A Statistical Model for Energy Intensity

Marjaneh Issapour
Farmingdale State College
Farmingdale N.Y. 11735, USA
&
Lori L. Scarlatos
Stony Brook University
Stony Brook, N.Y. 11790, USA
&
Herbert F. Lewis
Stony Brook University
Stony Brook , N.Y. 11790, USA

Abstract

A promising approach to improve scientific literacy in regards to global warming and climate change is using a simulation as part of a science education course. The simulation needs to employ scientific analysis of actual data from internationally accepted and reputable databases to demonstrate the reality of the current climate change situation. One of the most important criteria for using a simulation in a science education course is the fidelity of the model. The realism of the events and consequences modeled in the simulation is significant as well. Therefore, all underlying equations and algorithms used in the simulation must have real-world scientific basis. The "Energy Choices" simulation is one such simulation. The focus of this paper is the development of a mathematical model for "Energy Intensity" as a part of the overall system dynamics in "Energy Choices" simulation. This model will define the "Energy Intensity" as a function of other independent variables that can be manipulated by users of the simulation. The relationship discovered by this research will be applied to an algorithm in the "Energy Choices" simulation.

Keywords: (Educational games, Simulation, Energy choices, Linear Regression , Energy Intensity, Climate change, IPAT equation)

1. Introduction

The global rate of CO_2 emissions from fossil fuel combustion and industrial processes have been accelerating in the past decade. In reality, the emissions growth rate is greater than the most intensive fossil-fuel emissions scenario developed by the Intergovernmental Panel on Climate Change (IPCC) [1]. Global warming is a present and legitimate threat, but the situation is not hopeless. Scientific and technological research indicates that sustainable energy development strategies can provide a successful solution to global warming [2] & [3].

Scientists generally agree that the climate change is a complex socio-scientific issue in that human choices are a major factor in the relationship between global energy use and climate change. Humans are at the center of this socio-scientific challenge and need to understand their role as the decision makers within this system. The only known way to assist ordinary people in understanding the gravity of the situation is through education and scientific literacy [4].

One promising approach to scientific literacy is the use of simulation software as part of a science education course. Simulations often allow users to see how science applies to socio-scientific issues within their own lives [5] & [6]. Players can observe the predicted long term consequences of their decisions within a few minutes of using the simulation. The graphics used by simulations are a powerful tool to help people understand and internalize theoretical knowledge of a concept [7].The simulation needs to employ scientific analysis of actual data from internationally accepted and reputable databases in order to demonstrate the reality of the current situation.

"Energy Choices" is an example of such simulation [8]. This software is collaboratively designed by Dr. Scarlatos of Stony Brook University and Dr. Tomkiewicz of Brooklyn College. The "Energy Choices" simulation is designed to form one of the building blocks of an educational system, which addresses scientific literacy and quantitative skills. It also aims to assist educators in addressing the moral and ethical implications of climate change and global warming as a whole. This simulation teaches science while providing an opportunity for users to role-play, which may in turn trigger an attitude change. The software is designed for employment in a general science education course on energy and the environment [8].

The goal of the player in the "Energy Choices" simulation/game is to keep the population happy by keeping consumption high while minimizing environmental impact. The "Energy Intensity" factor determines how efficiently energy gets converted to gross domestic product (GDP), therefore minimizing "Energy Intensity" can help the player to achieve his/her goals.

The focus of this paper is the development of a mathematical model for "Energy Intensity" as a part of the overall system dynamics in the "Energy Choices" simulation. This model will define "Energy Intensity" as a function of other independent variables that can be manipulated by users of the simulation.

The relationship discovered by this research will be applied to an algorithm in the “Energy Choices” simulation.

2. Background and the Theoretical Foundation

There is growing concern among educational scholars about the scientific literacy of the general population. Scientific literacy enables individuals to make informed decisions in the socio-technological context of everyday life. Science and technology education alone do not address socio-technological and environmental issues of our future generations [9]. Understanding the latter area of science is necessary for dealing with the scientifically based personal and societal issues that confront our society. The current curriculums in science and technology education do not adequately deal with issues that relate to human values in decision making [10] & [4].

Scientific literacy requires more than learning the basic facts. Educational research on the learning process demonstrates that people learn best when they are actively engaged, and when learning is linked to “meaning making,” or seeing how the science being taught applies to their own lives [11] & [12]. The learner’s own life is where existing knowledge has been constructed, and in order for new knowledge to take root, it must connect with the everyday experiences of the learner [13], [6], [14]& [15].

