Peer-reviewed version available at Environmental Science and Pollution Research 2017, 24, , 6622-6633; doi:10.1007/s11356-016-8321-6

Article

The Impact of Energy, Agriculture, Macroeconomic and Human-Induced Indicators on Environmental Pollution from 1971 to 2011

Samuel Asumadu-Sarkodie * and Phebe Asantewaa Owusu

Middle East Technical University, Northern Cyprus Campus, Mersin 10, Turkey;
phebe.owusu@metu.edu.tr (P.A.O.)
Correspondence: samuel.sarkodie@metu.edu.tr

Abstract: In this study, the impact of energy, agriculture, macroeconomic and human-induced indicators on environmental pollution from 1971 to 2011 is investigated using the statistically inspired modification of partial least squares (SIMPLS) regression model. There was evidence of a linear relationship between energy, agriculture, macroeconomic and human-induced indicators and carbon dioxide emissions. Evidence from the SIMPLS regression shows that a 1% increase in crop production index will reduce carbon dioxide emissions by 0.71%. Economic growth increased by 1% will reduce carbon dioxide emissions by 0.46%, thus supports the environmental Kuznets curve hypothesis that an increase in a country's economic growth leads to a reduction in environmental pollution. An increase in electricity production from hydroelectric sources by 1% will reduce carbon dioxide emissions by 0.30%, thus increasing renewable energy sources in Ghana's energy portfolio will help mitigate carbon dioxide emissions. Increasing Enteric Emissions by 1% will increase carbon dioxide emissions by 4.22% and a 1% increase in the Nitrogen content of Manure Management will increase carbon dioxide emissions by 6.69%. The SIMPLS regression forecasting exhibited a 5% MAPE from the prediction of carbon dioxide emissions.

Keywords: SIMPLS; energy economics; econometrics; carbon dioxide emissions; Ghana

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 26 September 2016 doi:10.20944/preprints201609.0095.v1

Peer-reviewed version available at *Environmental Science and Pollution Research* **2017**, *24*, , 6622-6633; doi:10.1007/s11356-016-8321-6

INTRODUCTION

Climate change as gained global prominence as a result of its long-term effect on the globe. As a result, climate change mitigation through a sustainable global action deems essential to limit the rising levels of greenhouse gas emissions (Asumadu-Sarkodie and Owusu 2016b; Owusu and Asumadu-Sarkodie 2016). This global effort has propelled a lot of research interest in environmental, energy and agricultural sustainability. This paradigm shift in scientific research has increased the interest in using historical data to predict and/or explain the causal-effect between response and predictor variables.

Several studies have employed modern econometric techniques such as; Engle-Granger's method of cointegration (Engle and Granger 1987), Johansen's method of cointegration (Johansen 1995), Dynamic fixed effect model (DFE), Fully-Modified Ordinary Least Squares (FMOLS), Dynamic Ordinary Least Squares (DOLS), Dynamic Panel Data (DPD), Generalized Method of Moments (GMM), Generalized Estimating Equations (GEE), Generalized Least Squares (GLS) Least Squares (LS), Two-Stage Least Squares (TSLS), Limited Information Maximum Likelihood (LIML), Cointegration Regression, Autoregressive Conditional Heteroskedasticity (ARCH), Binary Choice (Logit, Probit, Extreme Value), Ordered Choice, Censored or Truncated Data (including Tobit), Stepwise Least Squares (STEPLS), Robust Least Squares (ROBUSTLS), Heckman Selection (Generalized Tobit), and Least Squares with Breakpoints (BREAKLS), Threshold Regression, Quantile Regression (QREG), Switching Regression (SWITCHREG), Vector Autoregression (VAR) Vector Error Correction Model (VECM) and Autoregressive Distributed Lag (ARDL) to examine the causal relationship between variables in a time series or panel data different disciplines. However, multicollinearity among study variables are problematic in the aforementioned models.

Peer-reviewed version available at Environmental Science and Pollution Research 2017, 24, , 6622-6633; doi:10.1007/s11356-016-8321-6

Two strands of studies are analysed in the study. The first strand of existing literature estimates the causal nexus between carbon dioxide emissions and gross domestic product (GDP), or carbon dioxide emissions and energy-intensity/consumption/production, or a combination of carbon dioxide emissions, energy consumption, GDP and/or population. Huang et al. (2008) examined the causal relationship between energy consumption and GDP in 82 countries using the GMM model which found no evidence of causal relationship between energy consumption and GDP. Soytas and Sari (2009) examined the relationship between CO_2 emissions, energy consumption and GDP using the long-run Granger causality test which found evidence of causal direction running from CO_2 emissions to energy consumption but not valid in the reverse. Lozano and Gutiérez (2008) examined the causal relationship between GDP, CO_2 emissions, energy consumption and population in USA using a non-parametric frontier method which found evidence of steady efficiency increase in the average of the modelled variables for the estimated period. Jammazi and Aloui (2015) examined the relationship between CO_2 emissions, energy consumption and economic growth for 6 countries spanning from 1980-2013 by employing a wavelet-window-cross-correlation approach. Their study found evidence of a bidirectional causality between energy consumption and GDP while there was evidence of a unidirectional causality running from energy consumption to CO_2 emissions. Zhang and Cheng (2009) examined the Granger-causality between CO_2 emissions, energy consumption, GDP, and population. Their study found no evidence of causality between CO_2 emissions and energy consumption and GDP. Gul et al. (2015) examined the relationship between energy consumption and carbon dioxide emissions in Malaysia with a data spanning from 1975-2013 by using the maximum entropy bootstrap method (MEBM). Their study found evidence of a unidirectional causality running from energy consumption to CO_2 emissions. Qureshi et al. (2016) examined the relationship between energy crisis, GHG emissions and GDP for the Caribbean, Europe, Asia, Africa and

Peer-reviewed version available at Environmental Science and Pollution Research 2017, 24, , 6622-6633; doi:10.1007/s11356-016-8321-6

Latin America with a panel data spanning from 1975-2012 by employing Johansen's method of cointegration and variance decomposition. Their study found evidence of negative significant relationship between electricity access and power shortage in certain regions of Europe and Asia. Remuzgo and Sarabia (2015) employed the factorial decomposition to examine the global distribution inequality of carbon dioxide emissions in IEA considered regions with data spanning from 1990-2010. Their study found evidence of causality between GDP per capita and global distribution inequality of CO_2 emissions. Ouyang and Lin (2015) examined the relationship between energy consumption and GDP in China with a data spanning from 1991-2010 by employing the Logarithmic Mean Divisa Index (LMDI) method. Their study found evidence of a long-run equilibrium relationship between CO_2 emissions, energy consumption, labour, industrial sector value added and fossil-fuel consumption. Ozturk and Acaravci (2010) examined the relationship between energy consumption and economic growth in Romania, Albania, Hungary and Bulgaria with a data spanning from 1980-2006 by employing the ARDL bounds testing method which showed evidence of a long-run equilibrium relationship. Their study found evidence of bidirectional causality between energy consumption and economic growth in Hungary. Ozturk and Acaravci (2011) examined the relationship between electricity consumption and economic growth in 11 Middle East and North Africa countries with a data spanning from 1971-2006 by employing the ARDL bounds testing method which showed no evidence of long-run equilibrium relationship for Iran, Morocco and Syria. Their study found evidence of long-run equilibrium relationship in Egypt, Oman, Israel and Saudi Arabia. The overall evidence of the 7 Middle East and North Africa countries showed no evidence of long-run equilibrium relationship. Acaravci and Ozturk (2010) examined the relationship between carbon dioxide emissions, electricity consumption and economic growth in 19 European countries with a data spanning from 1965-2005 by employing the ARDL bounds testing method which

Peer-reviewed version available at Environmental Science and Pollution Research 2017, 24, , 6622-6633; doi:10.1007/s11356-016-8321-6

showed no evidence of long-run equilibrium relationship for Switzerland, Portugal, Iceland, Greece, Italy, Denmark and Germany. Their study supported the validity of the environmental Kuznets curve hypothesis. Apergis and Payne (2011) examined the relationship between renewable energy consumption, labour force, gross fixed capital formation and economic growth in 6 Central American countries with a data spanning from 1980-2006 by employing the FMOLS method which showed evidence of a long-run equilibrium relationship for all the 6 Central American countries. Their study found evidence of a bidirectional causality between renewable energy consumption and economic growth. Asumadu-Sarkodie and Owusu (2016d) examined the multivariate causality analysis of the Kaya factors in Ghana by taking into consideration the total primary energy consumption, GDP, carbon dioxide emissions and population with a data spanning from 1980-2012 using the vector error correction model. Evidence from their study showed a long run causality from population, GDP and energy consumption to carbon dioxide and a bidirectional causality from CO_2 emissions to energy consumption.

