DOI: 10.2428/ecea.2012.19(11)134 ECOL CHEM ENG A. 2012;19(11):1383-1391

Alicja KOLASA-WIĘCEK[1]

REGRESSION MODELING OF AGRRICULTURE GREENHOUSE GASES EMISSIONS IN POLAND

REGRESYJNE MODELOWANIE ROLNICZYCH EMISJI GAZÓW CIEPLARNIANYCH W POLSCE

Abstract: Agricultural greenhouse gases emissions are mainly produced in direct emissions from plant and animal production as well as those associated with land use changes. Studies attempt to describe the variables correlated with agricultural greenhouse gas emissions using linear regression. The analysis covered two groups of independent variables such as the main crops and livestock. The analysis included the last 20 years and variables were set using Pearson's linear correlation. The resulting model concerns 87.5 % of the variability of agricultural greenhouse gases emissions, by cattle, horses, and rye. The study was conducted using the statistical package R-Project.

Keywords: greenhouse gases, agriculture emissions, linear regression, modeling, R-Project, livestock production, crops

Introduction

Agriculture constitutes a serious share in gas emissions. The dynamic growth of human population causes an increase in food demand, which in turn enforces an intensification of agricultural production and the associated environmental degradation.

Agricultural greenhouse gases (GHG) emissions are mainly produced in direct emissions from plant and animal production as well as those associated with land use changes. It should be emphasized that a significant indirect share in other economic sectors of manufacturing products for agriculture is a necessity. A significant negative contribution involved processes associated with livestock production [1]. As indicated by estimates the share of livestock production may be from 50 to 80 % [1]. Intensive livestock production has always been an excessive burden on the environment. 3/4 of

[1] Department of Economics and Regional Research, Faculty of Economy and Management, Opole University of Technology, Waryńskiego 4, 45–047 Opole, Poland, phone: +48 77 454 35 33 ext. 42, fax: +48 77 453 04 71, email: a.kolasa-wiecek@po.opole.pl

the world's agricultural land is devoted either to the production of animal feed or for grazing. A significant threat is currently observed, the displacement of the traditional method of livestock and small commercial farms by advanced industrial systems, which only leads to the intensification of rapid environmental degradation and an increase in hazardous pollutants. About half of manufactured meat is produced using factory farming methods of animals [2].

The sources of greenhouse gas emissions from agriculture in Poland include: enteric fermentation of animals (CH_4), animal manure (CH_4, N_2O), agricultural soils (N_2O), and the burning of plant residues (CH_4, N_2O). These gases represent a much larger global warming potential than emitted in the greatest amounts the CO_2. Their concentration is increasing rapidly. These gases have the ability to persist in the atmosphere for many years – up to 12 years for CH_4 and N_2O up to 120 years [3]. Besides that they are characterized by a considerably higher rate than the CO_2 of global warming GWP (*Global Warming Potential*) and absorbing the heat 23 times more intensely in the case of CH_4 and about 300 times for N_2O [4]. The biggest changes in emission reduction in Poland occurred in 1989–1993 when economic transformations were taking place [5].

N_2O emissions in Poland are predominantly in agriculture and in 2009, 75 thousand Mg (tons) of N_2O was emitted, of which approximately 77 % is accounted as from agricultural soils and about 13 % from livestock manure. In the case of CH_4 – 587 thousand Mg (tons) of the gas was emitted, and around 75 % came from enteric fermentation and 15 % from livestock manure. A fractional part is attributed to the burning of agricultural waste [6].

The size of CH_4 and N_2O emissions in recent years is presented in Fig. 1 and 2.

The necessity for modeling possible greenhouse gases emissions in relation to current and future crop and livestock production has become particularly important in recent years. A major challenge at the present time is to adjust agriculture to evolving climatic conditions and to support actions to reduce these changes. Studies on the impact of agricultural production on the environment, especially greenhouse gas emissions from this sector, is currently the issue undertaken by scientists and policy makers around the world [7–14].