The National Science Education Standards recommend that, “All students should develop an understanding of science and technology in local, national and global challenges” [16]& [17]. As a result of this, there has been a growing focus on the combined context of science, technology, and society, stressing the impact of science and technology decisions on society [18] & [19]. More recently, this idea has been extended to encompass a range of socio-scientific issues that allow students to consider the impact of science on a personal, as well as global, level. These issues include the nature of science, classroom discourse, cultural aspects, and moral issues raised on a case-by-case basis [20]. Within this context, the Nature of Science [NoS] has emerged as an interdisciplinary area of inquiry that draws input from both the technological and social sciences [4].

Global energy use, and its connections to climate change, is a prime example of a socio-scientific topic that requires examination and understanding of these issues within a single complex system. Although this system is dominated by humans, it must be subject to the same disciplined study that is applied to other physical systems. The study must be anchored by reproducible observations that give rise to theoretical understanding through testing. The challenge here is that the system is somewhat unique in that the investigators are actually part of the system that requires examination [21].The fact that humans now have a major influence on this interaction requires that moral and ethical implications, which have traditionally been difficult issues for scientists and educators, be part of such a system [22]& [23]In democratic societies, steps taken to ensure sustainable planetary equilibrium will be implemented through the political process. The way to translate science into electoral issues is through education.

A promising approach to improve scientific literacy is using a simulation as part of a science education course. The simulation needs to employ scientific analysis of actual data from internationally accepted and reputable databases to demonstrate the reality of the current climate change situation. This approach is the basis for the “Energy Choices” simulation. This simulation is designed to teach quantitative skills as well as to promote positive pro-environmental behavior for individuals as decision makers. This simulation offers a unified solution to deal with some of the challenges of the three research areas namely: 1) Science educations’ consideration for scientific literacy and Nature of Science, 2) Effects of role play in Simulation on individual’s behavior and attitude, and 3) Climate change as an example of a socio-scientific issue that allow students to consider the impact of science on a personal, as well as global, level.

One of the most important criteria for using a simulation in a science education course is the fidelity of the model. The realism of the events and consequences modeled in the simulation is significant as well. Therefore, all underlying equations and algorithms used in the simulation must have real-world scientific bases. This paper focuses on development of a statistical model for the Energy Intensity as a part of the overall system dynamics in the “Energy Choices” simulation.

3. Mathematical Model

The “Energy Choices” simulation uses the IPAT equation, a mathematical description of the following relationship (I stands for Impact, P for Population, A for Affluence and T for Technology):

*Impact = Population * Affluence * Technology*

For CO_2 emission, representing the impact on the planet, the identity takes the following form:

Equation 1

$$\frac{CO2}{year} = Population * \left(\frac{GDP}{population}\right) * \left(\frac{Energy}{GDP}\right) * \left(\frac{FossilFuels}{Energy}\right) * \left(\frac{CO2}{FossilFuels}\right)$$

Although this is a simplified model, it helps students visualize the factors influencing the climate change and causes of global warming [3]. All the quantities in the above equation, except for the population, are quantities per year. The Impact here is the environmental impact of CO_2 emission. The Affluence is measured by GDP/Capita (GDP/Population in the above equation). The rest of the terms refer to the Technology part of the acronym. *Energy/GDP* often referred to as Energy Intensity is a measure of how efficiently energy is used.

Player's goals in the "Energy Choices" simulation as well as the global challenge are to minimize CO_2emission while at the same time maximizing the GDP/capita.

The simulation offers each player the choice of the types of fuel that each country will use, which directly affects the CO_2 emission. Using an energy profile which includes a greater percentage of alternative and renewable sources will reduce the CO_2 emission, although it may increase the cost and therefore have a negative effect on GDP. Each player also gets to choose what percentage of the country's GDP is spent on investing in infrastructures which should indirectly lower the Energy Intensity and improve the efficient use of energy.

The purpose of this research is to investigate the factors that influence Energy Intensity (EI) and develop a statistical model for EI. This model will be used to come up with an algorithm to be used in the Energy Choices simulation.

Energy Intensity model development process

As stated earlier the Energy Intensity (EI) of a particular country is an indication of how efficiently a country uses energy to generate GDP. Therefore, one of the objectives of any country's leader is to reduce the EI. Theoretically this may be achieved by increased savings, which translates to investments in infrastructure. This paper focuses on development of a mathematical model for the Energy Intensity as a part of the overall system dynamics in the "Energy Choices" simulation. The approach will use the least square linear regression techniques to produce this mathematical equation.