The second strand of existing literature estimates the causal nexus between carbon dioxide emissions, energy consumption, GDP and/or population by adding more variables like; foreign investments, industrialization, trade, urbanization, financial development, etc.. Shahbaz et al. (2015) examined the relationship between energy consumption from road transportation, transport sector value added, fuel prices and CO_2 emissions in Tunisia with a data spanning from 1980-2012 by employing the ARDL bounds testing method in the presence of structural breaks which showed an evidence of a long-run equilibrium relationship. Their study found evidence of a bidirectional causality between energy consumption and CO_2 emissions while fuel prices exhibit a unidirectional causality on energy consumption, road infrastructure, CO_2 emissions and transport sector value added. Ben Abdallah et al. (2013) examined the relationship between energy consumption from road

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 26 September 2016 doi:10.20944/preprints201609.0095.v1

Peer-reviewed version available at Environmental Science and Pollution Research 2017, 24, , 6622-6633; doi:10.1007/s11356-016-8321-6

transportation, transport sector value added, fuel prices and CO_2 emissions in Tunisia with a data spanning from 1980-2010 by employing the Johansen's method of co-integration and Granger-causality test which showed an evidence of a long-run equilibrium relationship. Their study refuted the validity of the environmental Kuznets curve hypothesis and found an evidence of unidirectional causality running from fuel price to energy consumption from road transportation. Cerdeira Bento and Moutinho (2016) examined the relationship between CO_2 emissions per capita, non-renewable electricity production per capita, renewable electricity production per capita, GDP per capita and international trade with a data spanning from 1960-2011 in Italy by employing the ARDL bounds testing method in the presence of structural breaks which showed an evidence of a long-run equilibrium relationship. Their study confirmed the validity of the environmental Kuznets curve hypothesis and found an evidence of unidirectional causality running from GDP per capita to renewable electricity production per capita. Seker et al. (2015) examined the relationship between foreign direct investment and environmental quality with a data spanning from 1974-2010 in Turkey by employing the ARDL bounds testing method in the presence of structural breaks which showed an evidence of a long-run equilibrium relationship. Their study confirmed the validity of the environmental Kuznets curve hypothesis and found an evidence of unidirectional causality running from foreign direct investment, GDP per capita and energy consumption to CO_2 emissions. Rafindadi and Ozturk (2015) examined the relationship between natural gas consumption, capital, exports and labour with a data spanning from 1971-2012 in Malaysia by employing the ARDL bounds testing method in the presence of dual structural breaks which showed an evidence of a long-run equilibrium relationship. Their study found no evidence of causality running GDP per capita and natural gas consumption. Salahuddin et al. (2015) examined the relationship between CO_2 emissions, electricity consumption, economic growth and foreign direct investment with a data spanning from 1980-2012 in the Gulf

Peer-reviewed version available at Environmental Science and Pollution Research 2017, 24, , 6622-6633; doi:10.1007/s11356-016-8321-6

Cooperation Council by employing DOLS, FMOLS and DFE methods which showed an evidence of a long-run equilibrium relationship. Their study found evidence of unidirectional causality running from electricity consumption to CO_2 emissions and a bidirectional causality between economic growth and CO_2 emissions. Tiwari et al. (2013) examined the relationship between coal consumption, capital, trade openness, economic growth and CO_2 emissions with a data spanning from 1966-2011 in India by employing the ARDL bounds testing method in the presence of structural breaks which showed an evidence of a long-run equilibrium relationship. Their study confirmed the validity of the environmental Kuznets curve hypothesis and found an evidence of bidirectional causality between economic growth and CO_2 emissions, and between coal consumption and to CO_2 emissions. Shahbaz et al. (2012) examined the relationship between CO_2 emissions, energy consumption, trade openness and economic growth in Pakistan with a data spanning from 1972-2009 by employing the ARDL bounds testing method which showed an evidence of a long-run equilibrium relationship. Their study confirmed the validity of the environmental Kuznets curve hypothesis and found an evidence of unidirectional causality running from economic growth to CO_2 emissions.

Spurious regression occurs as a result of multicollinearity problems among study variables in a model. As a result, informative variables that are essential in explaining a specific response are dropped. Multicollinearity problems have been reported to exist when analysing variables that include environmental pollution, energy consumption and economic growth (Asumadu-Sarkodie and Owusu 2016a; Asumadu-Sarkodie and Owusu 2016e). As a result of multicollinearity problems, the inability for some econometric models (VAR, VECM and ARDL) to estimate coefficients per equation due to insufficient observations and the inability to estimate large number of candidate models due to the large number of regressors or the maximum number of lags used in a model, many studies employ few

Peer-reviewed version available at *Environmental Science and Pollution Research* **2017**, *24*, , 6622-6633; doi:10.1007/s11356-016-8321-6

predictor variables which may not be informative in explaining a response variable(s). Contrary to the aforementioned econometric models, the partial least squares regression is able to estimate the relationship between variables that exhibit strong multicollinearity, have the maximum number of lags and have more or equal number of variable observations than the predictor variables. The partial least squares regression analysis has been employed in previous studies outside the scope of the study (Ceglar et al. 2016; Mehmood et al. 2012; Xu et al. 2012), and has been labelled as a flexible method for multivariate analysis. Against the backdrop, the study explores the worth of estimating the impact of energy, agriculture, macroeconomic and human-induced indicators on environmental pollution in Ghana with 21 variables using the statistically inspired modification of partial least squares (SIMPLS). The study contributes to the global debate on climate change through the use of a versatile methodology and more informative variables that are essential in explaining environmental pollution in Ghana.

The remainder of the study comprises of: “Methodology”, “Results and Discussion”, and “Conclusion and Policy recommendations”.

METHODOLOGY

Data

The study examines the impact of energy, agriculture, macroeconomic and human-induced indicators on environmental pollution using the SIMPLS regression model. A time series data spanning from 1971 to 2011 were employed from World Bank (2014) and FAO (2015) as shown in Table 1. Table 2 presents the descriptive statistical analysis of the study variables. Evidence from Table 2 shows that with the exception of EPH, all the remaining variables are positively skewed. In addition, EN2OMA, ECH4MM, EPH, FOSI, GDPP, IND, MNMA,

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 26 September 2016 doi:10.20944/preprints201609.0095.v1

Peer-reviewed version available at Environmental Science and Pollution Research 2017, 24, , 6622-6633; doi:10.1007/s11356-016-8321-6

MNMM, RES and STOCK exhibit a leptokurtic distribution while the remaining variables exhibit a platykurtic distribution. From the Jarque-Bera test statistic, EN2OMA, ECH4MM, EPH, FOSI, GDPP, IND, MNMA, MNMM, RES and STOCK do not fit the normal distribution while the remaining variables are normally distributed based on 5% significance level. Figure shows the trend of the predictor variables versus the response variable. With the exception of EPH, it appears from Figure 1 that all the variables have a positive monotonic relationship with CO2.

