Proceedings of International Conference on Public Policy and Social Science, UiTM Melaka Malaysia, November 2012
ISBN 978-967-11354-5-7

THE ADVANTAGES OF GENETIC ALGORITHM AND NEURAL NETWORKS TO FORECAST AIR POLLUTION TREND IN MALAYSIA: AN OVERVIEW

Mohamad Idham Md Razak[1], Ismail Ahmad[2], Imbarine Bujang[3]
Adi Hakim Talib[4] and Nor Adila Kedin[5]

[1]Lecturer Faculty of Business Management, Universiti Teknologi MARA (UiTM) Melaka, Malaysia
iedham@melaka.uitm.edu.my

[2]Dr Ismail is a Professor Faculty of Business Management, Universiti Teknologi MARA (UiTM), Malaysia
drismailahamd@salam.uitm.edu.my

[3]Dr. Imbarine is a Senior Lecturer Faculty of Business Management, Universiti Teknologi MARA (UiTM), Malaysia

[4, 5]Lecturer Faculty of Computer and Mathematical Sciences, Universiti Teknologi MARA (UiTM) Melaka, Malaysia
adihakim@melaka.uitm.edu.my *and* *noradila_kedin@uitm.edu.my*

ABSTRACT

Genetic Algorithms (GAs) are adaptive heuristic search algorithm premised on the evolutionary ideas of natural selection and genetic. The basic concept of GAs is designed to simulate processes in natural system necessary for evolution, specifically those that follow the principles first laid down. As such they represent an intelligent exploitation of a random search within a defined search space to solve a problem. On the other hand, Neural Networks are composed of interconnecting artificial neurons. Artificial neural networks may either be used to gain an understanding of biological neural networks, or for solving artificial intelligence problems without necessarily creating a model of a real biological system. Thus, these two methods is the advance model for forecasting technique especially involving air pollution prediction and will be discuss as an overview of methodology in this particular research aspect.

Keywords : Air Pollution, Genetic Algorithm, Neural Networks

ISBN 978-967-11354-5-7

1.0 INTRODUCTION

In general, pollution can be described as the deliberate or accidental contamination of the environment with waste that is created by human action. A pollutant is a substance or effect, which adversely alters the environment by changing the growth rate of species, interferes with the food chains, is toxic or interferes with health, comfort, amenities or property values of people. The consequences from this particular activity will create negative externality. Negative externality can best be described as a spillover effect associated with production or consumption that extends to a third party outside the market (Callan SJ. and Thomas JM. 2004). There are various types of pollution, however for this study, the primary focus is on air pollution. Air pollution occurs when our air is contaminated probably caused by the natural and anthropogenic pollutants. Anthropogenic pollutants meaning, the contaminants associated with human activity, including polluting residuals from consumption and production (Callan SJ. and Thomas JM. 2004).

Malaysia is a prosperous country. In order to achieve high-income country by the year 2020, the government proposed aggressive growth from the various economic activities, more manufacturing output and encouraging additional population per year (Lehar, H. 2007). Due to this, the opportunity cost is air pollution emission. This research aims to fill the gap that failed to comply by the previous research. According to Sanglimsuan, K. (2012), he found the empirical evidence that global population change is significantly associated with an increase in carbon dioxide emissions. However, the study was conducted from year 1980 to the year 2007. Thus, the result was obsolete. The main aim of this study will cover the analysis from the year 1970 to the year 2011 in order to come out with the more recent result.

Under the 10th Malaysia Plan from the period 2011 to 2015, the government outlined several strategies to promote the sustainable development and environmental conservation. Environmental conservation cannot rely purely on a sense of responsibility. The Government will, therefore, promote economic opportunities that create value from conservation. For example, eco-tourism can generate income particularly for local communities to encourage the conservation of the country's flora and fauna. We must also seize opportunities that arise from emerging trends, where green products, services and technology are increasingly in demand. In this respect, I call upon industries to take this opportunity and use the incentives provided by the Government through the Green Technology Financing Scheme worth RM1.5 billion to enhance the application of green technology in the production of goods and provision of services (10th MP). In addition, The Government will also promote environmentally friendly housing by introducing guidelines and a green rating system. Putrajaya and Cyberjaya will serve as flagship green townships. The Government will take the lead in adopting green building standards. New Government buildings will be designed to meet green standards. Energy efficiency of existing buildings will be enhanced and as a showcase example, the Prime Minister's Office complex will be upgraded to meet the Gold Standard Green rating (10th MP). Thus, from this research, it will cater well the country with the relevent information needed with the concrete analysis in order to achieve the objective stated in 10th MP.

2.0 LITERATURE REVIEW

The research findings, presented by each of the authors cited in this chapter, were balanced against one another in order to establish a sound basis for this research and to provide a reference by which any contributions to knowledge, presented herein, could be measured. The literature review for this research was conducted over a period of several years. The objective of the review was acquire an understanding of the current state of

ISBN 978-967-11354-5-7

knowledge in the field in which this research was undertaken and to identify key research groups, seminal authors and forums where such research was presented and subjected to peer review. The review process involved an examination of numerous texts, refereed journal publications, conference proceedings, internet-sourced publications and trade journals.

The literatures were limited to three variables based on the IPAT Model proposed by Commoner, Corr and Stamler (1971) and Ehrlich and Holdren (1971). Based on their study, the equation is I = *f* (P, A, T), where I is environmental impact, P is Population, A is Affluence and T is Technology. However, since there is a limitation in term of availability of the data, the variables are as follow; I is Air Pollution Index (API), P is Population, A is Affluence and T is Manufacturing Industry. Specifically, the authors, and their influence on this research are summarized as follows.

Global population keeps increasing year by year. There are numerous literature found that direct correlation between increasing number of population will lead to high air pollution emission. According to Sanglimsuan K. (2012), he found the empirical evidence that global population change is significantly associated with an increase in carbon dioxide emissions. The study was achieved by analyzing the relationships between carbon dioxide emissions, population, and other related factors. The empirical study is based on cross-country data from 83 countries from 1980 to 2007. The results show that population pressure has impact on growth in carbon dioxide emissions. The researcher also added that the results has been confirmed that population is a critical factor for manipulation of carbon dioxide increase.

