throughout the confidence interval (+ for both, corresponding to a direct link with greenhouse gas emissions), while the estimators for the constant Intercept and for the variables which have failed the test of significance (renewable energy production and oil prices) pass from negative to positive, which means they can be even zero, while the margins indicate values significantly different from zero.

Therefore, from the new regression model explaining the variation of *y* will no longer contain the constant Intercept and the variables "energy production from renewable sources" (x1) and "oil prices" (x3):

$$\hat{y}_t = 0{,}6171x_2 + 434{,}5546x_4$$

The table of regression for the new model with two explanatory variables and constant ZERO, looks like this:

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0,9992
R Square	0,9984
Adjusted R Square	0,8732
Standard Error	5843
Observations	10

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression, SSE	2	170889676845	85444838422	2503	0,0000
Residual, SSR	8	273133263	34141658		
Total, SST	10	171162810108			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*
Intercept	0	#N/D	#N/D	#N/D	#N/D	#N/D
GDP	0,1622	0,0477	3,4032	0,0093	0,0523	0,2721
Energy intensity of the economy	261,9942	13,2984	19,7013	0,0000	231,3282	292,6603

Table 5. *Regression model with 2 variables*

The resulting model indicates a very strong correlation between the explanatory variables and the endogenous variable, the multiple correlation coefficient in this case being very close to 100%. Also, the linear model is valid, as confirmed by the coefficient of determination (R square = 0,9984).

The resulting regression is globally significant, the Fisher test indicating an F* much higher than the theoretical value (4,46), for a significance level of 0%.

The test of significance for the new estimators relative to the null value indicates coefficients significantly different from zero, the value t-State being higher than the theoretical value calculated for a significance level of α = 5% (2,31) at a significance level below 5% (P-value). Virtually all estimators are different from zero with a probability of nearly 100%. This is evidenced also by the confidence intervals for the estimators obtained because none change their sign (no null values) over the entire determined interval.

$$\hat{y}_t = 0,1622x_2 + 261,9942x_4$$

See below the table and graphical representation of greenhouse gas emissions determined with the two models obtained (4 or 2 variables) compared to the actual values:

Year	Greenhouse gas emissions [1000 tons CO2 equivalent]	Greenhouse gas emissions [1000 tons CO2 equivalent]	Greenhouse gas emissions [1000 tons CO2 equivalent]
	y	y theo t1	y theo t1
2004	141.221	142.868	145.025
2005	141.314	140.065	141.730
2006	144.777	142.335	139.468
2007	142.804	145.827	136.011
2008	139.812	136.239	130.488
2009	119.917	117.224	121.027
2010	115.799	119.737	123.941
2011	121.514	123.756	124.769
2012	118.764	116.307	120.947
2013	116.626	118.542	120.786

Table 6. *Real greenhouse gas emissions and adjusted greenhouse gas emissions determined using the two regression models*

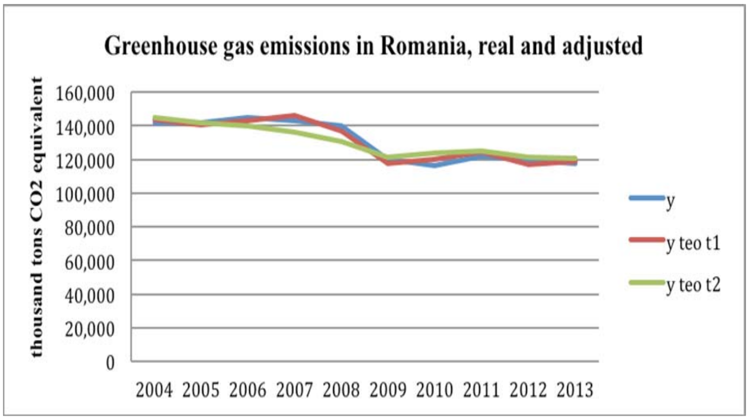

Fig. 10. *Greenhouse gas emissions in Romania, real and adjusted*

5. Outliers

The purpose of this analysis is to identify the correlation between greenhouse gas emissions and European oil prices, the latter being one of the most representative benchmarks that define the demand for renewable energy worldwide.

Year	Greenhouse gas emissions [1000 tons CO2 equivalent]	Oil prices [dollar/barrel]	Greenhouse gas emissions [1000 tons CO2 equivalent]	Deviations
	y	x3	y theo	ei
2004	141.221	38,26	142.024	-803
2005	141.314	54,57	137.442	3.872
2006	144.777	65,16	134.466	10.311
2007	142.804	72,44	132.421	10.383
2008	139.812	96,94	125.537	14.275
2009	119.917	61,74	135.427	-15.510
2010	115.799	79,61	130.407	-14.608
2011	121.514	111,26	121.514	0
2012	118.764	111,63	121.410	-2.646
2013	116.626	108,56	122.273	-5.647

Table 7. *Data regarding greenhouse gas emissions and oil prices before eliminating the outliers*

Previously confirmed by fig. 8, an inverse determination is noted between the two indicators so that a drop in oil prices will lead to increased emissions of greenhouse gases. However, in this case there are certain "extreme" points that deviate from the regression line, which will be highlighted and subsequently removed using the following regression model.

Regression Statistics	
Multiple R	0,5851
R Square	0,3423
Adjusted R Square	0,2601
Standard Error	10737
Observations	10

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression, SSE	1	479967588,2	479967588	4,1634	0,0756
Residual, SSR	8	922268711,4	115283589		
Total, SST	9	1402236300			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*
Intercept	152774	11529	13	0	126187	179361
Oilprices	-280,9628	137,6977	-2,0404	0,0756	-598,4942	36,5686

Table 8. *Regression model with 1 variable*

The first part of the regression table highlights that oil prices explain the variation on greenhouse gas emissions at a rate of only 34,23%.

