Forecast Global Carbon Dioxide Emission Using Swarm Intelligence

Alaa Abdelfatah, Sahar A. Mokhtar, Alaa Sheta
Computers and Systems Department
Electronics Research Institute (ERI)
Cairo, Egypt

Basma Solaiman
College of Computer Science and Information Technology
Sudan University of Science and Technology (SUST)
Khartoum, Sudan

ABSTRACT

The tremendous effects of air quality in large cities have been considered a severe environmental problem all over the world. Therefore, the international community agreed to develop air quality standards to monitor and control pollution rates around industrial communities. Harmful emission into the air is a sign that could extremely affect man's health, natural life and agriculture. Forecasting models is essential for predicting air quality. CO_2 emissions have been an international concern because of fossil fuels. In this study, Particle Swarm Optimization (PSO) is used for analyzing world CO_2 emission based on the global energy consumption. A parametric PSO model is developed to forecast CO_2 emission based set of attributes. They include: global oil, natural gas, coal, and primary energy consumption. A data set collected during the years 1980 and 2010 were used in this study. Experimental results show that PSO can provide good modeling results using a limited number of measurements compared to other linear models.

General Terms:

Climate Change, Carbon Dioxide Emission, Particle Swarm Optimization

Keywords:

Air Pollution, Carbon Emission, Swarm Intelligence

1. INTRODUCTION

Climate change towards the increase of air pollution presented a significant environmental harmfulness. This reason causes an increased interest from both academics and policy-makers in developing prediction model in the area of CO2 emissions. Air quality prediction model helps in the management of our environment. The air pollution prediction is a complex nonlinear process function of numerous parameters [16, 13, 17]. Air quality information is collected from various sources and sensors to be used in developing a successful relationship which helps producing accurate prediction models. Mathematical methods and tools were introduced to solve variety of problems in system identification, control, stock market prediction and air pollution modeling [10].
The complexity of air pollution data has been broadly discussed in [15], while the usage of various modeling tools was addressed in related literature [13], [21]. In [3], authors used a new panel data set covering the fifty United States and Washington D.C. from 1960-2001 to model CO_2 emission. Their objective was to measure of model performance using the squared out-of-sample prediction error of aggregate CO_2 emissions. CO_2 emissions and economic growth in Nigeria for the period 1971-2009, in a multivariate framework was presented in [8]. Authors in [19] provided a study to develop statistical models for predicting the carbon dioxide emissions and the atmosphere in the United States. A monthly emissions data from 1981 to 2003 that was collected by the Carbon Dioxide Information Analysis Center was used. The developed statistical models took into consideration various trends and seasonal effects. Forecasting the expected growth of China's CO_2 emissions using province-level information was presented in [2].
Air pollution has been explored in many articles based soft computing and computational intelligence techniques [15, 4]. For example, in [1] authors develop a non-parametric Artificial Neural Network (ANN) models to predict both the Particulate Matters (PM10) and Total Suspended Particles (TSP) in Salt, Jordan. Two Artificial Neural Network based AutoRegressive with eXternal (ANNARX) Input models were used to provide high performance modeling for the PM10 and the TSP air pollution parameters.
An integrated multi-layer perceptron neural network and Bees Algorithm was proposed to analyze the world CO_2 emissions [5]. A data set from the years 1980 to 2006 were used, the data for the years 1980-1999 were used for training and the testing data used was for the years 2000-2006. The model was used to predict the World CO2 emission up to year 2040. Several models have been proposed to investigate the causal relationships between energy consumption and CO_2 emission. For instance, the relatively new time series technique known as the Toda-Yamamoto method [12], and the Grey prediction model (GM) [17], have been applied to predict CO_2 levels. Neural networks have been used for short-term CO_2 prediction [5]. In addition methods like Genetic Algorithm (GA) [11] have been employed for predicting CO_2 emission concentrations.
In this paper, our objective is to develop a mathematical model for the global CO_2 emission based on a set of variables. They include global oil, natural gas, coal and primary energy consumption over the period from 1980 to 2010 using a Particle Swarm Optimization (PSO). The proposed model should be able to provide an accurate estimation results in our case.