Table 1. Data and Variable Definition

Abbreviation	**Variable Name**	**Unit**	**Source**
CO2	Carbon dioxide emissions	kt	World Bank (2014)
IND	Industry value added “as a proxy for industrialization”	current LCU	World Bank (2014)
POP	Population	NA	World Bank (2014)
RES	Total reserves (includes gold)	current US$	World Bank (2014)
TRD	Trade	% of GDP	World Bank (2014)
GDPP	GDP per capita	current LCU	World Bank (2014)
FOSI	Fossil fuel energy consumption	% of total	World Bank (2014)
LIVE	Livestock production index (2004-2006 = 100)	%	World Bank (2014)
EPH	Electricity production from hydroelectric sources	%	World Bank (2014)
CROPI	Crop production index (2004-2006 = 100)	%	World Bank (2014)
BBCRDM	Biomass burned crop residues	tonnes	FAO (2015)
ECH4En	Enteric Emissions	Gg	FAO (2015)
ECH4MM	Methane Emissions from Manure Management	Gg	FAO (2015)
EMCH4BCR	Methane Emissions from Burning crop residues	Gg	FAO (2015)
EMN2OBCR	Nitrous Oxide Emissions from Burning crop residues	Gg	FAO (2015)
EMN2OCR	Nitrous Oxide Emissions from Crop residues	Gg	FAO (2015)
EN2OMA	Nitrous Oxide Emissions from applied Manure	Gg	FAO (2015)
MNMA	Nitrogen content of Manure	kg	FAO (2015)
MNMM	Nitrogen content of Manure Management	kg	FAO (2015)
RCR	Crop Residue	tonnes	FAO (2015)
Stock	Stock of livestock	Head	FAO (2015)

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 26 September 2016 doi:10.20944/preprints201609.0095.v1

Peer-reviewed version available at Environmental Science and Pollution Research 2017, 24, , 6622-6633; doi:10.1007/s11356-016-8321-6

Table 2. Descriptive Statistical Analysis

Variable/Statistic	Mean	Median	Maximum	Minimum	Std. Dev.	Skewness	Kurtosis	Jarque-Bera	Probability†
BBCRDM	654173.9	665195	1135500	276160	220611.4	0.2879	2.3291	1.3355	0.5129
CO2	5035.238	4044.701	10102.58	2295.542	2459.736	0.6855	2.1929	4.3238	0.1151
CROPI	59.96049	52.8900	131.9200	25.1	32.10517	0.7125	2.2322	4.4757	0.1067
ECH4EN	63.7949	61.2759	96.8172	44.3663	15.28686	0.5197	2.1346	3.1247	0.2096
ECH4MM	2.960583	2.7357	5.0481	2.0038	0.802109	1.0549	3.1498	7.6433	0.0219
EMCH4BCR	1.76628	1.7960	3.0659	0.7457	0.595659	0.2879	2.3290	1.3357	0.5128
EMN2OBCR	0.045812	0.0466	0.0795	0.0193	0.015454	0.2882	2.3270	1.3413	0.5114
EMN2OCR	0.345305	0.3310	0.6326	0.1631	0.126716	0.4378	2.1803	2.4576	0.2926
EN2OMA	0.207612	0.1995	0.3506	0.1393	0.048843	1.4745	4.5270	18.8392	0.0001
EPH	91.02664	98.9276	100.0000	53.41072	13.10556	-1.2724	3.3097	11.2275	0.0036
FOSI	23.17066	21.99753	40.79426	11.5289	6.562491	0.9526	3.4830	6.5994	0.0369
GDPP	275.9624	16.13755	2399.515	0.028321	563.6219	2.4090	7.9861	82.1265	0.0000
IND	1.33E+09	41181200	1.43E+10	45700	2.87E+09	2.9282	11.9852	196.5128	0.0000
LIVE	79.1932	80.5400	127.5100	42.5700	22.3015	0.2495	2.5898	0.7129	0.7002
MNMA	9271656	8908200	15652400	6224200	2180329	1.4768	4.5301	18.9017	0.0001
MNMM	9963619	9621180	16636400	6715140	2295278	1.4411	4.4523	17.7943	0.0001
POP	15586682	15042736	24928503	8827273	4847325	0.3220	1.8888	2.8176	0.2444
RCR	17938625	17193600	32859500	8472450	6582574	0.4376	2.1798	2.4580	0.2926
RES	9.54E+08	4.37E+08	5.91E+09	43092215	1.30E+09	2.5046	8.9315	102.9695	0.0000
STOCK	24337141	18109700	63682000	11484700	13821325	1.3997	3.8776	14.7026	0.0006
TRD	54.05703	45.84812	116.0484	6.320343	29.8343	0.3255	2.0776	2.1777	0.3366

†denotes the rejection of the null hypothesis of normal distribution

 doi:10.20944/preprints201609.0095.v1

Peer-reviewed version available at *Environmental Science and Pollution Research* **2017**, *24*, , 6622-6633; doi:10.1007/s11356-016-8321-6

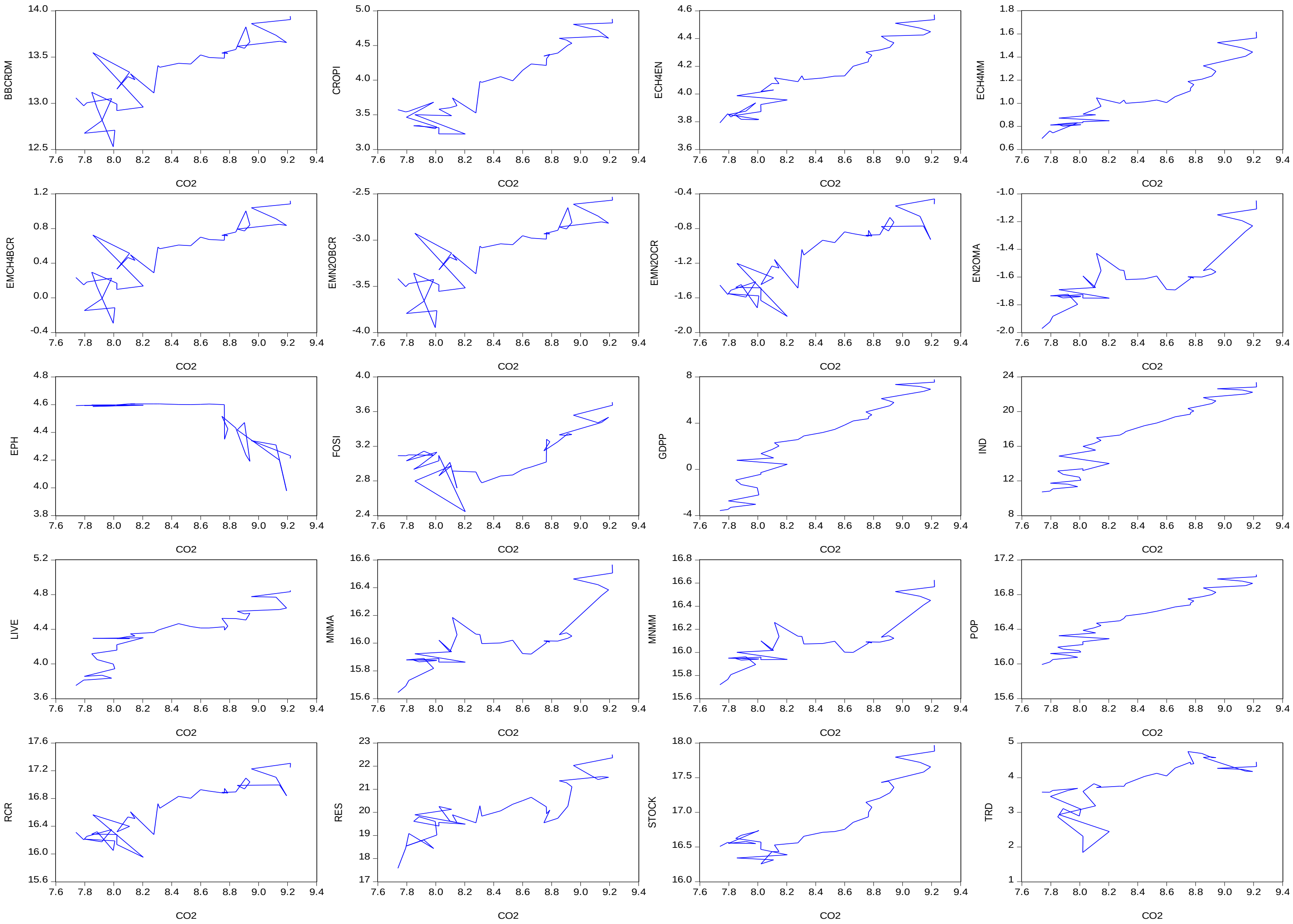

Figure 1. Trend of Predictor Variables versus Response Variable

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 26 September 2016 doi:10.20944/preprints201609.0095.v1

Peer-reviewed version available at Environmental Science and Pollution Research 2017, 24, , 6622-6633; doi:10.1007/s11356-016-8321-6

Partial Least Squares Regression

The partial least squares regression involves the use of principal component analysis to decide on a regression model with the latent variables and the response variables. The SIMPLS regression model directly computes the partial least squares factors as linear combination of the original study variables (De Jong 1993). The SIMPLS regression model was developed with the aim of explaining a specific optimality problem (i.e. maximizing covariance criterion between the predictor variables and the response variable, on condition that the predictor variable scores are orthogonal). As a result, it has been suggested that the SIMPLS regression model is to some extent superior over the non-linear iterative partial least squares (NIPALS) (Asumadu-Sarkodie and Owusu 2016c; Wise 2004).