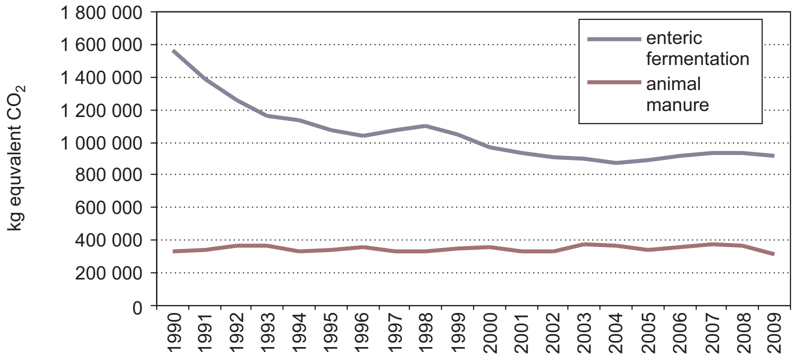

Fig. 1. CH_4 emissions from agriculture in the years 1990–2009 in CO_2 equivalent by major category

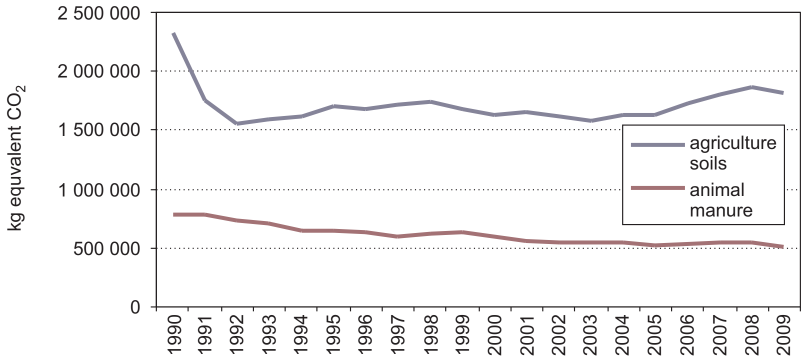

Fig. 2. N_2O emissions from agriculture in the years 1990–2009 in CO_2 equivalent by major category

Modeling is important both from a scientific perspective and practical, economic exploitation. In this context, efforts being made to bring additional information to this particular field of knowledge seem justified.

Only a comprehensive approach to land use and farming, including the principles of sustainable development, can bring a permanent solution to the problem of agricultural emissions.

Methodology and tools

This research attempts to describe the chosen factors responsible for agricultural GHG emissions (CH_4 and N_2O) using linear regression. Regression methods have been successfully used in similar studies which examined the influence of crops and husbandry and/or the participation of other factors connected with agricultural emissions [15]. The research was conducted at farm level. Testing covered, for example: the share of the green soil in the total area of agricultural land, livestock density, farmers education, the amount of nitrogen lodged with purchased fertilizers, the amount of nitrogen lodged with purchased feed, the amount of nitrogen lodged with purchased breeding animals [16] or the share of arable land in the area of farms, number of days of grazing, the livestock density [17].

There are a number of direct or indirect variables affecting the emissions. The study analyzes efforts to describe the impact of two important groups of independent variables such a major crops in Poland (wheat, barley, canola, rye, triticale, corn, oats, and use of mineral fertilizers) and livestock production (cattle, pigs, poultry, horses). The study used the Central Statistical Office [6], the Food and Agricultural Organization [18], the International Fertilizer Industry Association [19] and the United Nations Framework Convention on Climate Change [20] databases. The analysis included the last 20 years.

Correlations relative to the dependent variable (agricultural GHG emissions), were determined using Pearson's test. It is a commonly applicable linear correlation coefficient and is used in cases where both variables are measurable and have a distribution close to normal, and the relationship is linear. If the absolute value of the correlation coefficient is closer to 1, then the correlative relationship between the variables is stronger.

Linear regression analysis which was used in research assumes that between the input and output variables, there is a linear relationship. A straight line should reflect the best fit to the data set [21]. Most often the classical least squares method and its derivatives are used for this purpose. This method is the oldest and the easiest to implement [22]. The study was conducted using the statistical package R-Project.