All the data used for this analysis is from the World Bank database [24] as of Oct. 1st of 2010. This analysis considers the intersection of the 25 most populous countries and the 25 countries that use the most energy. The result is 32 countries that consume over 80% of the world's energy. The 2007 World Bank data on these set of countries was utilized in this model. The results of this entire analysis are only valid for this set of data.

Explanatory/Independent variables of interest

Ten variables were selected to be included. However, in the process of regression analysis some of these variables will be eliminated. Below is the list along with a brief description of Energy Intensity. Please note that the definition of all other variable is found in the World Bank database [24].

1. Energy Intensity (Scarlatos, et.al., 2009): Is the ratio of total energy use in kilo-tons (kt) of oil equivalent to total GDP in US dollars.
2. % of total energy spent on Alternative Energy sources
3. GDP per capita
4. Services correspond to ISIC divisions 50-99
5. Public spending on education as a percentage of Government spending
6. Literacy rate adults as a % of population
7. Electric Power Consumption (KWh per capita
8. Roads paved
9. Information and communication expenditure
10. Net official development and official aid
11. GNI Per Capita

Statistical Analysis

Initial step in the analysis was to check the correlation between EI and several of these independent variables. With the exception for *Net official development and official aid* the 2007 data for independent variables were used. The effect of such investments is not immediate and takes 3-4 years to come to fruition. Therefore the 2003 values for *Net official development and official aid* were used.

The correlation coefficients obtained from the initial step, are used as the criteria for exclusion of certain independent variables. If the absolute value of the correlation coefficient is small(0.0 – 0.4), which indicates a very weak linear correlation we can eliminate the variable at this step. If on the other hand the absolute value of the correlation coefficient is large(0.5 – 1.0) then we will proceed and look at the scatter plot of the each independent variable and EI. As can be seen from the following two tables there are 5 variables that show a promise.

Table 1-Correlation table EI and possible independent variables

Correlation Table	**Energy Intensity most recent (2007) Data Set #1**
Energy Intensity most recent (2007)	1.000
Alt.energy- % of total energy spenton alternative-and-nuclear-energy- 2007	-0.306
Pub.spnd.edu-public spending on education over the period of 2003	0.421
NODA-net-official-development-assist/10^6- 2003	0.142
roads-paved-percentage-of-total-latest reported	-0.314
INF&COM-information-and-communication-t- 2007	-0.543
gni-ppp-current-international-d-2007	-0.263
literacy 15-24(most recent)	-0.635
Electrification('electric-power-consumption-kwh- 2007)	-0.507
GDP-per-capita(2007)	-0.613
services-etc-value-added-percen-2007	-0.712

Before proceeding, a brief discussion on the sign and the direction of the possible relation between these explanatory variables and EI is in order. As shown in the above table all these variables are negatively correlated with EI. This means that for example as the country's electrification rate increases the EI for that country is reduced. This makes sense scientifically as well as intuitively. Theoretically the more developed countries have a higher Electrification index and the more developed countries invest more in infrastructure and therefore are able to use energy more efficiently. As a result EI is lower for more developed countries. Similar arguments may be used to validate the correlation direction for all other the remaining variables that will be considered for further analysis in this paper.

The next step is to look at the scatter plots and run a linear regression model. The details of the process and several iterations which lead to the final model are not included in this paper. Some of these variables were eliminated because one or several of the model assumptions were grossly violated. Some variables were eliminated because of the large *p* values. Some other were eliminated do to large VIF(Variance Inflation Factor) which indicates Colinearity. Only two of the above mentioned independent variables are included in the final model (GDP-per-capita and services-etc.-value added expenditure percentage.