Model Estimation

Prior to the SIMPLS regression model, a linear regression analysis is estimated to examine the relationship between energy, agriculture, macroeconomic and human-induced indicators and environmental pollution, expressed as:

$$CO2_t = \beta_0 + \beta_1 IND_t + \beta_2 POP_t + \beta_3 RES_t + \beta_4 TRD_t + \beta_5 GDPP_t + \beta_6 FOSI_t + \beta_7 LIVE_t + \beta_8 EPH_t + \beta_9 CROPI_t + \beta_{10} BBCRDM_t + \beta_{11} ECH4En_t + \beta_{12} ECH4MM_t + \beta_{13} EMCH4BCR_t + \beta_{14} EMN2OBCR_t + \beta_{15} EMN2OCR_t + \beta_{16} EN2OMA_t + \beta_{17} MNMA_t + \beta_{18} MNMM_t + \beta_{19} RCR_t + \beta_{20} Stock_t + \varepsilon_t \quad (1)$$

Where, β_0 represents the intercept, $\beta_1, \beta_2, \beta_3, \ldots, \beta_{20}$ represent the "projected change in the mean response of the predictor variables for each unit change in their value" and ε_t represents the white noise.

For brevity, the SIMPLS regression model is expressed as:

$$t = X_0 w \quad (2)$$

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 26 September 2016 doi:10.20944/preprints201609.0095.v1

Peer-reviewed version available at *Environmental Science and Pollution Research* **2017**, *24*, , 6622-6633; doi:10.1007/s11356-016-8321-6

Where, t is the score vector and w is its corresponding weight vector.

$$Let,\ X = X_0\ and\ Y = Y_0 \tag{3}$$

Where, $X = X_0$ is the centred and scaled matrix of the predictor variables and $Y = Y_0$ is the centred and scaled matrix of the response variable. X_0 and Y_0 are predicted by the partial least squares method via a regression on t:

$$\hat{X}_0 = tp', \text{ where } p' = (t't)^{-1}t'X_0 \tag{4}$$

$$\hat{Y}_0 = tc', \text{ where } c' = (t't)^{-1}t'Y_0 \tag{5}$$

Where, vectors p and c are the X-loading and Y-loading.

$t = X_0w$ in equation 2 is the "specific linear combination" with a maximum covariance, $t = t'u$ and a response linear combination $u = Y_0q$ which is characterized by the X-weight and Y-weight (w and q), that are proportional to the first left and right singular vectors of the covariance matrix $X_0'Y_0$. This cross-product matrix, $X_0'Y_0$ is deflated repeatedly for many latent variables/factors if is required. A detailed mathematical representation of the SIMPLS regression model can be found in the work of Boulesteix and Strimmer (2007); De Jong (1993); Wise (2004).

RESULTS AND DISCUSSION

Multicollinearity Examination

The study employs the linear regression analysis to examine the multicollinearity among variables using the variance inflation factor (VIF). Table 3 shows evidence of a linear relationship between energy, agriculture, macroeconomic and human-induced indicators and carbon dioxide emissions. However, there is evidence of a high multicollinearity among variables since VIF > 10 thus, the SIMPLS regression analysis in this study is worthwhile.

Peer-reviewed version available at Environmental Science and Pollution Research 2017, 24, , 6622-6633; doi:10.1007/s11356-016-8321-6

Table 3. Linear Regression Analysis

Source	DF	Adj SS	Adj MS	F-Value	P-Value
Regression	20	2.39E+08	11952080	80.47	0.0000
Term	Coef	SE Coef	T-Value	P-Value	VIF
Constant	5035.2	60.2	83.66	0.0000	
IND	580	645	0.9	0.3800	112.16
POP	3331	1127	2.95	0.0080	342.33
RES	-403	355	-1.14	0.2700	33.86
TRD	-117	285	-0.41	0.6850	21.91
GDPP	-901	1206	-0.75	0.4640	391.41
FOSI	-201	307	-0.65	0.5210	25.42
LIVE	-191	731	-0.26	0.7960	144
EPH	-699	223	-3.14	0.0050	13.39
CROPI	-2188	1088	-2.01	0.0580	318.81
EMCH4BCR	445299	1099936	0.4	0.6900	3.26E+08
EMN2OBCR	-22938	24729	-0.93	0.3650	164696.5
BBCRDM	-423457	1100536	-0.38	0.7040	3.26E+08
EMN2OCR	-104931	183343	-0.57	0.5730	9052993
RCR	106232	183191	0.58	0.5680	9037986
Stock	1274	3958	0.32	0.7510	4219.98
ECH4En	-33760	23169	-1.46	0.1610	144574.6
EN2OMA	-26835	57210	-0.47	0.6440	881463.7
MNMA	-308234	179926	-1.71	0.1020	8718736
ECH4MM	44522	33239	1.34	0.1950	297546
MNMM	323125	169114	1.91	0.0700	7702406

Assessment of SIMPLS Regression

A leave-one-out cross validation and a randomization test known as Van der Voet T^2 is used in the SIMPLS method to ascertain whether a model with a specific number of latent variables are different from the selected optimal model by the root mean predicted residual error sum of squares (PRESS). Evidence from Table 4 shows that the minimum root mean PRESS is 0.2407 and the minimizing number of factors in the SIMPLS method is 5. After the selection of the number of optimal latent variables, the study examines the percentage variation explained in SIMPLS model. Evidence from Figure 2 shows that almost 98% of variation is explained for the predictor variables while 97% of variation is explained for the response variable.

doi:10.20944/preprints201609.0095.v1

Peer-reviewed version available at *Environmental Science and Pollution Research* **2017**, *24*, , 6622-6633; doi:10.1007/s11356-016-8321-6

Table 4. Cross Validation with SIMPLS

Number of factors	**Root Mean PRESS**	**plot**	**van der Voet T^2**	**Prob > van der Voet T^2**
0	1.025000		17.695569	<.0001*
1	0.329389		4.668915	0.0240*
2	0.273027		1.041678	0.3350
3	0.246546		0.113571	0.7580
4	0.240888		0.000168	0.9920
5	0.240697		0.000000	1.0000
6	0.252542		1.384288	0.2740
7	0.253406		1.011143	0.3160
8	0.283166		2.229541	0.0740
9	0.298964		1.937773	0.1670
10	0.305272		1.512556	0.3370
11	0.340052		1.555519	0.3390
12	0.508279		1.162422	0.3520
13	0.745206		1.049135	0.3780
14	1.017027		0.998313	0.5350
15	0.533327		0.944965	0.6160

Note: The minimum root mean PRESS is 0.2407 and the minimizing number of factors is 5.