2. AIR POLLUTION

Fossil energy resources remain abundant but contain significant amounts of carbon that are normally released during combustion. The proven and probable reserves of oil and gas are enough to last for decades and in the case of coal for centuries. Possible undiscovered resources extend these projections even further. Fossil energy use is responsible for about 85% of the anthropogenic CO_2 emissions produced annually [14]. In Figure 1, we show the origin of CO_2 emissions from all fossil fuel combustion by region of the world. It was found that almost half of the total emissions come from Organization of Economic Cooperation and Development (OECD) countries, excluding Mexico, and only 20 percent are emitted in China, but only seven percent are from Latin America.

3. PARTICLE SWARM OPTIMIZATION

Particle Swarm Optimization (PSO) was developed in 1995 by James Kennedy (social-psychologist), and Russell Eberhart (electrical engineer) [9]. PSO is a robust stochastic nonlinear-optimization technique based on movement and intelligence of swarms. It is inspired from social behavior of bird or fish, where a group of birds randomly search for food in an area by following the nearest bird to the food. It combines local search methods with global search methods (exploration and exploitation). Moreover, it depends on social interaction between swarms to locate the best achieved position so far. The main idea of PSO is obtaining a number of particles that constitute a swarm moving around in the search space looking for the best solution. Each particle is treated as a point in a multi-dimensional space. The particles move around in the search space, and communicate either directly or indirectly with one another.

In the implementation of PSO, each particle tries to modify its position using information as: its current position, its current velocity, the distance between the current position and best solution individually found, and the distance between the current position and the best solution found in its neighborhood [7]. The basic PSO equations are given as follows:

$$v_{id}^{new} = v_{id}^{old} + \phi_1 * c_1 * (p_{id} - x_{id}) + \phi_2 * c_2 * (p_{gd} - x_{id}) \quad (1)$$

$$x_{id}^{new} = x_{id}^{old} + v_{id}^{new} \quad (2)$$

where:
v_{id} represents the velocity of particle i in dimension d,
x_{id} represents the position of particle i in dimension d,
ϕ_1, ϕ_2 are positive constants,
c_1, c_2 are random numbers
p_{id} is the best position reached so far by the particle, and
p_{gd} is the global best position reached by the neighborhood.

The previously shown velocity update equation is influenced by two components: $(p_{id} - x_{id})$ that represents the cognitive component, and $(p_{gd} - x_{id})$ that represent the social component. The velocity of PSO is prune to reach infinity which lead to causing the particles to reach a state of instability. Therefore, the velocity of the particle should not exceed a maximum value, v_{max}. The performance can suffer if maximum velocity is inappropriately set. If it is too high, the particles can fly past optimal solutions, and if it is too low, they can get stuck in local minimal. More enhancements on the basic PSO model are stated in [20].

The PSO algorithm is straightforward (see Algorithm 1). First, initialize particles with random position and velocity vectors. For each particle: evaluate the fitness and if it is better than the best individual fitness then update it. After that, update the best global fitness. Then obtain the new velocity and position for each particle. This procedure is repeated for a number of iterations (epochs) or until convergence is beyond a certain limit.

Algorithm 1: Basic steps describing the PSO algorithm

```
Begin PSO
  Randomly initialize the position and velocity of the particles:
  X_i(0) and V_i(0)
  While (While terminating condition is not reached) do
      for for i = 1 to number of particles
        Evaluate the fitness:= f(X_i)
        Update p_i and g_i
        Update velocity of the particle V_i
        Update position of the particle X_i
        Evaluate the population.
      Next for
  End While
End PSO
```

3.1 Nature of Data

The Particle Swarm Optimization (PSO) model was used to forecast the CO_2 emission using the four energy commodities that are considered the highest. Long-term (annual) forecasting was performed. The related data from 1980 to 2010 were used, partly for installing the models (finding candidates of best weighting factors for each model (1980-2003) and partly for testing the models (2004 to 2010). The values of global oil, natural gas, coal, and primary energy consumption are presented in [11, 18].