Furthermore, the population of industrialized countries such as the United States or of countries from the European Union spends approximately more than one hour each day in vehicles. In this respect, numerous studies have so far addressed outdoor air pollution that arises from traffic (Mueller D. et al. 2011). In addition, in both developed and rapidly industrializing countries, the major historic air pollution problem has typically been high levels of smoke and Sulphur Dioxide (SO2) arising from the combustion of sulphur-containing fossil fuels such as coal for domestic and industrial purpose. At the current status, the major threat to clean air is now posed by traffic emissions. Traffic congestion is the result of the relationship between transport and land-use, where impacts are obvious in several metropolitans (Sertel, E. et al. 2008).

The relationship between economic growth and the environment is controversial according to Ekins (1999) and Xepapadeas (2003). Grossman and Krueger (1991, 1994) and the World Bank (1992) are the pioneering studies, which brought the EKC to public attention. Grossman and Krueger’s (1991) study explored the relationship between economic growth and pollution measures for air quality, and their focus on the year 1994 studied on water quality. Since its discovery, much statistical evidence on the EKC has accumulated for many other pollution measures based on the research being done by Barbier (1997), Cole et al (1997), Suri and Chapman (1998). Bradford et al (2000), Harbaugh (2002), and Stern (2003).

According to the research being done by Selden and Song (1994), Stokey (1998), and Andreoni and Levinson (2001), they have also been many attempts to derive the EKC theoretically.The dominant theoretical explanation is that when GDP increases, the greater scale of production leads directly to more pollution, but, at a higher level of income per capita, the demand for health and environmental quality rises with income which can translate into environmental regulation, in which case there tend to be favorable shifts in the composition of output and in the techniques of production.

Greenhouse gases production is the only one element of the environmental footprint of conventional manufacturing process technologies. Volatile organic compounds (VOCs) are common by-products of a wide range of industrial processes in the United States including petroleum refining, chemical production, food & beverage production, forest & paper

ISBN 978-967-11354-5-7

products manufacturing, printing, packaging, and industrial coating. Next, the VOCs react with air when exposed to sunlight and are a major source of ground level ozone in example is smog. Additionally, many VOCs threaten human health, being linked to cancer, asthma and birth defects (Brunkreef and Holgate 2002). Because of these environmental and human health effects, VOC emissions are addressed by the Clean Air Act and tightly regulated by the EPA.

Industrial processes are the single largest source of Volatile Organic Compound (VOC) pollution emissions with nearly 6 million tons emitted into the atmosphere annually (USEPA 2006). The EPA and most states currently regulate VOC emissions, mandating that industrial sources exceeding some threshold, typically 25 tons per year are required to install some type of abatement equipment. Thermal oxidation systems using natural gas fueled burners to incinerate VOCs account for nearly 75% of abatement technologies installed in the US (CMR 2006). They are capital intensive, have high operating costs, and consume substantial energy. Due to the potential economic impact of tightening emission regulation, the EPA requires only the largest polluters to install abatement equipment, allowing smaller manufacturers to emit VOC pollutants up to a cap. The cost of pollution abatement equipment, both to install and to operate, is significant for manufacturers. And although VOC emission caps protect the smallest manufactures from economic hardship, medium sized manufacturers are often forced to limit their production output in order to stay under limits. Limiting production in an underutilized plant means less return on invested capital and loss of global manufacturing share to manufacturers operating in less regulated countries.

3.0 METHODOLOGY

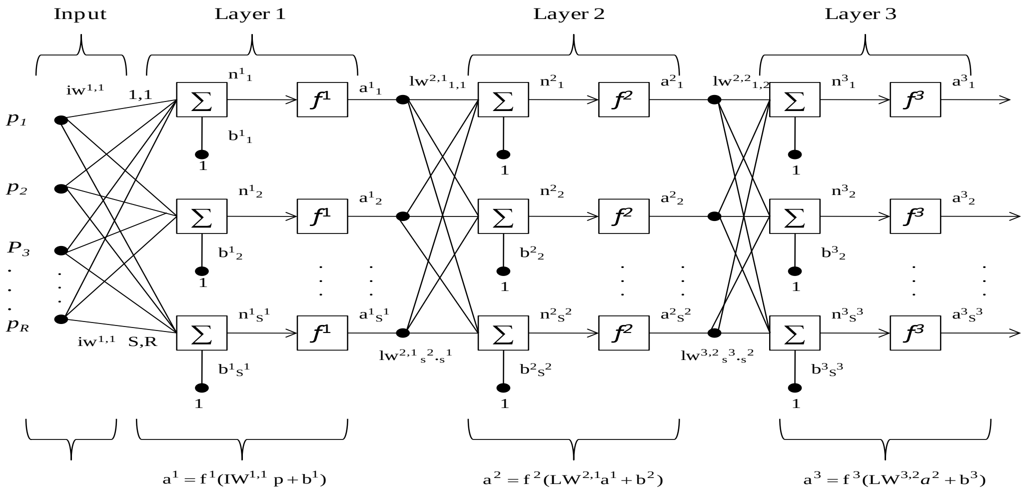
3.1 GENERAL NEURAL NETWORKS

An Artificial Neural Network (ANN), simply referred to as a Neural Network, is a computer program designed to model the human brain and its ability in terms of learning and information processing **(Haykin, 1994)**. Early artificial neural networks were inspired by biological nervous systems. The primary features of artificial neural networks are derived from two characteristics of the brain: the ability to 'learn' and to generalize from limited information **(Hewitson and Crane, 1994)**. In recent years the development in ANN technology has evolved into an applied mathematical technique that has some similarities to the human brain **(Easson, 1996)**.

Learning of a neural network can be divided into supervised learning and unsupervised learning. The former refers to the learning pattern that results in networks with adjustable parameters updated by a supervised learning rule **(Jang, 1997)**, and the latter consists of only input training data and is trained without human intervention. All the neural networks with supervised learning capabilities can be unified into the framework of adaptive neural network. An adaptive network is a neural network structure comprising a number of nodes connected through directional links **(Jang, 1997)**. Each node, or neuron, represents a process element, and the links between the nodes specify the relationship between the connected nodes. Adaptive neural networks are composed of a large number of interconnected neurons working in parallel to solve a specific problem. Two or more neurons can be combined to form a layer, and a network can contain several layers. Multiple-layer networks are very powerful. It is proved that a network of two layers with sigmoid as the function of the first layer and linear as the function of the second layer can be trained to approximate any function **(Demuth and Beale, 2000)**. Figure 3 displays a three-layer neural network.