Peer-reviewed version available at *Environmental Science and Pollution Research* **2017**, *24*, , 6622-6633; doi:10.1007/s11356-016-8321-6

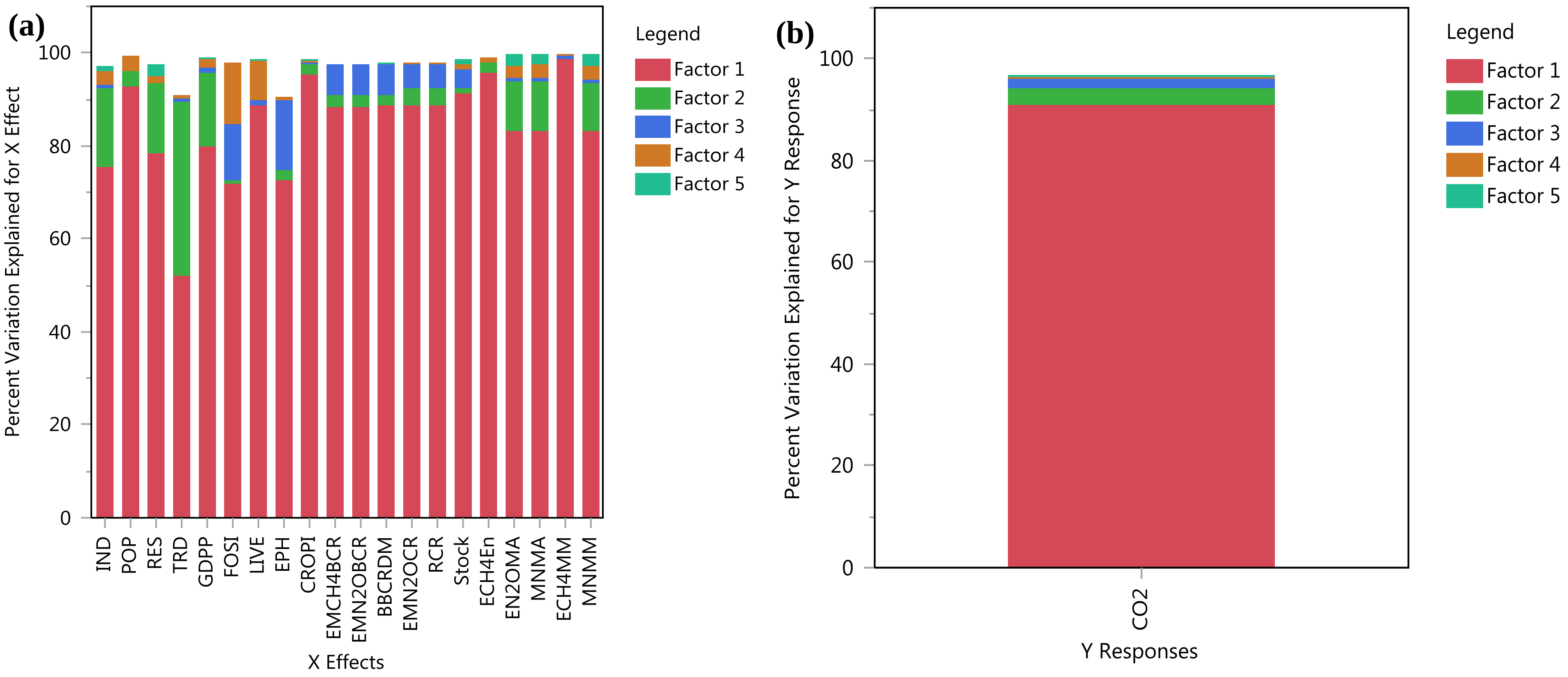

Figure 2. (**a**) Percent Variation Explained for X Effects and (**b**) Percent Variation Explained for Y Responses

Peer-reviewed version available at *Environmental Science and Pollution Research* **2017**, *24*, , 6622-6633; doi:10.1007/s11356-016-8321-6

The Variable Importance Analysis

The importance of the predictor variables in explaining carbon dioxide emissions is estimated using the variable importance of projection (VIP), which examines the contribution of each predictor variable based on the variance explained by each SIMPLS latent variable (Mehmood et al. 2012; Wold et al. 2001). The first analysis in the SIMPLS model led to the deletion of some variables that had VIP value less than 0.83, since literature considers them as irrelevant (Gosselin et al. 2010). Eriksson et al. (2001) categorizes the explanatory variables based on their VIP thus, "highly influential" ($VIP > 1$), "moderately influential" ($0.8 < VIP < 1$) and "less influential" ($VIP\ less\ than\ 0.8$). In the examination of the variable importance analysis in Table 5, the explanatory variables can be categorized based on their VIP value. IND, RES, TRD, GDPP, FOSI, EN2OMA, MNMA and MNMM can be classified as moderately influential variables while POP, LIVE, EPH, CROPI, EMCH4BCR, EMN2OBCR, BBCRDM, EMN2OCR, RCR, Stock, ECH4En and ECH4MM can be classified as highly influential variables. Figure 3 presents a plot of VIP versus coefficients for centred and scaled data. The black dotted lines represent the "threshold line" while the red-solid-circles represent the level of carbon dioxide emissions. Values of the coefficient below zero represent a negative contribution towards carbon dioxide emissions while values greater than zero represent a positive contribution towards carbon dioxide emissions. It appears from Figure 3 that, the higher the VIP, the higher the coefficient contribution and the lower the VIP, the lower the coefficient contribution.

Peer-reviewed version available at Environmental Science and Pollution Research 2017, 24, , 6622-6633; doi:10.1007/s11356-016-8321-6

Table 5. Variable Importance Analysis

X	VIP
IND	0.8905
POP	1.1438
RES	0.9083
TRD	0.9417
GDPP	0.9148
FOSI	0.9228
LIVE	1.0224
EPH	1.0765
CROPI	1.0728
EMCH4BCR	1.0008
EMN2OBCR	1.0028
BBCRDM	1.0012
EMN2OCR	1.0089
RCR	1.0099
Stock	1.0299
ECH4En	1.0842
EN2OMA	0.9331
MNMA	0.9356
ECH4MM	1.0927
MNMM	0.9591

Peer-reviewed version available at Environmental Science and Pollution Research 2017, 24, , 6622-6633; doi:10.1007/s11356-016-8321-6