3.2 Evaluation Criterion

In order to check the performance of the developed PSO model, Route Mean Square (RMS), the Euclidian distance (ED), the Manhattan distance (MD) and the Variance-Accounted-For (VAF) were measured. These performance criteria were assessed to measure how close the measured values to those developed using the PSO approach. This criterion helps in better evaluating the capabilities of the developed PSO model. The RMS, VAF, ED and MD are computed as follows:

(1) Route Mean Square (RMS):

$$RMS = \sqrt{\frac{\sum_{i=1}^{n} |(y_i - \hat{y}_i)|}{n}} \quad (3)$$

(2) Variance-Accounted-For (VAF)

$$VAF = \left[1 - \frac{var(y - \hat{y})}{var(y)})\right] \times 100\% \quad (4)$$

(3) Euclidian distance (ED):

$$ED = \sqrt{\sum_{i=1}^{n} (y_i - \hat{y}_i)^2} \quad (5)$$

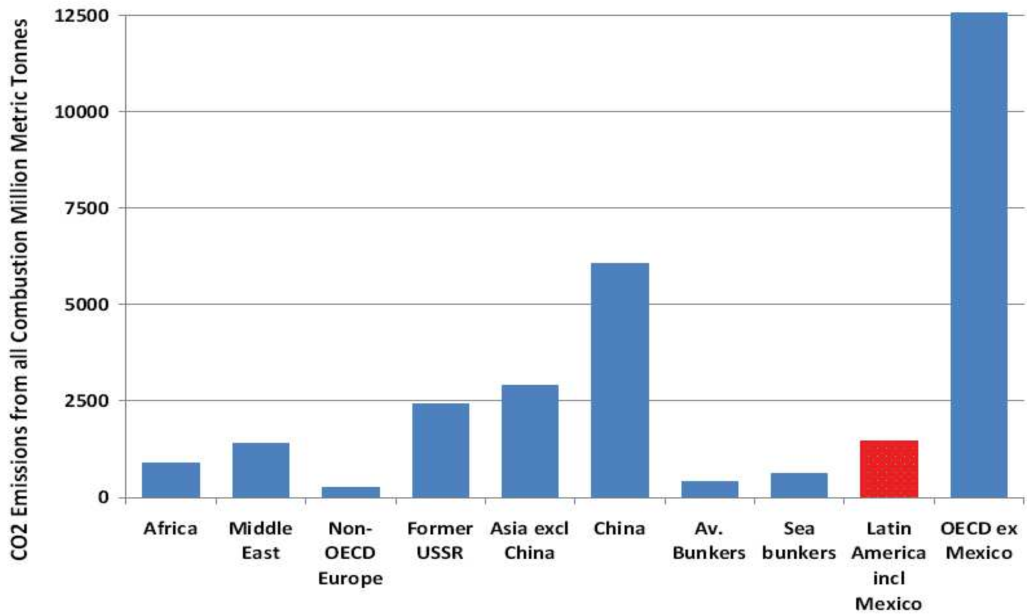

Fig. 1. CO_2 emissions from all fossil fuel combustion by country or region in 2006 (million metric tonnes).Source: International Energy Agency (IEA2008)

(4) Manhattan distance (MD):

$$MD = (\sum_{i=1}^{n} |y_i - \hat{y}_i|) \tag{6}$$

where y and $\hat{y}$ are the observed and estimated CO_2 of the proposed model and n is the number of measurements used in the experiments, respectively.

4. EXPERIMENTAL RESULTS

The input and output data used for the model is presented in Table 1. The proposed model is an exponential model of the input variables. The tuning parameters and search space for PSO is given in Table 2. A Matlab Toolbox for PSO was used to develop our results [6]. Candidate solutions (particles) in this case are just n-dimensional vectors of particles of the form:

α_1	β_1	α_2	β_2	α_3	β_3	α_4	β_4	γ

The search process begins by random distribution of the initial PSO population which represents a random sample of the PSO search space. The corresponding fitness of each particle is computed. Particles with higher fitness shall be selected to produce new population which enhances the features of their parents. The problem under consideration is to estimate the correct parameters α_1, β_1, α_2, β_2, α_3, β_3, α_4 and β_4 for the proposed exponential model.