In addition natural logarithmic transformation was found suitable for this model, since the EI data was positively skewed. Natural logarithmic transformation was also used for the independent variables. The final model is presented below:

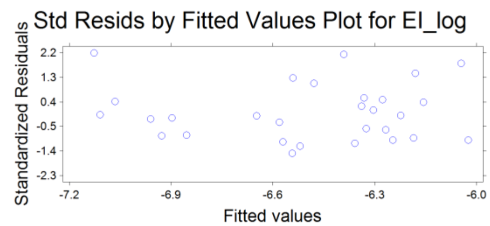
Final Model

Summary	**Multiple R**	**R-Square**	**Adjusted R-Square**	**StErr of Estimate**
	0.8912	0.7943	0.7785	0.401602

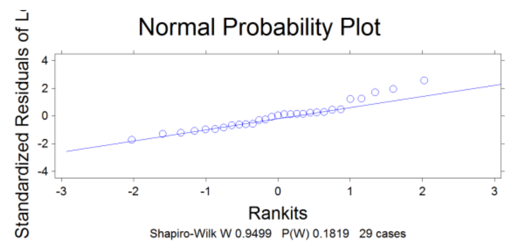
ANOVA Table	**Degrees of Freedom**	**Sum of Squares**	**Mean of Squares**	**F-Ratio**	**p-Value**
Explained	2	16.19298	8.096491	50.2002	< 0.0001
Unexplained	26	4.193385	0.161284		

Regression Table	**Coefficient**	**Standard Error**	**t-Value**	**p-Value**	**Confidence Interval 95% Lower**	**Upper**
Constant	-8.35392	1.149692	-7.2662	< 0.0001	-10.7171	-5.99069
Log(GDP-per-capita(2007))	-0.31931	0.065402	-4.8823	< 0.0001	-0.45375	-0.18488
Log(services-etc-value - 2004)	-0.97957	0.37183	-2.6345	0.0140	-1.74388	-0.21527

Below is the residual versus fitted values plot for the final model. This plot satisfies the linearity and homoscedasticity assumptions.



The following plot is the Normal probability plot.



Translating the output of the above table for the final model into a mathematical equation is the final step in the process. The statistical model for Energy Intensity using least squares linear regression technique based on World Bank data of year 2007 is as follows:

Equation 2

$$Ln(EI) = -8.354 - 0.3193 * Ln(GDP - per - capita) - 0.9796 * Ln(services)$$

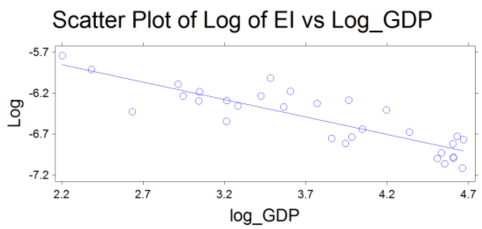
Using logarithmic laws the above equation may be simplified to the following form:

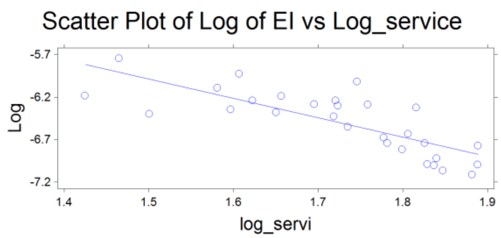
Equation 3

$$\boldsymbol{EI = (e^{-8.354} * GDP - per - capita^{-0.3193} * Services^{-0.9769})}$$

To check the reliability of this model, 2006 data from the World Bank was used to generate a set of calculated values for Energy Intensity. The calculated percentage error and the average value of the %error is less than 3% (=2.053%) . The largest % error was 8.7%.

Based on this model 79.4% of variation in Ln(EI) is explained by the two explanatory variables (independent variables) in the above equation. Earlier in the paper there was a brief discussion on the sign and the direction of the relationship. The fact that EI is negatively correlated with GDP-per-capita and the services makes sense scientifically as well as intuitively. Theoretically the more developed countries have a higher GDP-per-capita and spend more money on Services. More developed countries also invest more in infrastructure and therefore are able to use energy more efficiently. As a result EI is lower for more developed countries .Below are the plots that graphically express this relationship separately for each independent variable:





4. Conclusion

In conclusion this exploratory research has lead to discovery of the relationship between EI, GDP-per-capita and the country's expenditure on Services. This relationship was the result of Least Squares Linear Regression process. The relationship is as indicated in Eq.(2) & Eq.(3) in the above section :

$$Ln(EI) = Ln[(e^{-8.354}) * (GDP - per - capita)^{-0.3193} * (Services)^{-0.9796}]$$

$$\boldsymbol{EI = (e^{-8.354} * GDP - per - capita^{-0.3193} * Services^{-0.9769}}$$

Although this relationship will be used in the "Energy Choices" simulation, the reader must bear in mind that this relationship is based on the stated assumptions in the above sections and it only explains 79% of the variation in EI for year 2007.