ISBN 978-967-11354-5-7

Figure 3: A Three Layer Neural Network



In order to utilize an ANN to solve a problem, the first step is to train the ANN to 'learn' the relationship between the input and output. This is accomplished by presenting the network with examples of known inputs and outputs, in conjunction with a learning rule **(Easson, 1996)**. The learning rule specifies how the network parameters, i.e., weights and biases, should be updated to minimize the discrepancy between the network's actual output and the desired output **(Jang, 1997)**. In the learning process, initial weights to the connections in the architecture of the ANN are assigned, and then the ANN iteratively adjusts the interconnection weights until the ANN can successfully produce output values that match the original values. These weighted matrixes of interconnections allow the neural network to recognize and respond to patterns in data **(Obermeier and Barron, 1989)**, and create a mapping of input-output variables. The mapping can be used to make simulation and prediction according to other input data within the range.

The ANN models have been developed in various disciplines to recognize patterns or approximate functions from complicated data and to make predictions. Recently, this approach has been applied to air pollution simulation and prediction **(Comie, 1997; Gardner and Dorling, 1998; Ruiz-Suarez et al, 1995)**. The ANN can be trained to screen data to detect patterns, to identify potential problems or opportunities, or to discover similarities between current and past situations. Air pollution mechanism is complex and dynamic in nature, and can be represented by various interactions that operate on different spatial-temporal scales where different aspects can lead to the same impact in example, the emission of acid deposition pollutants. Since air pollution formation and transportation are complex and nonlinear processes, and the air pollution data may be imprecise and complicated, the neural network approach is appropriate for air pollution prediction.

As one of the nontraditional modeling techniques, ANN has been used to deal with a variety of problems in the environmental field. When traditional methods produce unsatisfactory results, the ANN could provide an alternative, or in some cases, represents a significant improvement **(Robinson, 1991)**. The development of statistical models for air quality prediction has been a subject that involved extensive research efforts, resulting in many models **(Korsog and Wolff, 1991; Feister and Balzer, 1991; Abdul-Wahab et al., 1996; Katsoulis, 1996; Fiore et al., 1998)**. However, the main limitations of statistical techniques are the rigid assumptions that are essential for justifying their applications, such as those of sample size, linearity, and continuity (Huang, 2003). Neural networks, with their ability to derive meaning from complicated or imprecise data, have been shown as an effective alternative to more conventional statistical techniques (Schalkoff, 1992).

ISBN 978-967-11354-5-7

The neural network approach exhibits several advantages over traditional phenomenological models. The most important advantage is that it can solve problems that are too complex for conventional technologies such as statistical methods. These problems include pattern recognition and function approximation. Their applicability is increasing in air quality predictions because of their ability to handle uncertainties and complex relationships. Other advantages include rapid information procession, the ability to develop a mapping of the input and output variables. The ANN has been applied in the field of air quality prediction, where there is a lack of understanding of the complex nonlinear relationships between meteorology and pollution. ANN models have been used to forecast air quality behaviour for NO*x* and NO2 **(Gardner and Dorling 1999, Hasham et al, 2004)**, SO2 **(Fernandez de Castro et al. 2003; Boznar et al. 1993)**, O3 **(Cannon and Lord 2000; Comrie 1997;Yi and Prybutok 1996)**, PM10 **(McKendry 2002; Chelani et al. 2002)**, PM2.5 **(McKendry 2002; Perez et al. 2000)**, and H2S and NH3 **(Rege and Tock 1996)**. In addition, **Benvenuto and Marani (2000)** used neural networks for data quality control of environmental time series and reconstruction of missing data.

Although there have been research papers in using neural network approach to make predictions in atmospheric field, more researchers are expected to take advantage of ANN in this area. Here is some research work on neural network in the air pollution prediction area. **Comrie (1997)** compared neural networks and multi-regression models for ozone forecasting. In this study, multiple regression models and neural networks are developed and examined for eight cities under different meteorological conditions and ozone concentrations. Model comparison statistics indicate that neural network techniques are consistently better than regression models for daily ozone prediction to some extent, and neural network models have some advantage over regression models in terms of inherently incorporating nonlinear relationships and thus make somewhat more accurate predictions of ozone than regression models using the same set of input data. It is also concluded that all types of models are sensitive to different weather-ozone regimes and the role of persistence in aiding predictions.

Gardner and Dorling (1999) applied the neural network to model and predict hourly NOx concentrations in an urban area of London. In this study, Multilayer perceptron(MLP) neural networks with meteorological input data were used to predict the concentrations. They were trained using the scaled conjugate gradient algorithm network with two hidden layers. Based on the data used by **Shi and Harrison (1997)**, Gardner and Dorling conducted a comparison study with a regression model developed by Shi and Harrison. Gardner and Dorling concluded that MLP neural networks could accurately model the relationship between local meteorological data and NO2/NO*x* concentrations in an urban environment. The neural network models are seen to learn the underlying pattern of emissions without any external guidance. This enables the models to be easily constructed. Gardner and Dorling also noticed that MLP neural networks have several advantages over traditional multiple linear regression models. These include the ability of MLP models to make efficient use of proxy data when the optimum predictor variables are unavailable.

Grubert (2003) used neural network approach to predict acid deposition in the Eastern United States. The author employed backpropagation training algorithm to train neural networks on total sulphur dioxide emissions from power plants and measured field data on precipitation chemistry. These trained networks were then applied to predict seasonal changes in sulphate, hydrogen, nitrate, and ammonium ion concentrations caused by projected decreases in sulphur dioxide emissions by the power plants in the eastern United States. After comparing forecasted acid ion concentration results with actual monitoring site values, Grubert concluded that neural networks could predict future wet deposition trends provided that the extrapolation process was done in steps.