A data set adopted from [11] for the model development process was used. We defined the parameter space and other PSO setup tuning parameters are given in Table 2. The fitness (i.e. quality) of a particular estimate is obtained by observing the behavior of the system with the estimated parameters, and using the RMS error between the actual and predicted CO_2.

$$CO_2 = \alpha_1 Oil^{\beta_1} + \alpha_2 NG^{\beta_2} + \alpha_3 Coal^{\beta_3} + \alpha_4 PE^{\beta_4} + \gamma \tag{7}$$

Table 1. Inputs and Output for the Exponential PSO model

Inputs	Oil, NG, $Coal$, PE
Output	CO_2

Table 2. The tuning parameters for the PSO

Operator	Value
Acceleration constant	[2.1,2.1]
Inertia Weight	[0.9,0.6]
Maximum no. of Iteration	1000
Maximum Velocity	150
No. of Particles	9
Search space for α and β	[-10,10]
Search space for γ	[-1000,1000]

The computed model parameters using PSO was found as given in Equation 8.

$$\begin{aligned} CO_2 &= 8.0639\ Oil^{-4.2278} + 5.9998\ NG^{-0.63625} \\ &+ 4.4493\ Coal^{-0.14562} + 2.4769\ PE^{1.1308} + 3.7837 \end{aligned} \tag{8}$$

Previous research work presented in [11] explored the use of Genetic Algorithms to find the best tuning parameters for the same exponential model. The developed model is presented in Equation 9.

$$\begin{aligned} CO_2 &= 0.3427\ Oil^{2.0509} + 0.9188\ NG^{0.9115} \\ &+ 0.0092\ Coal^{0.9096} + 0.25644\ PE^{0.2827} + 0.3597 \end{aligned} \tag{9}$$

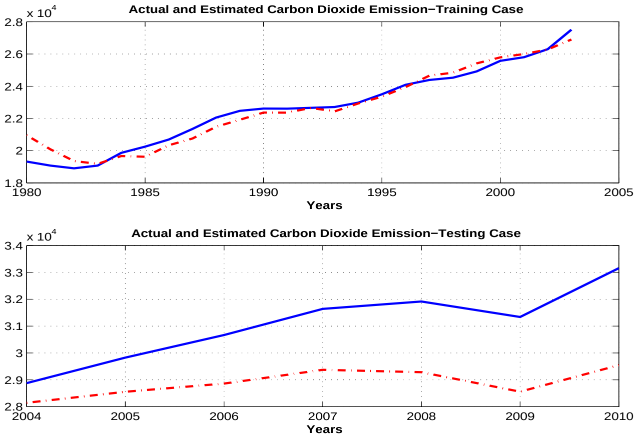

Fig. 2. Observed and Estimated Exponential PSO Model

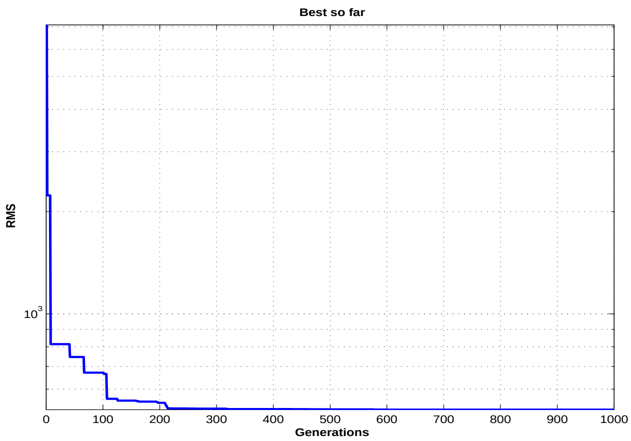

Fig. 3. RMS convergence of the PSO process

Table 3. Evaluation criteria of the PSO model

Criteria	RMS	VAF	ED	MD
Training	537.86	92.978%	2464.8	401.14
Testing	2121.4	72.913%	6708.4	1740.2

It was found that the model developed using PSO is much stable and can provide better evaluation criterion which was adopted in this study.