Results and discussion

A sample result returned from the Pearson correlation test and generated by program R is given in Fig. 3. In case a), the p-value complies with the hypotheses and is lower than the established significance level*,* which is 0.05.

a)
```
cor.test(Cattle_Pcs,Agri.emission.GHG_Gg,method="pearson",
alternative = "greater")

        Pearson's product-moment correlation

data:  Cattle_Pcs and Agri.emission.GHG_Gg
t = 8.6262, df = 18, p-value = 4.122*10^-8
alternative hypothesis: true correlation is greater than 0
95 percent confidence interval:
 0.7854564 1.0000000
sample estimates:
      cor
0.8973402
```

b)
```
cor.test(Rapeseed,Agri.emission.GHG_Gg,method="pearson",
alternative = "greater")

        Pearson's product-moment correlation

data:  Rapeseed and Agri.emission.GHG_Gg
t = -0.3929, df = 18, p-value = 0.6505
alternative hypothesis: true correlation is greater than 0
95 percent confidence interval:
 -0.4553343  1.0000000
sample estimates:
        cor
-0.09221111
```

Fig. 3. The exemplary results of Pearson's test: a) fulfilling the assumptions, b) not fulfilling the assumptions, where: t – value of t statistics investigating the significance of the correlation coefficient, df – the size of the group, p-value – boundary of significance level

Table 1 presents the others linear correlation coefficients for the rest considered variables.

As follows from the p-value levels and correlations, the Pearson test showed positive correlations with the test variable for such variables like: cattle, horses, sugar beet, and oats. Other variables like poultry, fertilizer use, wheat, and rapeseed were not statistically significant and they have been omitted in the next step of studies. In the first

Table 1

Pearson test result for all variables

Variables	t	p-value	Correlations
Cattle	8.6262	$4.122 \cdot 10^{-8}$	0.897
Poultry	–1. 6888	0.9457	–0.37
Horses	4.8174	$6.903 \cdot 10^{-5}$	0.75
Pigs	2.2179	0.01983	0.463
Fertilizer use	–3. 0141	0.9963	–0.579
Pasture&meadows	1.7006	0.05311	0.372
Wheat	–0.1542	0.5604	–0.036
Barley	1.9099	0.03611	0.41
Triticale	–1.5996	0.9365	–0.353
Sugar beet	2.4844	0.01152	0.505
Oats	6.9028	$9.361 \cdot 10^{-7}$	0.852
Maize	–3.1109	0.997	–0.591
Rapeseed	–0.3929	0.6505	–0.092
Rye	2.3795	0.01430	0.489

approach a model for all variables satisfying the conditions was formulated. Some variables proved to be statistically insignificant in this regression model (Fig. 4a).

A rejection of statistically insignificant variables results in a higher coefficient of determination. In case 4b the Adjusted R-Squared takes into account the number of variables in the model, which is higher and amounts to 0.875.

The function lm (linear model) performs a linear model adjustment, appoint residua, adjusted coefficients of the model (where the Estimate – represents the assessment of the value of the regression coefficients, Std. Error – provides information about the standard error of this assessment, t-value is the value of statistics test for this factor and Pr(>|t|) – is a p value determined for the t-student test). Multiplate R-Squared coefficient of determination indicates a compatibility model fits to the actual data and a proportion of the total variability of the dependent variable is explained by the resulting model. The value of Adjusted R-Squared is usually smaller than the previous one. The estimated model can be written as:

$$\textit{Agricultural emission GHG} = 1{,}740e{+}04{+}4{,}251e{-}03{*}cattle{-}6{,}911e{-}03{*}horses{-}2{,}598e{-}03{*}rye \quad (1)$$

The resulting model explains over 87.5 % of the variability of agricultural GHG emissions. The remaining 12.5 % can be explained by others variables not included in this study. The obtained results indicate the highest proportion of cattle. It is well known that, cattle-farming creates a substantial emission of mostly CH_4 and nitrogen compounds. The dominating source of emissions of ammonia NH_3 is livestock production [23]. Ruminant animals emit much more gases than those with single

