

# Dragonfly Optimization based ANN Model for Forecasting India's Primary Fuels' Demand

J. Kumaran  
Assistant Professor in  
Computer Science and  
Engineering, Pondicherry  
Engineering College,  
Puducherry, India

J. Sasikala  
Assistant Professor  
Computer Science and Engineering,  
Annamalai University  
India

## ABSTRACT

This paper presents a dragonfly optimization (DFO) based ANN model for predicting India's primary fuel demand. It involves socio-economic indicators such as population and per capita GDP and uses two ANNs, which are trained through DFO algorithm. The method optimizes the connection weights of ANN models through effectively searching the problem space in finding the global best solution. Primary fuel demands during the years 1990-2012 forms the data for training and validating the model. The proposed model requires an input, the year of the forecast, and predicts the primary fuels' demand. The forecasts up to the year 2025 are compared with that of the RM with a view to illustrate the accuracy.

## General Terms

forecasting, primary energy fuels

## Keywords

artificial neural network, dragonfly optimization, regression model.

## Nomenclature

ARIMA	autoregressive integrated moving average
ANN	artificial neural network
DFO	dragonfly optimization
$no$	number of outputs
$F_i$	demand of $i$ -th primary fuel
GDP	gross domestic product
MSE	mean squared error
MAPE	mean absolute percent error
N	number of training data set
$no$	number of outputs
$O_i^n(k)$	output of $i$ -th neuron for $n$ -th training data for ANN model-k
$Pop$	population
PM	proposed model
RM	regression model
$W_{ih}$	weight matrix in between input and hidden layers
$W_{ho}$	weight matrix in between hidden and output layers
$W_{bias\_h}$	bias vector for hidden layer
$W_{bias\_o}$	bias vector for output layer
$y_i^n(k)$	targeted value for $i$ -th neuron for $n$ -th training data for ANN model-k

## 1. INTRODUCTION

Energy is one of the most fundamental parts of our universe. Energy has come to be known as a 'strategic commodity' and any uncertainty about its supply can threaten the functioning of the economy, particularly in developing economies. The energy-mix of any country comprises both non-renewable (coal, lignite, crude oil and natural gas) and renewable energy sources (wind, solar, small hydro, biomass, cogeneration bagasse etc.) [1].

The total consumption of raw coal by industries in India has increased from 407.04 MTs during 2005-06 to 570.23 MTs during 2012-13. The annual growth rate for coal consumption from 2011-12 to 2012-13 is 6.41%. Electricity generation is the biggest consumer of coal, followed by steel and washery (15.88 MTs), cement (13.55 MTs) and paper industries (2.13 MTs). The consumption of Lignite has increased from 30.23MTs in 2005-06 to 46.41 MTs in 2012-13 registering a compound growth of 5.51%. The major consumption of Lignite is in the sector of electric power generation, accounting for about 80.38% of the total lignite consumption. The consumption of crude oil has a steady increase, from 130.11 MMTs during 2005-06 to 219.21 MMTs during 2012-13 with annual growth rate of 6.74%. The consumption of Natural Gas has a firm increase, from 31.03 BCMs during 2005-06 to 38.4 BCMs during 2012-13. The maximum use of Natural Gas is in power generation (33.46%) followed by fertilizers industry (27.87%) and domestic use (5.20%) [2].

Coal is the most abundant conventional source of energy in the country. However, the average quality of the Indian coal is not very high compared to those available in Australia or Canada. Further, the coal washing capacity in the country has not increased sufficiently, due to various reasons, to generate the required quantity of washed coal for consumption, particularly in steel plants. This necessitates the import of high quality coal to meet the requirements of steel plants. There has been an increasing trend in the import of coal. This is evident from the fact that the gross import of coal has steadily increased from 36.60 MTs during 2005-06 to 134.73 MTs during 2012-13 [2].

India is highly dependent on import for crude oil. Oil import has been steadily rising over the years. There is hardly any export of crude oil from India. Therefore, both gross import of crude oil have increased from 11.68 MTs during 1970-71 to 159.26 MTs during 2009-10. There has been an annual increase of 19.9% during 2009-10 over 2008-09, as the net import increased from 99.41MTs during 2005-06 to 184.80 MTs during 2012-13. Although more than 70% of its crude oil requirements and part of the petroleum products is met from imports, India has developed sufficient processing

capacity over the years to produce different petroleum products so as to become an exporter of petroleum products [2].

Economic development and energy are deeply inter-related. Information on the future demands of non-renewable sources of energy is a pre-requisite for ascertaining the country's economy through allocating appropriate funds for future procurement of energy fuels. Affordable energy is critical for sustaining economic growth and for improved living standards of its population, which in turn is driven by increasing energy consumption. Hence, it is recognized that energy security in a fast developing economy like India is of strategic importance and energy needs of the country are likely to increase in the coming decades. Achieving energy security in this strategic sense is of fundamental importance not only to India's economic growth but also for the human development objectives that aim at alleviation of poverty and unemployment. Holistic planning for achieving these objectives requires quality energy statistics that is able to address the issues related to future demand for primary fuels.

There was little or no systematic procedure for forecasting before 1950s. Although methodologies such as regression and time-series decomposition were available, their applications were severely hampered by a lack of timely data and the tediousness of the required computations.

In the mid 1950s two major breakthroughs dramatically changed the forecasting field. The first breakthrough was the introduction of a broad range of exponential smoothing techniques, which were empirically based and practically oriented. The greatest advantages of these methods were their simplicity in concept and their ease of computations. They were difficult to apply on a grand scale. If forecasts were needed for several thousand items, an enormous amount of work was required to maintain data files, make the required computations, and simply transcribe the results. The second breakthrough was the introduction of the computer, which made it possible to use not only exponential smoothing but also a host of other forecasting methods on a much more continuous basis [3].