ISBN 978-967-11354-5-7

3.2 GENERAL GENETIC ALGORITHM

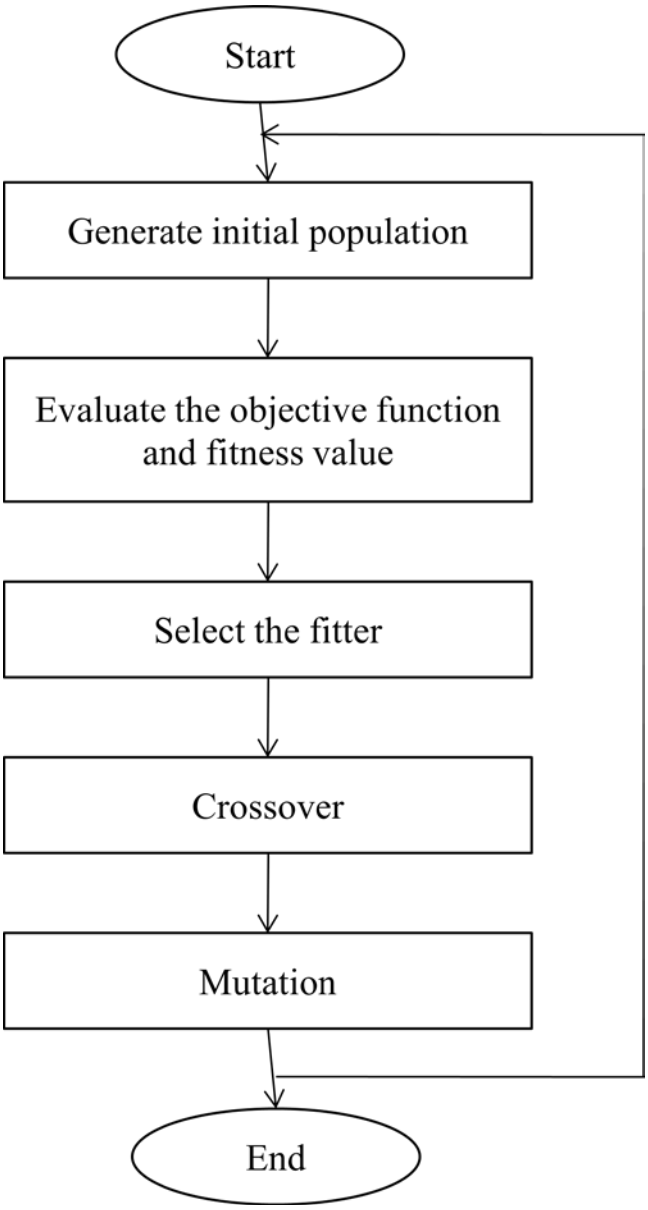
Gas can be applied to obtain good solutions for many problems to which traditional optimization approaches have not proven successful. Several GA approaches have been developed and applied into the water management field **(Scott, 1995; Liong, 1995; Brian 1994).** However, there are almost no such cases in air quality management especially in Malaysia. GAs is class of probabilistic procedures that search for good solutions to problems by emulating the "survival to the fittest" concept seen in nature. The principle idea of the Gas can be summarized as follows.

In a GA, a potential solution to a problem is most often represented as a vector of values or genes. In the context of this model, each gene may represent the allowable emission level at a controlling emission source in the study region, based on this context is Malaysia. In GA the set of potential strategies are also called a population, generally consisting of about 50 to 200 strategies, which are generated at random or seeded with good solutions. The problem is subjected to several probabilistic operators that are analogous to natural selection, mating (including genetic combination) and mutation.

In the selection step, pairs of strategies may then undergo mating or crossover to form two new strategies. The new strategies are then ordered to create a new population. The selective and mating steps continue until the new population is the same size as the current population. The performance of each strategy in the population has been generated, use Monte Carlo simulation to evaluate the fitness of a strategy. For example, the vector representing an air control strategy is first decoded to determine which emission level are allowable at each source. An air quality model is then run to determine the resulting air quality. The strategy is assigned a fitness that is a function of how effectively it meets the ambient target as well as other modeled objectives and constraints. After repeating the process for required number of generations, best strategies are sure to be found. The GA search process is depicted in Figure 4.

ISBN 978-967-11354-5-7

Figure 5: The Flowchart of GA



3.3 NEURAL NETWORKS AND GENETIC ALGORITHM

In forecasting one usually starts with the linear regression model, given by the following equation.

$$Y_t = \Sigma \beta_k XL_{k,t} + \varepsilon_t \quad (1)$$

$$\varepsilon_t \sim N(0, \sigma^2) \quad (2)$$

Where the variable εt is a random disturbance term, normally distributed with mean zero and variance σ2, and {βκ} represent the parameters to be estimated. The set of estimated parameters is denoted {βeκ} while the set of forecasts of Y generated by the model with the coefficients {βeκ} is denoted by {Yet}. The goal is to select {βeκ} in order to minimize the sum of squared differences between the actual observations Y and the observations predicted by the linear model Ye. While there are a number of different computational methods for selecting {βeκ} depending on the structure of the model, the basic

ISBN 978-967-11354-5-7

concept is to use the implied stochastic process of the error term to generate a maximum likelihood solution to the problem. Commonly this results in the autoregressive linear forecasting model:

$$Y_t = \sum_{t=1}^{K^*} \beta_i Y_{t-i} + \sum_{j=1}^{K} \gamma_j X_{j,t} + \varepsilon_t \qquad (3)$$

in which there are K independent X variables, with coefficient γj for each xj, and K* lags for the dependent variable y with, of course, K+K* parameters {β} and {γ} to estimate. Thus the longer the lag structure, the larger the number of parameters to estimate, and the smaller are the degrees of freedom of the overall regression estimates.

The number of output variables may be more than one. In the case of orthogonal error terms ε, each may be estimated as a single equation as above. In the case of correlation between the error terms, a vector auto- regression approach is appropriate. The linear model has the useful property of having a closed form solution for solving the regression problem of minimizing the sum of squared differences between {Y} and {Ye}. Thus the linear method is quick and easily understood.

For short run forecasting, the linear model is a reasonable starting point or benchmark. This is because in many markets one observes only small symmetric changes in the variable to be predicted about the long-term trend. From the Weierstrass Theorem, a polynomial expansion around a set of inputs x with a progressively larger power P is capable of approximating to a given degree of precision any unknown but continuous function y = g(x). Consider for example a second degree polynomial approximation of three variables, [x1t , x2t, x3t] and a continuous but unknown function g. The approximation formula for g would be as equation (16):

$$Y_t = {}_0 + B_1x_{1t} + B_2x_{2t} + B_3x_{3t} + B_4x^2_{1t} + B_5x^2_{2t} + B_6x^2_{3t}\ B_7x_{1t}x_{2t} + B_8x_{1t}x_{3t} + B_9x_{2t}x_{3t} \qquad (4)$$

Note that the second-degree polynomial expansion with three arguments or dimensions has three cross terms {β7, β8, β9} which represent the cross effects of the exogenous variables on y. This simple expansion requires the estimation of ten parameters and the degrees of freedom in the estimate will decline exponentially as the degree of the polynomial in the expansion increases. This “curse of dimensionality” is an exacting price to pay for accuracy in the fitting of the non-linear model to the data in the econometric estimation.