a)

```
lm(Agri.emission.GHG_Gg~Cattle_Pcs+horses+pigs+pastures.meadows+Barley+
Sugar.beet+Oats+Rye_Ha,data=dane2)
summary(modelPP)

Call:
lm(formula = Agri.emission.GHG_Gg ~ Cattle_Pcs + horses + pigs +
pastures.meadows + Barley + Sugar.beet + Oats + Rye_Ha, data = dane2)

Residuals:
     Min       1Q   Median       3Q      Max
-2079.66  -868.08    46.62   748.24  2133.40

Coefficients:
                   Estimate Std. Error t value Pr(>|t|)
(Intercept)       2.113*10^-4 9.812*10^3 2.154  0.05427 .
Cattle_Pcs        4.397*10^-3 1.055*10^-3 4.169  0.00157 **
horses           -5.605*10^-3 4.873*10^-3 -1.150  0.27446
pigs             -2.304*10^-4 2.808*10^-4 -0.821  0.42931
pastures.meadows  6.165*10^-4 2.419*10^-3 0.255  0.80355
Barley           -1.971*10^-3 7.836*10^-3 -0.252  0.80603
Sugar.beet       -2.861*10^-3 1.307*10^-2 -0.219  0.83075
Oats             -5.096*10^-4 1.754*10^-2 -0.029  0.97734
Rye_Ha           -2.649*10^-3 2.846*10^-3 -0.931  0.37196
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1481 on 11 degrees of freedom
Multiple R-squared: 0.9055,     Adjusted R-squared: 0.8368
F-statistic: 13.18 on 8 and 11 DF,  p-value: 0.0001245
```

The following result was obtained after removing them from the model:

b)

```
modelPP<-
lm(Agri.emission.GHG_Gg~Cattle_Pcs+horses+Rye_Ha,data=dane2)
summary(modelPP)

Call:
lm(formula = Agri.emission.GHG_Gg ~ Cattle_Pcs + horses + Rye_Ha, data = dane2)

Residuals:
     Min       1Q   Median       3Q      Max
-2017.16  -748.16   -23.37   749.11  2505.53

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)  1.740*104 2.648e+03   6.572 6.43*10-6 ***
Cattle_Pcs   4.251*10-3 6.191*10-4   6.867 3.79*10-6 ***
horses      -6.911*10-3    3.853*10-3  -1.794   0.0918 .
Rye_Ha      -2.598*10-3 1.083*10-3  -2.399   0.0290 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1296 on 16 degrees of freedom
Multiple R-squared: 0.8947,     Adjusted R-squared: 0.875
F-statistic: 45.32 on 3 and 16 DF,  p-value: 4.797*10-8
```

Fig. 4. The results of linear model

chamber stomach [24]. The largest amount of nitrogen excreted in manure comes from cattle in the first place, especially cows with high milk efficiency than from horses [23]. Estimates of greenhouse gas emissions in Poland are presented [25], and indicate that a high proportion of CH_4 emissions are from enteric fermentation of cattle – 79 %. The second source is anaerobic manure decay responsible for 10–15 % of the emissions [1]. The last one important factor in the model is the share of rye. Analysis of the rye crop in the last 20 years indicates that the beginning of the 90s was characterized by the highest areas cultivated in Poland. For instance the size of wheat areas in 1990 amounted to

2.28 million hectares and rye 2.31 million hectares [5]. Rye crop area over the last two decades has been declining. The share of this variable in the model can be justified by the highest CH_4 emission factor for this type of grain (Table 2) [26].

Table 2

CH_4 emission factor values for major crops in Poland

Grains	CH_4 emission factor in kg/TJ
Wheat	0.1816
Barley	0.1473
Maize	0.0367
Oats	0.1506
Rye	0.2004
Other grains	0.1618

Based on the modeling results the observed relationships between variables were presented. For this purpose, a scatterplot was prepared, which is a graphical representation of the correlation and facilitates visual assessment to determine the strength and type of relationships between variables. If the correlations of the test points are grouped along a hypothetical straight line – a regression line or a curve, this illustrates the existence of a relationship. When the increase of the independent feature value generates growth of the independent variable we can talk about the directly proportional connection. Otherwise, there is an inverse relationship. Using the scatterplot function it was possible to generate a graphical interpretation of the variables (Fig. 5).