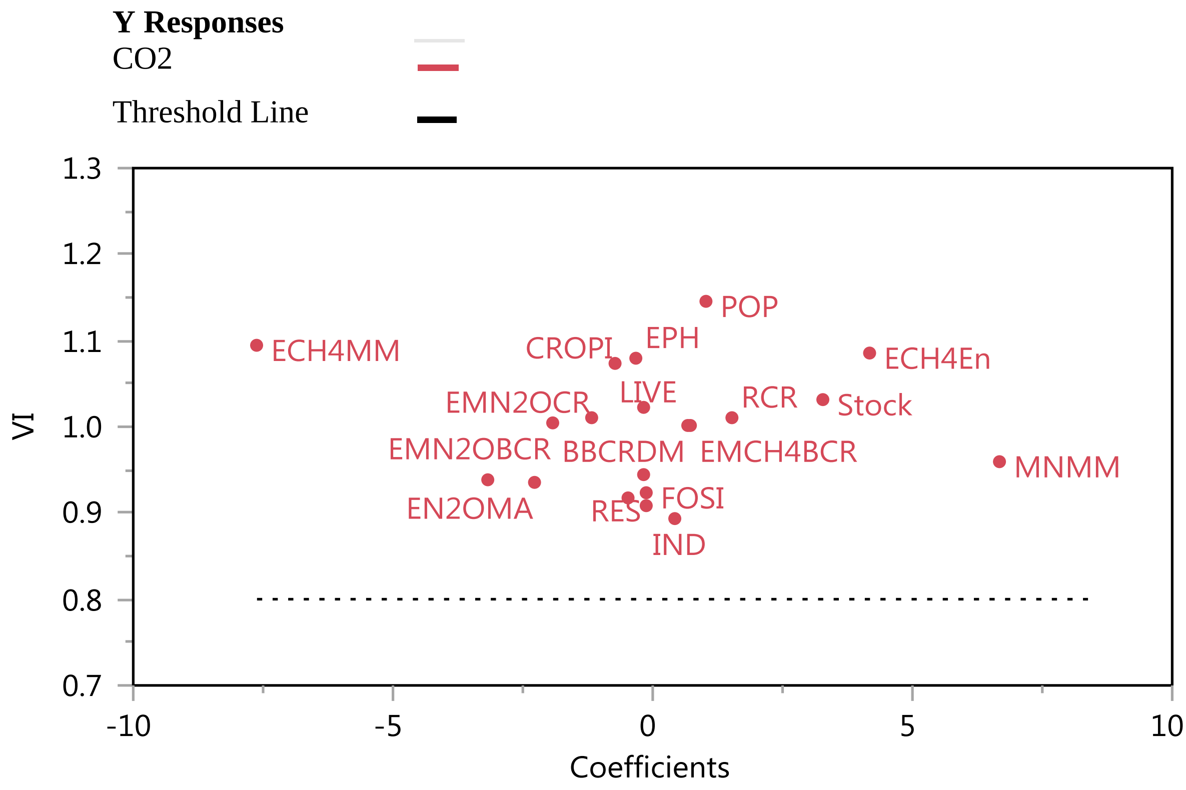

Figure 3. VIP versus Coefficients for Centred and Scaled Data

Impact Analysis

The impact analysis of energy, agriculture, macroeconomic and human-induced indicators on environmental pollution is evident in Table 6. Evidence from Table 6 shows that, a 1% increase in Methane Emissions from Manure Management will reduce carbon dioxide emissions by 7.62%. According to EPA (2016), the emissions from methane can be reduced and captured if manure management strategies are altered through animal feeding practices or at the livestock operations. Nitrogen content of Manure increased by 1% will reduce carbon dioxide emissions by 3.13%. An increase in Nitrous Oxide Emissions from applied Manure by 1% will reduce carbon dioxide emissions by 2.27%. Nitrous Oxide Emissions from Burning crop residues increased by 1% will reduce carbon dioxide emissions by 1.91%. Nitrous Oxide Emissions from Crop residues increased by 1% will reduce carbon dioxide emissions by 1.13%. Crop production index increased by 1% will reduce carbon dioxide

Peer-reviewed version available at *Environmental Science and Pollution Research* **2017**, *24*, , 6622-6633; doi:10.1007/s11356-016-8321-6

emissions by 0.71%. GDP per capita increased by 1% will reduce carbon dioxide emissions by 0.46%, thus supports the environmental Kuznets curve hypothesis that an increase in a country's economic growth leads to a reduction in environmental pollution. Electricity production from hydroelectric sources increased by 1% will reduce carbon dioxide emissions by 0.30%, thus increasing renewable energy sources in Ghana's energy portfolio will help mitigate carbon dioxide emissions. Livestock production index increased by 1% will reduce carbon dioxide emissions by 0.14%. Increasing Trade by 1% will reduce carbon dioxide emissions by 0.13%. Fossil fuel energy consumption increased by 1% will reduce carbon dioxide emissions by 0.12%. Increasing the total reserves including gold by 1% will reduce carbon dioxide emissions by 0.10%, thus halting the illegal mining and chain saw operations will help reduce carbon dioxide emissions.

In contrast, a 1% increase in industrialization will increase carbon dioxide emissions by 0.44%. A 1% increase in Biomass burned crop residues will increase carbon dioxide emissions by 0.69%. Methane Emissions from Burning crop residues increased by 1% will increase carbon dioxide emissions by 0.74%. population increased 1% will increase carbon dioxide emissions by 1.03%. A 1% increase in Crop Residue will increase carbon dioxide emissions by 1.56%. A stock of animals increased 1% will increase carbon dioxide emissions by 3.29%. Increasing Enteric Emissions by 1% will increase carbon dioxide emissions by 4.22% and a 1% increase in the Nitrogen content of Manure Management will increase carbon dioxide emissions by 6.69%. However, the value of the intercept indicates that when there is no increase in the energy, agriculture, macroeconomic and human-induced indicators, there will be no carbon dioxide emissions.

Table 6. Model Coefficients for Centred and Scaled Data

Coefficient	CO2
Intercept	0.0000
IND	0.4366
POP	1.0303
RES	-0.1047
TRD	-0.1301
GDPP	-0.4626
FOSI	-0.1176
LIVE	-0.1439
EPH	-0.2976
CROPI	-0.7094
EMCH4BCR	0.7443
EMN2OBCR	-1.9058
BBCRDM	0.6877
EMN2OCR	-1.1297
RCR	1.5621
Stock	3.2883
ECH4En	4.2232
EN2OMA	-2.2673
MNMA	-3.1275
ECH4MM	-7.6173
MNMM	6.6878

Diagnostics of the SIMPLS Method

The SIMPLS regression method is subjected to diagnostic checks to estimate the independence of the residuals in the model. Evidence from the diagnostic plot s shows that the actual response and the predicted variable nearly fit on the regression line along with evidence of normal distribution by the residual normal quantile plot.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 26 September 2016 doi:10.20944/preprints201609.0095.v1

Peer-reviewed version available at *Environmental Science and Pollution Research* **2017**, *24*, , 6622-6633; doi:10.1007/s11356-016-8321-6

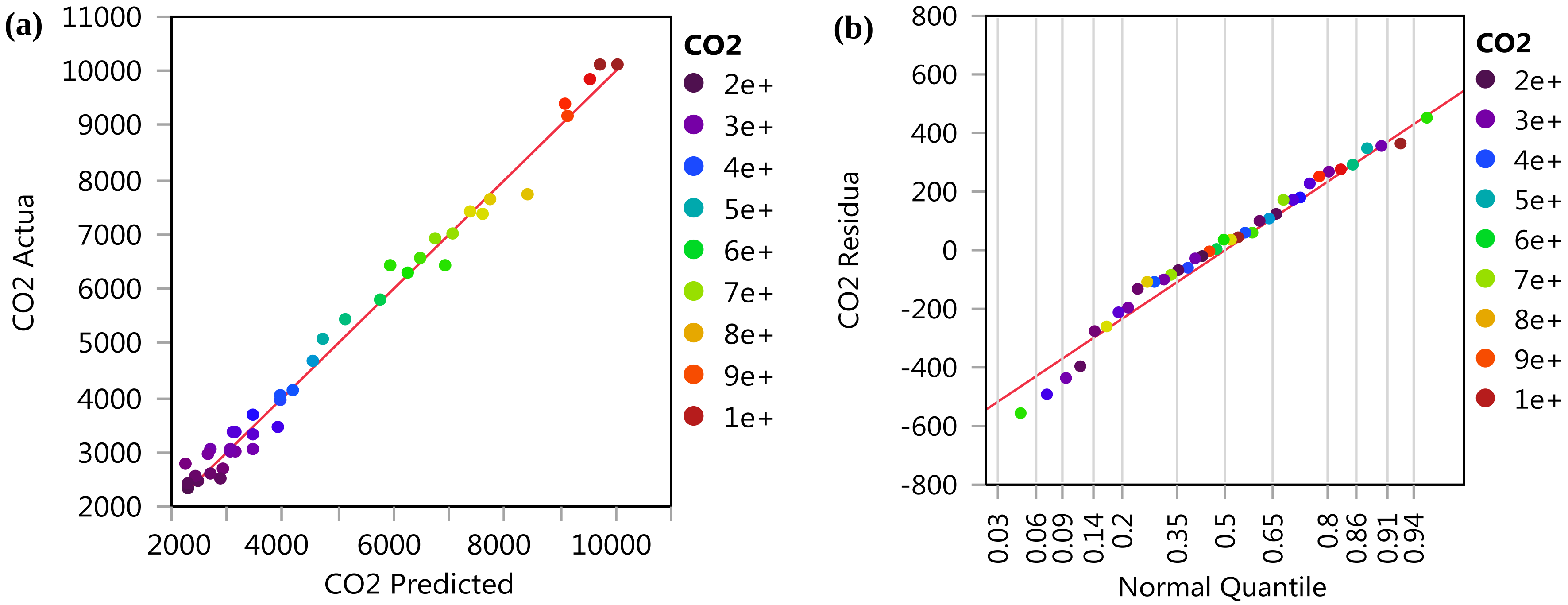

Figure 4. Diagnostics Plots (a) Actual by Predicted Plot (b) Residual Normal Quantile Plot

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 26 September 2016 doi:10.20944/preprints201609.0095.v1

Peer-reviewed version available at Environmental Science and Pollution Research 2017, 24, , 6622-6633; doi:10.1007/s11356-016-8321-6

SIMPLS Forecasting and Performance

Figure 5 shows the Actual versus SIMPLS Predicted CO2. The study explores the usefulness of the SIMPLS regression method in predicting the response variable. Evidence from Figure 5 shows that the SIMPLS predicted carbon dioxide emissions nearly fits the original data with a mean absolute percentage error (MAPE) of 5.16% which is acceptable. By the addition of a trend line, the following equation is generated based on sixth-polynomial: $y = 3E\text{-}05x^6 - 0.003x^5 + 0.1045x^4 - 1.0881x^3 - 2.7697x^2 + 127.4x + 2200.3$ with an $R^2 = 0.9742$. 97% of the SIMPLS predicted carbon dioxide emissions were explained by the sixth-polynomial equation.