This is not a general relationship for EI, but it allows the simulation to use a simplified scientifically based relationship to relay an important message about the climate change and global energy usage in the current context of today's science education.

5. References

1. *Global and Regional drivers of accelerating CO2 emission.* **Raupach, M., Marland, G.,Ciais, P., Que' re,C.L., Canadel, J.G., Klepper G.,& Field C. B.** 2007, PNAS, pp. 10288-10293.

2. *Renewable energy strategies for sustainable development.* **Lund, H.** 2007, ScienceDirect Energy, pp. 912-919.

3. *On the Feasibility of a Timely Transition to a More Sustainable Energy Future.* **Tomkiewiccz, M.** 2010, Sustainabiility, pp. 204-214.

4. *"Beyond STS: A Research-Based Framework for socioscientific Issues.* **Zeidler, D.L. Sadler, T.D., Simmons, M.L. Howes, E.V.** s.l. : Science Education, 2005, Scikence Education, Vol. Vol.89, pp. 357-377.

5. *Simulation Games: one more tool on the pedagogical Shelf. Teaching Socialogy.* **Dorn, D. S.** 1989, Teaching Socialogy, pp. 17,1, 1-18.

6. **Dewey, J.** *Experience and education.* s.l. : Kappa Delta Pi, 1938.

7. *Teaching with Simulation Games.* **Greenblat, C. C.** 1973, Teaching Sociology, pp. 1,62-83.

8. *Intelligent Energy Choices.* **Scarlatos, L.i L. , Tomkiewicz, M. , Bulchandani, A., Srinaivasan, K. , & Naik, P.** Alberta, Canada : Proceedings of the IASTED International Symposium on Modeling and Simulation, 2009. International Symposium on Modeling and Simulation. pp. 670-675.

9. *Time for action: Science education for an alternative future.* **Hodson, D.** 2003, International Journal of Science of Education, pp. 645-670.

10. **Zeidler, D.L., & Keefer, M.** *The role of moral reasoning and the status of socioscientific issues in science education: Philosophical, psychological and pedagogical considerations.* s.l. : Kluwer Academic Publishers, 2003.

11.**Donovan, M.S. and Bransford, J.D. (eds.)** *How Students Learn: History, Mathematics, and Science in the Classroom*Washington D.C.National Academies Press2005

12. **Bransford, J. D., Brown, A.L., & Cocking, R.R.** *How People Learn:Brain,Mind, Experience and school.* s.l. : National Academy Press, 1999.

13. **Vygotsky, L.** *Mind in Society: The Development of Higher Psychological Processes.* s.l. : Harvard, 1978.

14. **Bruner, J.** *Acts of Meaning.* Cambridge : Harvard University, 1990.

15. **Kolb, D. A.** *Experiential learning: experience as the source of learning and development.* N.J. : Englewood Cliffs, 1984.

16. **National Research Council.** *National Science Education Standards.* Washington D.C. : National Academies Press, 1996.

17. *Science Technology and Society as a reform.* **Yager, R.E.** Iowa : s.n., March 1993, pp. 145-152.

18. *The Nature of Science in International Society Education Standard Documents.* **McComas, W.F., & Olson, J.K.** [ed.]

McComas W.F. 2. Netherlands: : Kluwer, 2000, The Nature of Science in Science Education: Rationales and Strategies, pp. 41 – 52.

19. **Yager, R. E.** *How Student Achievement is Affected by STS.* Baltimore : s.n., 2005.

20. *Nature Of Science: Past , Present and Future.* **Lederman, N.G.** s.l. : Handbook of Research on Science and Education, Larwrence Erlbaum associates, 2007, Handbook of Research on Science Education, Lawrence Erlbaum Assoc., Vols. pages:831-879, pp. 831-879.

21. *Global Warming – Science, Money, and Self Preservation.* **Tomkiewicz, M.** 2006, Compte Rendus Chimie, pp. 9, 172-178.

22. **Pedretti, E. G., Bencze L., & Hewitt, J.** *Promoting Issues-based STSE Perspectives in teaching education: Issue of Teacher's Identity and Ideology.* s.l. : Science & Education, 2008.

23. *Research in STS Science education.* **Aikenhead, G. S.** 2005, Educación química, garritz.com, pp. 384-397.

24. **World Bank.** *World Bank Data indicators.* s.l. : available at http://data.worldbank.org/indicator, 2010.