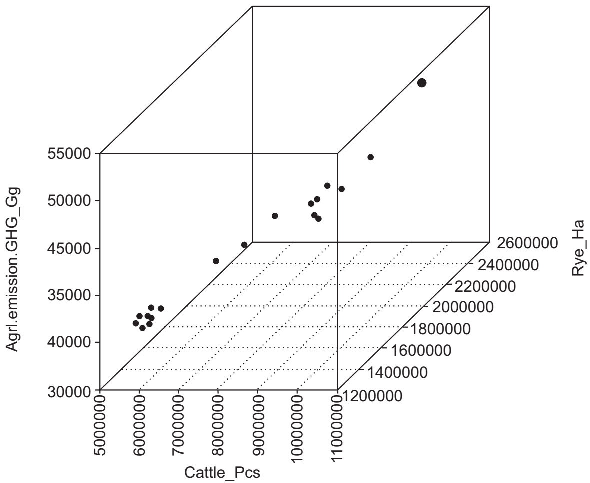

Fig. 5. Scatterplot 3D for GHG emissions vs cattle and rye

With the increase of both variables – cattle and rye, agricultural emissions increasing.

Conclusion

In order to properly shape environmental policy, its inventories and modeling the effects of efforts purposive to improve air quality, appear to be an adequate instrument of the assessment.

The variables provided by the Pearson test correlating with the dependent variable were used to build a linear regression model. Based on the coefficient of determination it can be concluded that the model explains 87.5 % of the variability of agricultural GHG emissions. The most important model variables are cattle and horse production, and also rye cultivation. The remaining 12.5 % can be attributed to the variables, which were not included in this study. The resulting form of the model appears to be justified because of high emissions, which is a breeding of these two groups of animals. Participation in the model of rye cultivation substantiated a high rate emission of CH_4 which is characteristic of this kind of grain.

References

[1] Roszkowski A. Technologie produkcji zwierzęcej a emisje gazów cieplarnianych. Probl Inż Roln. 2011;2:83-97.
[2] Oenema O, Ouendag DA, Velthof GL. Nutrient losses from manure management in the European Union. Livestock Science. 2007;112:261-272. DOI: 10.1016/j.livsci.2007.09.007.
[3] Zaliwski AS. Emisja gazów cieplarnianych przez rolnictwo. Studia i raporty IUNG-PIB. 2007;4:35-47.
[4] Greenhouse Gas Mitigation Potential in U.S. Forestry and Agriculture, Chapters 1–4, Washington: United States Environmental Protection Agency Office of Atmospheric Programs; 2005.
[5] Bebkiewicz K, Cieślińska J, Dębski B, Kanafa M, Kargulewicz I, Olecka A, Olendrzyński K, Skośkiewicz J, Żaczek M. Krajowy Raport inwentaryzacyjny 2011, Inwentaryzacja gazów cieplarnianych w Polsce dla lat 1988-2009. Warszawa: KOBIZE; 2011.
[6] Ochrona środowiska 2011. Warszawa: GUS; 2011.
[7] Li C, Narayanan V, Harriss RC. Model estimates of nitrous oxide emissions form agricultural lands in United States Global. Bioch Cyc. 1996;10(2):297-306. DOI: 10.1029/96GB00470.
[8] Dämmagen U, Hutchings NJ, The Assessment of emissions of nitrogen species froam agriculture using the methodology of the Atmospheric Emission Inventory Guidebook. In: European Emissions from agriculture. Kuczyński T, Dämmagen U, Weeb J, Myczko A, editors. Wageningen Academic Publishers. The Netherlands. 2005;71-96.
[9] Jugowar JL. Ekspertyza Aktualne kierunki badań nad redukcją gazów i odorów z budynków inwentarskich. Poznań: IBMER w Warszawie Oddział w Poznaniu; 2008.
[10] Klimont Z, Brink C. Modelling of Emissions of Air Pollutants and Greenhouse Gases from Agricultural Sources in Europe. Interim Report IR-04-048. Laxenburg: International Institute for Applied Systems Analysis; 2004.
[11] Leip A, Marchi G, Koeble R, Kempen M, Britz W, Li C. Linking an economic model for European agriculture with a mechanistic model to estimate nitrogen and carbon losses from arable soils in Europe. Biogeochemistry. 2008;5:73-94. DOI: 10.5194/bg-5-73-2008.
[12] Myczko A, Karłowski J, Szulc R., Detailed Investigations of Methan and Dinitrogen Oxide Emissions form Enteric Fermentation and from Animal Excreta. In: Emission of agricultural risks to health and environment. Myczko A, editor. Center of Excellence Tragen. 2003; 119-126.
[13] Smith WN, Grant BB, Desjardins RL, Worth D, Li C, Boles SH, Huffman EC. A tool to link agricultural activity data with the DNDC model to estimate GHG emission factors in Canada. Ecosys Environ. 2010;136(3-4):301-309. DOI: 10.1016/j.agee.2009.12.008.
[14] Frolking SE, Harriss RC, Terry RE. Modeling nitrous oxide emissions from agriculture: A Florida case study. Chemosphere. 1994;28(7):1401-1415. DOI.org/10.1016/0045-6535(94)90081-7.