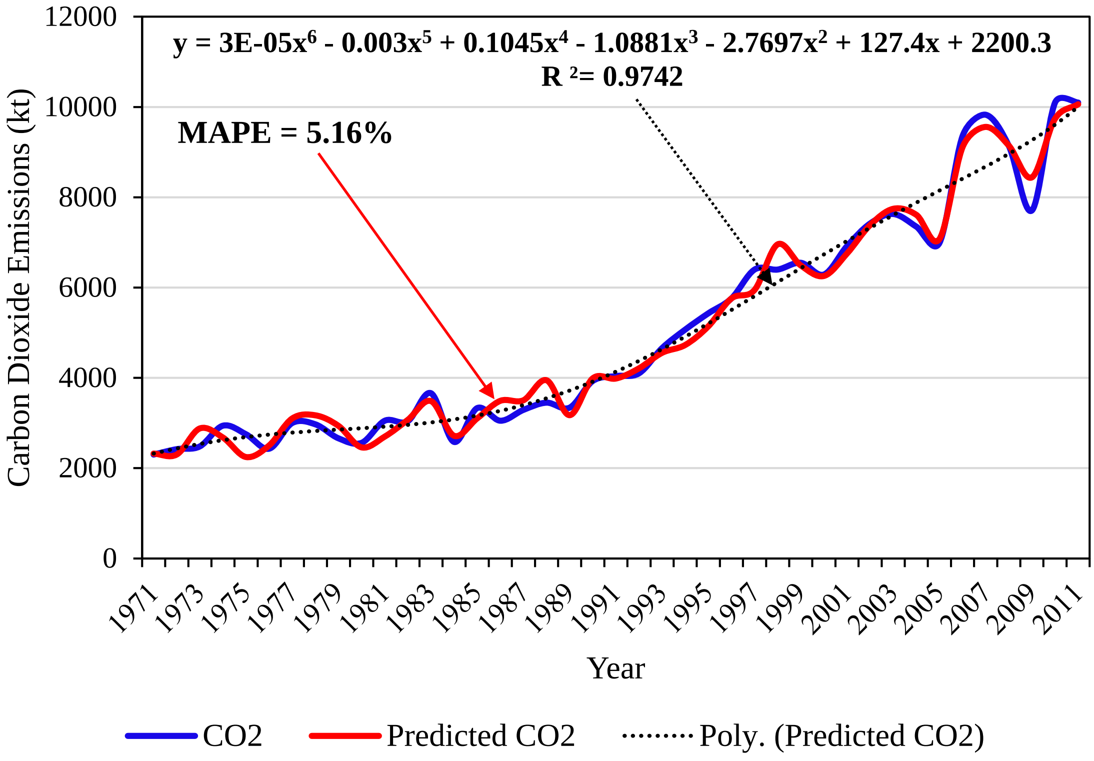

Figure 5. Actual versus SIMPLS Predicted CO2

Peer-reviewed version available at *Environmental Science and Pollution Research* **2017**, *24*, , 6622-6633; doi:10.1007/s11356-016-8321-6

CONCLUSION AND POLICY RECOMMENDATION

The study examines the impact of energy, agriculture, macroeconomic and human-induced indicators on environmental pollution from 1971 to 2011 using the SIMPLS regression model. There was evidence of a linear relationship between energy, agriculture, macroeconomic and human-induced indicators and carbon dioxide emissions.

Evidence from the SIMPLS regression model shows that, Crop production index increased by 1% will reduce carbon dioxide emissions by 0.71%. GDP per capita increased by 1% will reduce carbon dioxide emissions by 0.46%, thus supports the environmental Kuznets curve hypothesis that an increase in a country's economic growth leads to a reduction in environmental pollution. Electricity production from hydroelectric sources increased by 1% will reduce carbon dioxide emissions by 0.30%, thus increasing renewable energy sources in Ghana's energy portfolio will help mitigate carbon dioxide emissions.

In addition, a 1% increase in industrialization will increase carbon dioxide emissions by 0.44%. A 1% increase in Biomass burned crop residues will increase carbon dioxide emissions by 0.69%. Methane Emissions from Burning crop residues increased by 1% will increase carbon dioxide emissions by 0.74%. population increased 1% will increase carbon dioxide emissions by 1.03%. A 1% increase in Crop Residue will increase carbon dioxide emissions by 1.56%. A stock of animals increased 1% will increase carbon dioxide emissions by 3.29%. Increasing Enteric Emissions by 1% will increase carbon dioxide emissions by 4.22% and a 1% increase in the Nitrogen content of Manure Management will increase carbon dioxide emissions by 6.69%. However, the value of the intercept indicates that when there is no increase in the energy, agriculture, macroeconomic and human-induced indicators, there will be no carbon dioxide emissions.

Evidence from the SIMPLS regression forecasting shows that the predicted carbon dioxide emissions fit the original data with a performance of 5% MAPE and a 97% R squared value

Peer-reviewed version available at Environmental Science and Pollution Research 2017, 24, , 6622-6633; doi:10.1007/s11356-016-8321-6

from the performance of a trend line data analysis based on sixth-polynomial equation. The following policy recommendations are proposed in the study:

Integrating climate change measures into national energy policies, sustainable agricultural policies, strategies and planning will increase institutional capacities towards adopting climate change measures, adaptation, early warnings and impact reduction.

Industrial policies that aims at promoting local technological development, scientific research, innovation and creativity, access to affordable internet and the provision of conducive environmental policies will propel Ghana's effort towards achieving sustainable industrialization.

Finally, there is the need for an enhanced local financial institutional capacity to boost access to financial services, insurance and banking in order to attract foreign investments, create jobs and propel the country's financial development.

Conflicts of Interest: The authors declare no conflict of interest.

REFERENCES

Acaravci A, Ozturk I (2010) On the relationship between energy consumption, CO 2 emissions and economic growth in Europe Energy 35:5412-5420

Apergis N, Payne JE (2011) The renewable energy consumption–growth nexus in Central America Applied Energy 88:343-347 doi:http://dx.doi.org/10.1016/j.apenergy.2010.07.013

Asumadu-Sarkodie S, Owusu PA (2016a) Carbon dioxide emissions, GDP, energy use and population growth: a multivariate and causality analysis for Ghana, 1971-2013 Environmental Science and Pollution Research International 23:13508–13520 doi:http://dx.doi.org/10.1007/s11356-016-6511-x

Asumadu-Sarkodie S, Owusu PA (2016b) Feasibility of biomass heating system in Middle East Technical University, Northern Cyprus Campus Cogent Engineering 3:1134304 doi:http://dx.doi.org/10.1080/23311916.2015.1134304

Asumadu-Sarkodie S, Owusu PA (2016c) A Multivariate Analysis of Carbon Dioxide Emissions, Electricity Consumption, Economic Growth, Financial Development, Industrialization and Urbanization in Senegal Energy Sources, Part B: Economics, Planning, and Policy doi:http://dx.doi.org/10.1080/15567249.2016.1227886

Peer-reviewed version available at *Environmental Science and Pollution Research* **2017**, *24*, , 6622-6633; doi:10.1007/s11356-016-8321-6