As an alternative to these traditional methods, the neural network approach is a more parsimonious estimation technique. The reason why one uses the neural network is simple and straightforward. The goal is to find an approach or method which forecasts well data that are generated by often unknown and highly non0linear processes with as few parameters as possible.

Like the linear and polynomial approximation methods, a neural network relates a set of input variables {xi}, i=1,...,k to a set of one or more output variables{yj}, j=1,...,k*. The difference between a neural network and the other approximations methods is that the neural network makes use of one or more “hidden layers”, in which the input variables are transformed by a logistic or log-sigmoid transformation. The appeal of the log- sigmoid transform function comes from its “threshold behavior” which characterizes many types of economic responses to changes in fundamental variables.

Furthermore, the shape of the logsigmoid function reflects a kind of learning behavior. Often used to characterize learning by doing, the function becomes increasingly steep until some inflection point. Thereafter the function becomes increasingly flat up and its slope

ISBN 978-967-11354-5-7

moves exponentially to zero. Following the same example, as interest rates begin to increase from low levels, consumers will judge the probability of a sharp uptick or downtick in the interest rate based on the currently advertised financing packages. The more experience they have, up to some level, the more apt they are to interpret this signal as the time to take advantage of the current interest rate, or the time to postpone a purchase. The results are markedly different than those experienced at other points on the temporal history of interest rates. The following system of equations are most commonly used in the "feed forward" neural network:

$$n_{k,t} = \omega_{k,0} + \sum_{i=1}^{i*} \omega_{k,i} x_{i,t} \qquad (5)$$

$$N_{k,t} = 1 / (e^{-ni,t}) \qquad (6)$$

$$y_t = \gamma_0 + \sum_{k=1}^{k*} \gamma k N_{k,t} \qquad (7)$$

In this system there are i* input variables {x} and k* neurons. Equations (49) through (51) are interpreted as: at any time t the a convex combination of the input variables x (equation (49)) are transformed by the logsigmoidal transform (equation (50)) and are input into the output equation (51) to generate a forecast of the variables of interest y. It is easy to see that this is simply a nonlinear expansion of the function g for purposes of estimation. Researchers have used other transform functions such as the hyperbolic tangent function as well, however the logsigmoidal remains the most successful transform function to date.

4.0 CONCLUSION

Neural networking promises to provide computer science breakthroughs that rival anything we have yet witnessed. Once neural networks are trained properly, they can replace many human functions in targeted areas. We hope that our application will provide a small but important step in that journey.

Now the question remains, what is the difference between human and neural networks? Both can learn and become expert in an area and both are mortal. The main difference is, humans can forget but neural networks cannot. Once fully trained, a neural net will not forget. Whatever a neural network learns is hard-coded and becomes permanent. A human's knowledge is volatile and may not become permanent. There are several factors that cause our brain cells to die and if they do, the information that is stored in that part is lost and we start to forget.

The other difference is accuracy. Once a particular application or process is automated through a neural network, the results are repeatable and accurate. Whether the process is replicated one thousand times or one million times, the results will be the same and will be as accurate as calculated the first time. Human beings are not like that. The first 10 processes may be accurate, but later we may start to make mistakes in the process. Another key difference is speed. Neural networks can be hardware or software. It is obvious that neural networks are much faster than humans in processing data and information. Therefore, by using this particular method, we expect that the outcome for forecasting is more accurate and transparent.

ISBN 978-967-11354-5-7

REFERENCES

Abdul-Wahab S., Bouhamra W., Ettouney H., Sowerby B. and Crittenden B.D. (1996). *Predicting ozone levels: a statistical model for predicting ozone levels.* Environmental Science and Pollution Research*.* 3, 195-204.

Acharyya J. (2009). FDI, Growth and the environment: Evidence From India On CO2 emission during the last two decades. *Journal of Economic Development*, volume 28, 43-56.

Ahn, S.K. and C, G. Reinsel (1990). Estimation for Partially Nonstationary Multivariate Autoregressive Models. *Journal of the American Statistical Association*, 85, 813-823.

Andreoni, J. and A. Levinson. (2001). "The Simple Analytics of the Environmental Kuznets Curve," *Journal of Public Economics*, 80, 269-86.

Barbier, E. (1997), "Environmental Kuznets Curve Special Issue: Introduction," *Environment and Development Economics*, 2, 369-381.

Blevins, J. R. (2011). Lecture 16: Heteroskedasticity. *Econ 444: Elementary Econometrics, The Ohio State University* , 1-2.

Bongaarts J. (1992). Population growth and global warming. *Popul Dev Rev.* vol.18, 299 – 319.

Bongaarts J, O'Neill BC, and Gaffin SR. (1997). Global warming policy: population left out in the cold. *Environment,* 39, 40–41.

Boznar, M., Lesjak, M., and Mlakar, P. (1993). *A neural network-based method for short term predictions of ambient SO2 concentrations in highly polluted industrial areas of complex terrain*. Atmos. Environ. 27B(2): 221–230.

Bradford, D., R. Schlieckert and S. Shore (2000), "The Environmental Kuznets Curve: Exploring a Fresh Specification, NBER Working Paper No.8001.

Breusch, T., and A. Pagan. (1979). A Simple Test of Heteroskedasticity and Rrandom Coefficient Variation. *Econometrica 47*, 1287-1294.

Brunekreef B. and Holgate ST. (2002). Air Pollution and Health. 360 (9341), 1233-42.

Callan SJ. and Thomas JM. (2004). Environmental Economics & Management.Theory, Policy, and Applications, 3d ed, *Thomson South-Western*, 7.

Campbell, John Y., and Pierre Perron. (1991). Pitfalls and opportunities: what macroeconomists should know about unit roots. *NBER Macroeconomics* Annual 6: 141-201.

Cannon, A.J., and Lord, E.R. (2000). *Forecasting summertime surface level ozone concentrations in the Lower Fraser Valley of British Columbia: an ensemble NN approach*. J. AirWaste Manag. Assoc.50: 322–339.

ISBN 978-967-11354-5-7

Castillo-Velaquez, JI., Hugo-Tulio Rubio-Rodriguez, Carmen-Araceli Eudave-Loera. (2007).

Air Pollution in Puebla City: A mathematical model, 1-4.

Chelani, A.B., Gajghate, D.G., and Hasan, M.Z. (2002). *Prediction of ambient PM10 and toxic metals using artificial neural networks*. J. AirWaste Manag. Assoc. **52**: 805–810.
Chris C. Park, Acid *Rain Rhetoric and Reality*. Methuen & Co. Ltd.