[15] Verbruggen I, Carlier L, Van Bockstaele E. Surplus of nutrients on dairy farms in Belgium. In Grassland and socjety. Proc. 15th General Meet. EGF. 1994:463-465.
[16] Pietrzak S, Sapek A, Oenema O. Ocena emisji podtlenku azotu (N_2O) ze źródeł rolniczych w Polsce. Zesz Edu. 8. Falenty: IMUZ. 2002;23-36.
[17] Giustini L, Acciaioli A, Argenti G. Apparent balance of nitrogen and phosphorous in dairy Mugello (Italy). Ital J Anim Sci. 2007;6:175-185. DOI: 10.4081/ijas.2007.175.
[18] Food and Agricultural Organization [http://faostat.fao.org/site/573/DesktopDefault.aspx?PageID=573#ancor, 17.04.2012].
[19] International Fertilizer Industry Association [http://www.fertilizer.org/ifa/ifadata/serach, 17.04.2012].
[20] United Nations Framework Convention on Climate Change [http://unfccc.int/di/DetailedByGas/Event.do?event=go#, 17.04.2012].
[21] Faraway J J. Practical Regression and Anova using R. July 2002.
[22] Faraway JJ. Linear Models with R, University of Bath, UK, Chapman and Hall/CRS Texts in Statictics Science, 2004.
[23] Pietrzak S. Metoda inwentaryzacji emisji amoniaku ze źródeł rolniczych w Polsce i jej praktyczne zastosowanie, Water-Environ-Rural. 2006;1(16):319-334, 2006.
[24] Podkówka Z, Podkówka W. Emisja gazów cieplarnianych przez krowy. Przegląd hod. 2011;3:1-4.
[25] Zaliwski AS, Purchała L. Oszacowanie emisji podtlenku azotu i metanu z rolnictwa w przekroju województw za 2005. In: Emisja gazów cieplarnianych i amoniaku w rolnictwie. Acta Agrophysica, Rozprawy i Monografie, 4. Czyż EA, Jugowar JL, Sławiński C, editors. Lublin: PAN Lublin; 2007;76-83.
[26] Olendrzyński K, Kargulewicz I, Skośkiewicz J, Dębski B, Cieślińska J, Olecka A, Kanafa M, Kania K, Sałek P. Krajowa inwentaryzacja emisji i pochłaniania gazów cieplarnianych za rok 2007. KASHUE, 2009.

REGRESYJNE MODELOWANIE ROLNICZYCH EMISJI GAZÓW CIEPLARNIANYCH W POLSCE

Katedra Ekonomii i Badań Regionalnych, Wydział Ekonomii i Zarządzania
Politechnika Opolska

Abstrakt: Emisje rolnicze głównych gazów cieplarnianych to ważniejsze bezpośrednie emisje wytwarzane w produkcji roślinnej i zwierzęcej, jak również te związane ze zmianami w sposobie użytkowania gruntów. W badaniach podjęto próbę opisu zmiennych skorelowanych z rolniczymi emisjami gazów cieplarnianych z wykorzystaniem regresji liniowej. Do badań przyjęto dwie grupy zmiennych niezależnych, tj. główne uprawy oraz zwierzęta hodowlane. Analizie poddano okres ostatnich 20 lat. Korzystając z testu Pearsona, wyznaczono liniowe korelacje między zmiennymi. Otrzymany model w 87,5 % wyjaśnia zmienności rolniczych emisji GHG zmiennością udziału bydła, koni oraz żyta. Badania prowadzono z wykorzystaniem pakietu statystycznego R-Project.

Słowa kluczowe: gazy cieplarniane, rolnicze emisje, regresja liniowa, modelowanie, R-Project, produkcja zwierzęca, zboża