5. CONCLUSIONS

This paper provided an exponential model for the global carbon emission based on global oil, natural gas, coal, and primary energy consumption attributes. A data set of the years (19802010) was used for estimating nine parameters of the model using PSO in both the training and testing cases. The developed results showed the advantages of PSO in tuning the model than GAs. Future work will focus on exploring the advantage of evolutionary computation techniques such as genetic programming to develop a mathematical model for the global carbon emission.

6. ACKNOWLEDGEMENT

Basma Solaiman is a PhD student with the College of Computer Science and Information Technology, Sudan University of Science and Technology (SUST), Khartoum, Sudan. She is on leave from the New and Renewable Energy Authority (NREA), Cairo, Egypt.

7. REFERENCES

[1] Mouhammd Al-kasassbeh, Alaa F. Sheta, Hossam Faris, and Hamza Turabieh. Prediction of pm10 and tsp air pollution parameters using artificial neural network autoregressive, external input models: A case study in salt, jordan. *Middle-East Journal of Scientific Research*, 14(7):999–1009, 2013.

[2] Maximilian Auffhammer and Richard T. Carson. Forecasting the path of china's CO2 emissions using province-level information. *Journal of Environmental Economics and Management*, 55(3):229–247, May 2008.

[3] Maximilian Auffhammer and Ralf Steinhauser. Forecasting the path of U.S. CO2 emissions using state-level information. *The Review of Economics and Statistics*, 94(1):172–185, February 2012.

[4] A. Baareh. Solving the carbon dioxide emission estimation problem: An artificial neural network model. *Journal of Software Engineering and Applications*, 6(7):338–342, 2013.

[5] M. A. Behrang, E. Assareh, M. R. Assari, and A. Ghanbarzadeh. Using bees algorithm and artificial neural network to forecast world carbon dioxide emission. *Energy Sources, Part A: Recovery, Utilization, and Environmental Effects*, 33(19):1747–1759, 2011.

[6] B. Birge. PSOt - a particle swarm optimization toolbox for use with Matlab. In *Proceedings of the 2003 IEEE Swarm Intelligence Symposium (SIS03)*, pages 182–186, 2003.

[7] D. Bratton and J. Kennedy. Defining a standard for particle swarm optimization. In *IEEE Swarm Intelligence Symposium*, pages 120 – 127, 2007.

[8] E. Nnaji Chibueze, O. Chukwu Jude, and Nnaji Moses. Electricity supply, fossil fuel consumption, CO2 emissions and economic growth: Implications and policy options for sustainable development in nigeria. *International Journal of Energy Economics and Policy*, 3(3):262–271, 2013.

[9] Andries P. Engelbrecht. *Computational Intelligence: An Introduction*. Wiley Publishing, 2nd edition edition, 2007.

[10] Kostas D. Karatzas, George Papadourakis, and Ioannis Kyriakidis. Understanding and forecasting atmospheric quality parameters with the aid of ANNs. In *IEEE International Joint Conference on Neural Networks (IJCNN 2008)*, pages 2580–2587. IEEE, 2008.

[11] H Kavoosi, MH Saidi, M Kavoosi, and M Bohrng. Forecast global carbon dioxide emission by use of genetic algorithm. *International Journal of Computer Science Issues (IJCSI)*, 9(1), 2012.

[12] Mohammad Reza Lotfalipour, Mohammad Ali Falahi, and NULL NULL. Economic growth, CO2 emissions, and fossil fuels consumption in Iran. *Energy*, (35):5115–5120, October 2010.

[13] Marek Makowski. Modeling paradigms applied to the analysis of european air quality. *European Journal of Operational Research*, 122(2):219–241, 2000.