Cheung, Y.W. and K.S. Lai. (1993). Finite-sample sizes of Johansen's likelihood ratio tests for
cointegration. *Oxford Bulletin of Economics and Statistics*.

Cohen JE. (1995). How Many People Can The Earth Support? *New York: Norton*.

Commoner, B., Corr, M., & Stamler, P.J. (1971). The causes of pollution. Environment. 13(3), 2-19.

Comrie, A.C. (1997). *Compare neural neural networks and regression models for ozone forcasting*, Journal of Air and Waste Management Association, 47, 653-663

Cole, M. A., A. J. Rayner and J. M. Bates. (1997). "The Environmental Kuznets Curve: An Empirical
Analysis," *Environment and Development Economics*, 2, 401-416.

Cole MA and Neumayer E. (2004). Examining the impacts of demographic factors on air pollution.
Popul Environ, vol.26, 5–21.

Cooley, T.F. and LeRoy, S.F. (1985), Atheoretical Macroeconomics; A Critique. *Journal of Monetary Economics*, vol. 16, no. 3, 283-308.

Cramer, J. C. (1998). Population growth and local air pollution: methods, models, and results. In W.
Lutz, A. Prskawetz & W. C. Sanderson (Eds.), *Population and environment* (A supllement to Vol. 28).
Population and development review, pp. 22-52. New York: population Council.

Curran S, de Sherbinin A. (2004). Completing the picture: the challenges of bringing 'consumption'
into the population-environment equation. *Popul Environ*, vol. 26 no. 2, 107–31.

Dalton M, O'Neill BC, Prskawetz A, Jiang L, Pitkin J. (2007). Population aging and future carbon emissions in the United States. *Energy Econ.*

Demuth Howard and Beale Mark (2000). *Neural Network Toolbox User's Guide*. The MathWorks, Inc.

Department of Statistics (DOS); http://www.statistics.gov.my/portal/index.php?lang=en

Dietz, T., & Rosa, E. A. (1997). Effects of population and affluence on CO_2 emissions. *Proceedings of
the National Academy of Sciences*, 94(1), 175-179.

ISBN 978-967-11354-5-7

Dietz T, Rosa EA, York R. (2007). Driving the human ecological footprint. *Front Ecol Environ*, vol. 5, no.1, 13–18.

Dockery DW and Pope CA 3rd. (1993). Acute respiratory effects of particulate air pollution. *Annu Rev Public Health*, 15, 107–32.

Dockery DW, Pope CA 3rd, Xu X, Spengler JD, Ware JH, Fay ME, Ferris BG Jr, Speizer FE. (1993). An association between air pollution and mortality in six U.S. cities. *N Engl J Med*, 329(24), 1753-9.

DOEC NSW. (2005). Health cost of Air Pollution in the greater Sydney Metropolitan Region, 1.

Durbin, J. and G. S. Watson. (1951) Testing for Serial Correlation in Least-Squares Regression, *Biometrika, 38,* 159 -177.

Eagle, R. F. and C.W. Granger (1987). Co-integrated and error correction: representation, estimation and testing. *Econometrica*, 55, 251-76.

Easson, Gregory L. (1996). *Integration of Artificial Neural Networks and Geographic Information Systems for Engineering Geological Mapping*, Unpublished Doctoral Dissertation. University of Missouri--Rolla, 154p.

Ehrlich PR. (1968). The Population Bomb*. New York: Ballantine Books*.

Ehlirch, P.R., & Holdren, J.P. (1971). Impact of population and growth. *Science*, 171, 1212-1217.

Ekins, P. (1999). Economic Growth and Environmental Sustainability – The Prospects for Green Growth. *London: Routledge*.

Ericsson NR. (1998). Empirical modeling of money demand, Empirical Economics. *Springer*, vol. 23(3), pages 295-315.

E. Sertel, H. Demirel, , S. Kaya, I.Demir. (2008). Spatial prediction of transport related urban air quality. 805-808.

Feister U. and Balzer K.(1991). *Surface ozone and meteorological predictors on a subregional scale*. Atmospheric Environment 25, 1781-1790.

Fernandez de Castro, B.M., Sanchez, J.M.P., Manteiga,W.G., Bande, M.F., Bermudez Cela, J.R., and Fernandez, J.J.H. (2003). *Prediction of SO2 levels using neural networks*. J. Air Waste Manag. Assoc. 53: 532–539.

Fiore A.M., Jacob D.J., Logan J.A. and Yin J.H. (1998). *Long-term trends in ground level ozone over the contiguous United States*, 1980-1995. Journa*l of Geophysical Research* 103, 1471-1480.

ISBN 978-967-11354-5-7

Furno, M. (2005). The Glejser Test and the Median Regression. *The Indian Journal ofStatistics, 67(2)* , 335-358.

García-Pérez, J., Elena Boldo, Rebeca Ramis, Marina Pollán, Beatriz Pérez-Gómez, Nuria Aragonés and Gonzalo López-Abente. (2007). *BMC Public Health,* 1-13.

Gardner, M.W., and Dorling, S.R. (1999). *NN modeling and prediction of hourly NOx and NO2 concentrations in urban air in London*. Atmos. Environ. 33: 709–719.

Glejser, H. (1969). A new test for heteroscedasticity. *J. Amer. Statist. Assoc.*, 64, 316- 323.

Glinianaia SV, Rankin J, Bell R, Pless-Mulloli T, Howel D. (2004). Particulate air pollution and fetal health: a systematic review of the epidemiologic evidence. *Epidemiology*, 15(1), 36-45.

Godfrey, L. G. (1978). Testing Against General Autoregressive and Moving Average Error Models When the Regression Includes Lagged Dependent Variables, *Econometrica, 46,* 1293-1302.

Godfrey, L. G. (1996). Some results on the Glejser and Koenker test for heteroscedasticity. *J. Econometrics*, 72, 275-299.

Gonzalo, J. and C.W.J. Granger. (2001). Estimation of Common Long-Memory Components in Cointegrated Systems. *Harvard University Press Cambridge*, MA, USA.

Greene, W. H. (2003). *Econometric Analysis* (Fifth ed.). New York: Prentice Hall.

Grossman, G. M. and A. B. Krueger. (1991). “Environmental Impacts of a North American Free Trade Agreement,” NBER Working Paper No.3914, *Cambridge: National Bureau of Economic Research*.