Asumadu-Sarkodie S, Owusu PA (2016d) Multivariate co-integration analysis of the Kaya factors in Ghana Environmental Science and Pollution Research International 23:9934-9943 doi:http://dx.doi.org/10.1007/s11356-016-6245-9
Asumadu-Sarkodie S, Owusu PA (2016e) Recent Evidence of the Relationship between Carbon Dioxide Emissions, Energy use, GDP and Population in Ghana: A Linear Regression Approach Energy Sources, Part B: Economics, Planning, and Policy doi:https://dx.doi.org/10.1080/15567249.2016.1208304
Ben Abdallah K, Belloumi M, De Wolf D (2013) Indicators for sustainable energy development: A multivariate cointegration and causality analysis from Tunisian road transport sector Renewable and Sustainable Energy Reviews 25:34-43 doi:http://dx.doi.org/10.1016/j.rser.2013.03.066
Boulesteix A-L, Strimmer K (2007) Partial least squares: a versatile tool for the analysis of high-dimensional genomic data Briefings in bioinformatics 8:32-44
Ceglar A, Toreti A, Lecerf R, Van der Velde M, Dentener F (2016) Impact of meteorological drivers on regional inter-annual crop yield variability in France Agricultural and Forest Meteorology 216:58-67
Cerdeira Bento JP, Moutinho V (2016) CO2 emissions, non-renewable and renewable electricity production, economic growth, and international trade in Italy Renewable and Sustainable Energy Reviews 55:142-155 doi:http://dx.doi.org/10.1016/j.rser.2015.10.151
De Jong S (1993) SIMPLS: an alternative approach to partial least squares regression Chemometrics and Intelligent Laboratory Systems 18:251-263
Engle RF, Granger CW (1987) Co-integration and error correction: representation, estimation, and testing Econometrica: Journal of the Econometric Society:251-276
EPA (2016) Methane Emissions. https://www3.epa.gov/climatechange/ghgemissions/gases/ch4.html. Accessed July 7th, 2016
Eriksson L, Johansson E, Kettaneh-Wold N, Wold S (2001) Multi-and megavariate data analysis: principles and applications. Umetrics,
FAO (2015) FAO Statistical Yearbooks - World food and agriculture. http://faostat3.fao.org/home/E. Accessed October 24, 2015
Gosselin R, Rodrigue D, Duchesne C (2010) A Bootstrap-VIP approach for selecting wavelength intervals in spectral imaging applications Chemometrics and Intelligent Laboratory Systems 100:12-21
Gul S, Zou X, Hassan CH, Azam M, Zaman K (2015) Causal nexus between energy consumption and carbon dioxide emission for Malaysia using maximum entropy bootstrap approach Environmental Science and Pollution Research:1-13
Huang B-N, Hwang MJ, Yang CW (2008) Causal relationship between energy consumption and GDP growth revisited: A dynamic panel data approach Ecological Economics 67:41-54 doi:http://dx.doi.org/10.1016/j.ecolecon.2007.11.006
Jammazi R, Aloui C (2015) On the interplay between energy consumption, economic growth and CO2 emission nexus in the GCC countries: A comparative analysis through wavelet approaches Renewable and Sustainable Energy Reviews 51:1737-1751 doi:10.1016/j.rser.2015.07.073
Johansen S (1995) Likelihood-based inference in cointegrated vector autoregressive models OUP Catalogue
Lozano S, Gutiérrez E (2008) Non-parametric frontier approach to modelling the relationships among population, GDP, energy consumption and CO2 emissions Ecological Economics 66:687-699 doi:http://dx.doi.org/10.1016/j.ecolecon.2007.11.003

Peer-reviewed version available at Environmental Science and Pollution Research 2017, 24, , 6622-6633; doi:10.1007/s11356-016-8321-6

Mehmood T, Liland KH, Snipen L, Sæbø S (2012) A review of variable selection methods in partial least squares regression Chemometrics and Intelligent Laboratory Systems 118:62-69

Ouyang X, Lin B (2015) An analysis of the driving forces of energy-related carbon dioxide emissions in China's industrial sector Renewable and Sustainable Energy Reviews 45:838-849 doi:http://dx.doi.org/10.1016/j.rser.2015.02.030

Owusu P, Asumadu-Sarkodie S (2016) A Review of Renewable Energy Sources, Sustainability Issues and Climate Change Mitigation Cogent Engineering 3:1167990 doi:http://dx.doi.org/10.1080/23311916.2016.1167990

Ozturk I, Acaravci A (2010) The causal relationship between energy consumption and GDP in Albania, Bulgaria, Hungary and Romania: Evidence from ARDL bound testing approach Applied Energy 87:1938-1943

Ozturk I, Acaravci A (2011) Electricity consumption and real GDP causality nexus: Evidence from ARDL bounds testing approach for 11 MENA countries Applied Energy 88:2885-2892

Qureshi MI, Rasli AM, Zaman K (2016) Energy crisis, greenhouse gas emissions and sectoral growth reforms: repairing the fabricated mosaic Journal of Cleaner Production 112:3657-3666 doi:10.1016/j.jclepro.2015.08.017

Rafindadi AA, Ozturk I (2015) Natural gas consumption and economic growth nexus: Is the 10th Malaysian plan attainable within the limits of its resource? Renewable and Sustainable Energy Reviews 49:1221-1232 doi:http://dx.doi.org/10.1016/j.rser.2015.05.007

Remuzgo L, Sarabia JM (2015) International inequality in CO2 emissions: A new factorial decomposition based on Kaya factors Environmental Science & Policy 54:15-24 doi:10.1016/j.envsci.2015.05.020

Salahuddin M, Gow J, Ozturk I (2015) Is the long-run relationship between economic growth, electricity consumption, carbon dioxide emissions and financial development in Gulf Cooperation Council Countries robust? Renewable and Sustainable Energy Reviews 51:317-326 doi:http://dx.doi.org/10.1016/j.rser.2015.06.005

Seker F, Ertugrul HM, Cetin M (2015) The impact of foreign direct investment on environmental quality: A bounds testing and causality analysis for Turkey Renewable and Sustainable Energy Reviews 52:347-356 doi:http://dx.doi.org/10.1016/j.rser.2015.07.118

Shahbaz M, Khraief N, Jemaa MMB (2015) On the causal nexus of road transport CO2 emissions and macroeconomic variables in Tunisia: Evidence from combined cointegration tests Renewable and Sustainable Energy Reviews 51:89-100 doi:http://dx.doi.org/10.1016/j.rser.2015.06.014

Shahbaz M, Lean HH, Shabbir MS (2012) Environmental Kuznets Curve hypothesis in Pakistan: Cointegration and Granger causality Renewable and Sustainable Energy Reviews 16:2947-2953 doi:http://dx.doi.org/10.1016/j.rser.2012.02.015

Soytas U, Sari R (2009) Energy consumption, economic growth, and carbon emissions: Challenges faced by an EU candidate member Ecological Economics 68:1667-1675 doi:http://dx.doi.org/10.1016/j.ecolecon.2007.06.014

Tiwari AK, Shahbaz M, Adnan Hye QM (2013) The environmental Kuznets curve and the role of coal consumption in India: Cointegration and causality analysis in an open economy Renewable and Sustainable Energy Reviews 18:519-527 doi:http://dx.doi.org/10.1016/j.rser.2012.10.031

Wise BM (2004) Properties of Partial Least Squares (PLS) Regression, and Differences between Algorithms. Technical Report,

Peer-reviewed version available at *Environmental Science and Pollution Research* **2017**, *24*, , 6622-6633; doi:10.1007/s11356-016-8321-6

Wold S, Sjöström M, Eriksson L (2001) PLS-regression: a basic tool of chemometrics Chemometrics and Intelligent Laboratory Systems 58:109-130
World Bank (2014) World Development Indicators. http://data.worldbank.org/country. Accessed October 24, 2015
Xu C, Yue D, Deng C (2012) Hybrid GA/SIMPLS as alternative regression model in dam deformation analysis Engineering Applications of Artificial Intelligence 25:468-475
Zhang X-P, Cheng X-M (2009) Energy consumption, carbon emissions, and economic growth in China Ecological Economics 68:2706-2712 doi:10.1016/j.ecolecon.2009.05.011