Table 4. Actual and Estimated CO_2 Values Using the Exponential PSO Model

Year	Oil Consumption $(Mtoe)^a$	NG Consumption $(Mtoe)$	Coal Consumption $(Mtoe)$	PE Consumption $(Mtoe)$	CO_2 Emission $(Mt)^b$	CO_2 Estimated $(Mt)^b$
1980	2972.2	1296.9	1806.4	6624	19322.4	20953
1981	2863	1309.5	1820.6	6577.5	19073.2	20084
1982	2770.7	1312.5	1846.9	6548.4	18900.7	19354
1983	2748.3	1329	1897.7	6638.2	19072.1	19177
1984	2810.1	1440	1983.2	6960.2	19861	19665
1985	2804.7	1488.3	2056	7137.5	20246.7	19623
1986	2894.1	1503.6	2089.2	7307.5	20688.3	20331
1987	2946.8	1579.6	2169	7555.7	21344.5	20750
1988	3038.8	1654.9	2231.7	7833.5	22052.2	21484
1989	3093	1729.2	2251.2	8001.7	22470.2	21918
1990	3148.6	1769.5	2220.3	8108.7	22613.2	22364
1991	3148.2	1807.5	2196.4	8156	22606.5	22360
1992	3184.8	1817.9	2174.6	8187.6	22656.7	22655
1993	3158	1853.9	2187.6	8257.5	22710.6	22439
1994	3218.7	1865.4	2201.9	8357.6	22980.3	22927
1995	3271.3	1927	2256.2	8577.9	23501.7	23351
1996	3344.9	2020.5	2292.2	8809.5	24089.8	23946
1997	3432.2	2016.8	2301.8	8911.6	24387.1	24654
1998	3455.4	2050.3	2300.2	8986.6	24530.5	24842
1999	3526	2098.4	2316	9151.4	24922.7	25417
2000	3571.6	2176.2	2399.7	9382.4	25576.9	25789
2001	3597.2	2216.6	2412.4	9465.6	25800.8	25998
2002	3632.3	2275.6	2476.7	9651.8	26301.3	26285
2003	3707.4	2353.1	2677.3	9997.8	27508.7	26900
2004	3858.7	2431.8	2858.4	10482	28875.2	28145
2005	3908.5	2511.2	3012.9	10800.9	29826.1	28556
2006	3945.3	2565.6	3164.5	11087.8	30667.6	28860
2007	4007.3	2661.3	3305.6	11398.4	31641.2	29373
2008	3996.5	2731.4	3341.7	11535.8	31915.9	29284
2009	3908.7	2661.4	3305.6	11363.2	31338.8	28557
2010	4028.1	2858.1	3555.8	12002.4	33158.4	29545

[14] Bert Metz. *Climate Change 2007-Mitigation of Climate Change: Working Group III Contribution to the Fourth Assessment Report of the IPCC*, volume 4. Cambridge University Press, 2007.

[15] Francesco Carlo Morabito and Mario Versaci. Fuzzy neural identification and forecasting techniques to process experimental urban air pollution data. *Neural Networks*, 16(3-4):493–506, 2003.

[16] A. Newcomer, S. Blumsack, J. Apt, L.B. Lave, and M.G. Morgan, editors. *Electricity Load and Carbon Dioxide Emissions: Effects of a Carbon Price in the Short Term*. 41st Hawaii International Conference on System Sciences, Computer Society Press, 2008.

[17] Hsiao-Tien Pao and Chung-Ming Tsai. Modeling and forecasting the CO2 emissions, energy consumption, and economic growth in brazil. *Energy*, 36(5):2450 – 2458, 2011.

[18] Reza Samsami. Application of ant colony optimization (ACO) to forecast CO2 emission in Iran. *International Journal of Environment, Pharmacology and Life Sciences, Bulletin of Environment, Pharmacology and Life Sciences*, 2:95–99, 2013.

[19] Shou Hsing Shih and Chris P. Tsokos. Prediction models for carbon dioxide emissions and the atmosphere. *Neural, Parallel Sci. Comput.*, 16(1):165–178, January 2008.

[20] Ioan Cristian Trelea. The particle swarm optimization algorithm: convergence analysis and parameter selection. *Information Processing Letters*, 85(6):317–325, 2003.

[21] Takashi Washio, Einoshin Suzuki, Kai Ming Ting, and Akihiro Inokuchi, editors. *Advances in Knowledge Discovery and Data Mining, Proceedings of the 12th Pacific-Asia Conference, PAKDD 2008, Osaka, Japan, May 20-23*, volume 5012 of *Lecture Notes in Computer Science*. Springer, 2008.