Grossman, G. M. and A. B. Krueger (1994), “Economic Growth and the Environment,” *Quarterly Journal of Economics*, 110, 353-377.

Grubert John P. (2003), *Acid deposition in the eastern United States and neural network predictions for the future.* NRC Research Press, http://jees.nrc.ca/

Gujarati, D. M., and D. C. Porter. (2010). Essential of Econometrics. New York:McGraw Hill.

Harbaugh, W., A. Levinson and D. M. Wilson. (2002), “Reexamining the Empirical Evidence for an Environmental Kuznets Curve,” *The Review f Economics and Statistics*, 84(3), 541-551.

Hasham Faizal A., Kindzierski Warren B., and Stanley Stephen J. (2004). *Modeling of hourly NOx concentrations using artificial neural networks*. NRC Research Press, http://jees.nrc.ca/.

Haykin, Simon (1994). Neural Networks, *A Comprehensive Foundation*. New York: Macmillan College Publishing Company, 696p.

Heinrich J, Hoelscher B, Wjst M, Ritz B, Cyrys J, Wichmann H. (1999). Respiratory diseases

ISBN 978-967-11354-5-7

and
allergies in two polluted areas in East Germany. *Environ Health Perspect*,107(1), 53-62.

Hewitson, Bruce C. and Crane Robert G. (1994). *"Looks and Uses" in Hewitson*, *Neural Nets: Applications in Geography*. Boston: Kluwer Academic Publishers, pp. 1-9.

H. Merbitz, M. Buttstädt, S. Michael, W. Dott, and C. Schneider. (2012), GIS-based Identification of Spatial Variables Enhancing Heat and Poor Air Quality in Urban Areas Applied Geography. Volume 33.

Hug, A.A. (1996). Tests for cointegration: a Monte Carlo comparison. *Journal of Econometrics.*

Jang Jyh-Shing Roger, Sun Chuen-Tsai, Mizutani Eiji (1997). Neuro-Fuzzy Soft Computing, A Computational Approach to Learning and Machine Intelligence, Prentice Hall, Upper Saddle River, NJ 07458

Jiang L and O'Neill BC. (2004). The energy transition in rural China. *Int J Glob Energy Issue,* 21, 2 26.

Jing- Yuan, G. and Jia-Qing, L. (2011);
http://ieeexplore.ieee.org/xpl/freeabs_all.jsp?arnumber=6070027

Johansen, S. (1988). Statistical Analysis of cointegration vectors. *Journal of Economic Dynamics and Control*, 12. 231-54.

Johansen, S. and Juselius, K. (1990). Maximum Likelihood Estimation and Inference on Cointegration – with Applications to the Demand for Money. *Oxford Bulletin of Economics and Statistics* 52, 169210.

Katsoulis, B. D. (1996) *The relationship between synoptic, mesoscale and microscale meteorological parameters during poor air quality events in Athens*, *Greece*. Science of the Total Enironment 181, 13 – 24.

Künzli N, Kaiser R, Medina S, Studnicka M, Chanel O, Filliger P, Herry M, Horak F Jr, Puybonnieux Texier V, Quénel P, Schneider J, Seethaler R, Vergnaud JC, Sommer H. (2000). Public-health impact of outdoor and traffic-related air pollution: a European assessment. *Lance*t, 356(9232), 795-801.

Korsog P.E. and Wolff G.T. (1991). *An examination of urban ozone trends in the northeastern U.S. (1973-1983) using a robust statistical method. Atmospheric Environment* 25, 47-57.

Kwiatkowski D., Phillips P., Schmidt P. and Shin Y. (1992). Testing the null hypothesis of stationary against the alternative of a unit root: how sure are we that economic time series have a unit root?,*Journal of Econometrics*, 54, 159-178.

ISBN 978-967-11354-5-7

Lantz V and Feng Q. (2006). Assessing income, population, and technology impacts on CO_2 emission
in Canada: Where's the EKC? *Ecol Econ,* 57, 229–38.

Lehar, H. (2007). Malaysia Economics Past and Present. *UPENA*.
Lutz W, Sanderson WC, Scherbov S. (2001). The end of world population growth. *Nature,* 412, 543– 45.

Malthus TR. (1798). An Essay on the Principle of Population as It Affects the Future Improvement of Society. *London: Johnson*.

Marco, L. (2011); http://www.brighthub.com/environment/science-environmental/articles/17296.aspx

MacKellar FL, Lutz W, Prinz C, Guojon A. (1995). Population, households, and CO_2 emissions. *Popul Dev Rev*, vol. 21, 849–65.

McKendry, I.G. (2002). *Evaluation of artificial neural networks for fine particulate pollution (PM10 and PM2.5) forecasting*. J. Air Waste Manag. Assoc. 52: 1096–1101.

Meadows DH, Meadows DL, Randers J, Behrens WW. (1972). The Limits to Growth: A Report for the Club of Rome's Project on the Predicament of Mankind. *New York: Universe Books*.

Meyerson FAB. (1998). Population, carbon emissions, and global warming: the forgotten relationship at Kyoto. *Popul Dev Rev,* vol. 24,115–30.

Müller D, Klingelhöfer D, Uibel S, Groneberg DA. (2011). Car indoor air pollution - analysis of potential
sources. *J Occup Med Toxicol*, 6(1), 33.

Nakicenovic N. (2000). Special Report on Emissions Scenarios (SRES) Cambridge. *UK: Cambridge Univ. Press.*

Neumayer E. (2002). Can natural factors explain any cross-country differences in carbon dioxide emissions? *Energy Policy,* vol. 30, 7–12.

Neumayer E. (2004). National carbon dioxide emissions: geography matters. *Area,* vol. 36, 33–40.

Newey, W. and K. West. (1987). A simple, positive semi-definite, heteroscedasticity and autocorrelation consistent covariance matrix, *Econometrica 55(3),* 703–708.

Obermeier *Klaus K. and J* & Barron *anet J.* (1989). *Time to get fired up,* Byte, 14(8), p217-224.

Odhiambo, NM. (2012). http://journals.cluteonline.com/index.php/JABR/article/view/6682

Perez, P., Trier, A., and Reyes, J. (2000). *Prediction of PM2.5 concentrations several hours in advance using neural networks in Santiago, Chile*. Atmos. Environ. 34: 1189– 1196.

ISBN 978-967-11354-5-7

Pesaran, H.M. and Shin, Y. (1995), "Autoregressive Distributed Lag Modelling Approach to Cointegration Analysis". *DAE Working Paper Series* No. 9514, Department of Applied Economics,University of Cambridge.