The regression ANOVA table reflects a reduced global significance, as follows:

- SSR is greater than the SSE
- Theoretical F (for α = 5% and degrees of freedom 1 and 8) = 5,32 which is greater than the calculated value F* = 4,1634 for a significance level of 7,56% (Significance F = 0,0756)

The third part of the table highlights our regression model identified based on the data used, as follows: $\hat{y}_t$= 152774 - 280,9628xt

The Student Ratio (t State) for this model has a value of 2,31 (for α = 5% and 8 degrees of freedom) and highlights that oil prices are not a significant explanatory variable for the variation of greenhouse gas emissions. Also, P-value is greater than 5% which means that the probability of the variable xt taking a coefficient significantly different from zero is below 95%, a fact emphasized also by the confidence interval that has margins of different signs.

The above detailed indicators show the existence of so-called "outliers", as it could be seen from the graph that shows the correlation. In order to precisely identify them, calculations have been made in order to reveal the differences between the value of the variable y resulting from the regression model (y theo) and the actual values (see Table 7).

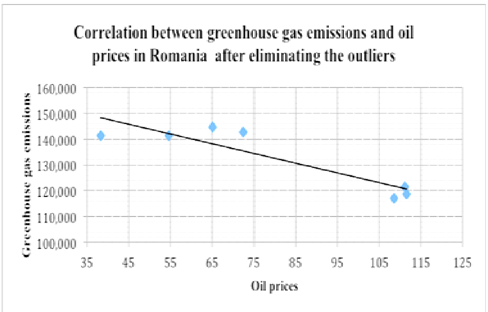

Fig. 11.*Correlation between greenhouse gas emissions and oil prices in Romania after eliminating the outliers*

The graphic representation of the correlation reflects a higher concentration of cloud points around the regression line and thus a stronger correlation between the greenhouse gas emissions and European oil prices.

The model obtained shows a more significant overall regression of 81,64%, and a stronger correlation between the two variables studied, of more than 90%.

Also, the sum of the squared regression factors (SSE) is much greater than the sum of squared residual factors (SSR) and F*>Ftheo (6,61 for α = 5% and degrees of freedom 1 and 5) for a significant overall regression starting from the threshold of 0,53%.

The regression model identified, namely: $\hat{y}_t$= 162904 - 378,9609xt can be considered good considering the significance of the variable xt given by the student test and an almost 100% probability that its coefficient is significantly different from zero. The confidence interval finishes with negative values on both sides.

6. Conclusions

The everyday activity of the modern man requires constant energy consumption. We refer here not to each person's own energy, but to that external energy so necessary in the daily life, be it in the form of economic, recreational or creative activities. Unfortunately, we are still dependent on traditional energy sources, and while we make every effort to get this energy from long-term, sustainable sources, the needs of humanity are constantly expanding, due to a growing population but also due to the economical development of more and more states.

To reflect the impact of human activity on the environment we used as the analyzed variable greenhouse gas emissions. This was then placed in conjunction with a range of macroeconomic indicators reflecting on the one hand the level of development (GDP) and the energy necessary to produce it (energy intensity of the economy), and on the other hand the efforts to extract energy from alternative sources (that are also environmentally friendly - energy production from renewable sources) and the cost of the one of the most traditional and well known energy source (oil price).

The result of this analysis was that all the indicators selected are globally significant in determining how much greenhouse gases are left in Romania's atmosphere. Going further into detail, we could observe some inconsistencies in the influence of oil prices that seem to have a direct influence on the emissions (which is invalid both logically and economically, considering that if oil prices decreased, this should result in an increase in exhaust emissions due to increased demand for this resource and the zero economic advantages for the use of renewable energy). And indeed, the coefficient of the variable "oil prices" would go from negative to positive, with a significant probability of being even zero. In this case it is clear that the significance of regression is adversely affected by this variable. The same line of reasoning applied, somewhat surprisingly, for the variable "energy production from renewable sources".

The constant Intercept had the same flaws, the result being a significant increase of the functions validity if the curve obtained should pass through the origin.

The resulting model had a significance of above 99%, which validates the claim that in Romania the emissions of greenhouse gases is explained in a very high proportion by GDP and the energy intensity of the economy.

The presence of GDP is totally explainable among the crucial variables because, although its impact on the environment is not directly quantifiable, obtaining it requires a great deal of energy, which is the greater as GDP increases.

Given the fact that from the final equation oil prices were eliminated, we attempted to make a further analysis considering only this variable and its influence on greenhouse gas emissions but eliminating the outliers affecting the correlation. The result was that removing a relatively small number of values from the initial statistical series, the regression has become increasingly significant, being one of obvious indirect influence. Nevertheless, a correlation of such low intensity between these two essential indicators (and apparently very tight connected) was quite surprising, demonstrating furthermore that the interconnections which govern the factors with a significant impact on the environment are much more complex (and sometimes unexpected) than we think.

7. References

Anghelache Constantin, Mitrut Cristina, (coordinators), Bugudui Elena, Deatcu Catalin, Dumbrava Madalina. 2009. *Econometrie*. *Teorie, sinteze si studii de caz*., Bucharest, 3rd Edition, Artifex.

Duguleană, Liliana. 2012. *Statistica în cercetare*. Brasov: Transylvania University Press, PhD School.

Duguleană, Liliana. 2011. *Metode de previziune economică*. Brasov: Transilvania University Press.

Petcu, Nicoleta. 2009. *Statistica.* Brasov: Transilvania University Press

http://ec.europa.eu/eurostat