Phillips, P.C.B, and P. Perron (1988).Testing for a Unit Root in Time Series Regressions. *Biometrika* 75, 335-346.

Phillips, P. and B. Hansen (1990). Statistical inference in instrumental variables regression with I(1) processes*. Review of Economic Studies* 57, 99-125.

Pope CA 3rd, Burnett RT, Thun MJ, Calle EE, Krewski D, Ito K, Thurston GD. (2002). Lung cancer, cardiopulmonary mortality, and long-term exposure to fine particulate air pollution. 287(9), 11341.

Rahbek, Anders and Hansen, Henrik. (2002). Approximate Conditional Unit Root Inference. *Journal of Time Series Analysis*, Vol. 23, pp. 1-28.

Ramsey, J. B. (1969) Tests for specification errors in classical least-squares regression analysis, *Journal of the Royal Statistical Society, Series B*, 31, 350-71.

Ramsey, J. B. and A. Alexander. (1984).The Econometric Approach to Business Cycle Analysis Reconsidered, *Journal of Macroeconomics, 6,* 347–356

Rege, M.A., and Tock, R.W. (1996). *A simple neural network for estimating emission rates of hydrogen sulfide and ammonia from single point sources.* Journal of Air and Waste Management Association., 46, 953-962.

Reimers, H.E. (1992). Comparisons of Test for Multivariate Cointegration. *Statistical Papers* 33, 335 359.

Robert, T.D (2011); http://www.sciencedirect.com/science/article/pii/S0143622810000743

Robinson, R. (1991). *Neural networks offer an alternative to traditional regression,* Geobyte, 14-19.

Ruiz-Suarez J.C., Mayora-Ibarra O.A., Torres-Jimenez J., and Ruiz-Suarez L.G.(1995). *Short-term ozone forecasting by artificial neural networks*, Advances in Engineering Software, 23, 143-149.

Saldiva PH, Pope CA 3rd, Schwartz J, Dockery DW, Lichtenfels AJ, Salge JM, Barone I, Bohm GM. (1995). Air pollution and mortality in elderly people: a time-series study in Sao Paulo, Brazil. *Arch Environ Health*, 50(2), 159-63.

Sanglimsuwan, K. (2012). The impact of population pressure on carbon dioxide emission: evidence from a panel-econometric analysis. *International Research Journal of Finance and Economics*, 82(2012), 90-93.

ISBN 978-967-11354-5-7

Schalkoff, R. (1992). *Pattern Recognition: Statistical, Structural and Neural Approaches*, Wiley, New York.

Schmidheiny, K. (2010). Heteroskedasticity in the Linear Model. *Short Guides to Microeconometrics Fall 2010, Unversitat Pompeu Fabra* , 1-10.

Schmidt P. (2007). A Robust Version of the KPSS Test Based on Indicators. *Journal of Econometrics*, 137.

Schwartz J and Marcus A. (1990). Mortality and air pollution in London: a time series analysis. *Am J Epidemiology*, 131(1), 185-94.

Schwartz J, Dockery DW, Neas LM, Wypij D, Ware JH, Spengler JD, Koutrakis P, Speizer FE, Ferris BG Jr. (1994). Acute effects of summer air pollution on respiratory symptom reporting in children. *AmJ Respir Crit Care Med*, 150(5 Pt 1), 1234-42.

Seldon, T. and D. Song. (1994). “Environmental Quality and Development: Is There a Kuznets Curve for Air Pollution Emissions?” *Journal of Environmental Economics and Management*, 27, 147-162.

Sims, C. A.: 1980, Macroeconomics and reality. *Econometrica* 48, 1—48.

Stern, D. I. (2003). “The Rise and Fall of the Environmental Kuznets Curve,” Rensselaer Polytechnic Institute, Working Paper No.0302, October.

Stokey, N. (1998). “Are There Limits to Growth,” *International Economic Review*, 39(1), 1-31.

Stock, J. and M. Watson (1993). A simple estimator of cointegrating vectors in higher order integrated systems*. Econometrica* 61(4), 783-820.

Suri, V. and D. Chapman. (1998). “Economic Growth, Trade and Energy: Implications for the Environmental Kuznets Curve,” *Ecological Economics*, 25(2), 147-160.

The McIlvaine Company. (2006).Thermal / Catalytic Markets Report; http://home.mcilvainecompany.com/.

US Environmental Protection Agency. (2006). Air Emission Sources; www.epa.gov/air/emissions/voc.htm.

Van Vuuren DP and O’Neill BC. (2006). The consistency of IPCC’s SRES scenarios to recent literature and recent projections. *Clim Change,* 75, 9–46.

Vlachokostas C., Charisios Achillas, Theodora Slini, Nicolas Moussiopoulos, Georgios Banias, Ioannis Dimitrakis. (2011). Willingness to pay for reducing the risk of premature mortality attributed to air pollution: a contingent valuation study for Greece. *Athmospheric Pollution Research*, 276-281.

Wasserman, N. and M. Kunter. (1990). Applied Linear Statistical Models, 3rd ed., Irwin.

White, H. (1980). A heteroskedasticity-consistent covariance matrix estimator and a

ISBN 978-967-11354-5-7

direct test for heteroskedasticity. *Econometrica 48(4)*, 817–838.

World Bank. (1992). “Development and the Environment,” *World Development Report*.

Xepapadeas, A. (2003). “Economic Growth and the Environment,” in preparation for the Handbook of Environmental Economics, J. Vincent and Karl Goran Male eds.

Yi, J., and Prybutok, V.R. (1996). *A neural network model forecasting prediction of daily maximum ozone concentration in an industrialized urban are*a. Environ. Pollut. 92: 349–357.

York, R., Rosa, E.A., & Dietz, T. (2003a). STRIPAT, IPAT and ImPACT: analytical tools for unpacking the driving forces of environmental impacts. *Ecological Economics*, 46(3), 351-365.

York, R., Rosa, E.A., & Dietz, T. (2003b). Footprints on the earth: the environmental consequences of modernity. *American Sociology Review*, 68(2), 279-300.

Zhu. (2012); http://in.news.yahoo.com/hong-kong-air-pollution-worst-levels-ever-report 024302043.html 10 Malaysia Plan (MP) 2010 - 2015; http://www.epu.gov.my/html/themes/epu/html/RMKE10/rmke10_english.html