A number of variations and extensions of smoothing methods have been suggested since the initial work of smoothing methods in 1950. These methods do not even require the user to specify the parameter values for the smoothing model; they are computed and updated automatically. Not long after smoothing methods began to become popular, decomposition methods experienced a rise in popularity [4,5]. In 1960s, more statistically sophisticated forecasting methods such as multiple regression and econometric models were used to quantify and test economic theory with qualitative data.

The Box-Jenkins methodology involving a systematic procedure for the analysis of time series that was sufficiently general to handle virtually all empirically observed time-series data patterns had been outlined. The method had become popular when several comparative studies of forecasting methods, done in 1970s, indicated that Box-Jenkins approach was as accurate as econometric approaches [6].

Variations of the autoregressive integrated moving average method (ARIMA) developed by Box and Jenkins began to emerge in the mid 1970s, overcome some of the computational difficulties of Box-Jenkins method and aid in the interpretation of the results [7].

Qualitative technological approaches such as Delphi approach and cross-impact matrices were in number of organizations by the early 1980s. They attempted to deal with long-term trends where the historical data and patterns necessary to apply statistical forecasting approaches simply were not available or did not apply.

In the recent decades, intelligent techniques involving artificial neural networks (ANNs) and fuzzy logic (FL) have been suggested for forecasting problems. ANNs are especially useful when there is a large number of historical events in which an input vector and a target variable are provided. They can approximate any universal function by an arbitrary grade of accuracy. Therefore it is possible to learn nonlinear correlations between input variables and target. They have several advantages over conventional statistical models: they handle noisy data better, do not have to fulfill any statistical assumptions, and are generally better at handling large amounts of data with many variables. FL resembles human reasoning, but uses estimated information and vagueness in a better way. The answers to real-world problems are rarely black or white, true or false, or start or stop. By using FL, knowledge can be expressed in a more natural way. FL employs linguistic terms which deal with casual relationships between the input and output variables. It becomes easier to manipulate and rig out solutions, particularly where the mathematical model is not explicitly known or is difficult to solve.

Besides hybrid versions blending ANN and FL with classical approaches have been developed for forecasting in all the fields.

Recently, a Dragonfly Optimization (DFO) has been presented for solving optimization problems in [8] and has been found to yield satisfactory results [9].

The focus of this paper is to develop an intelligent tool for forecasting the future demands for the prime energy fuels of coal, Lignite, Crude Oil and Natural Gas. The paper is organized with four sections containing introduction, proposed model (PM), results and discussions, and conclusion.

## **2. PROPOSED MODEL**

The focus of this paper is to develop an elegant forecasting model for predicting the future demands of the primary fuels of coal, lignite, crude oil and natural gas in the upcoming years with least input data. The ANNs are popular for pattern mapping or function approximation and frequently preferred for solving forecasting problems. Usually, the multi-layer feed forward networks, comprising of an input layer, an output layer and a hidden layer, each with a set of neurons, are used in forecasting models [10].

In the light of the fact that the per capita GDP and population establish a good relationship with the primary fuels' demand of any country in the long-term forecasting model, they are not readily available for the future years [10]. However, they can be forecasted for future years. The PM uses two multi-layer feed forward networks, each possessing an input layer, an output layer and a hidden layer, comprising of a set of neurons. The structure of PM is shown in Fig. 1. The first ANN predicts the population and per capita GDP for a given future year and the second one forecasts the future demand for coal, lignite, crude oil and natural gas by considering the output of the former as input. The input required for the first ANN model is just the year of forecast and the output of the second model is the forecasted demand of the four primary

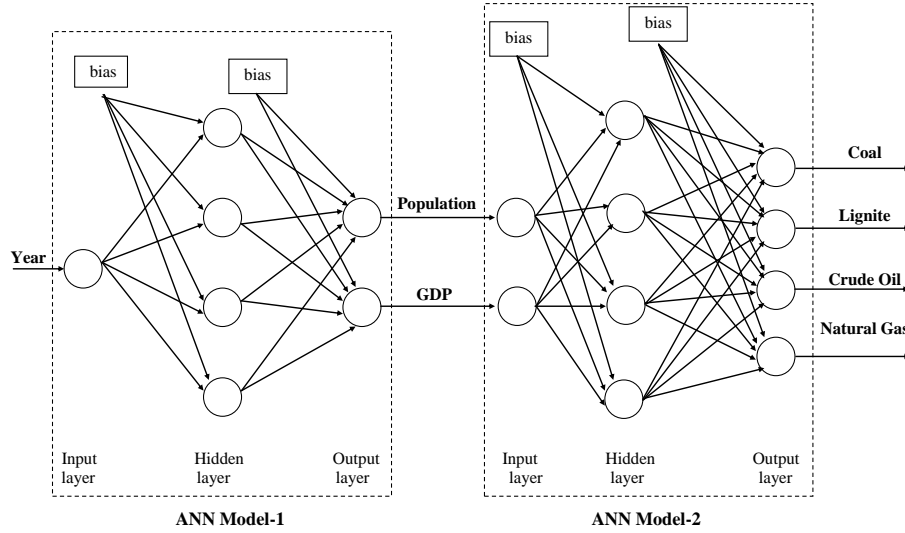


Fig. 1 ANN based forecasting model

fuels. The training and testing data set comprising the input ( $X$ ) and the target ( $T$ ) vectors for both ANN models are as follows.

$$\{X(1) \leftrightarrow T(1)\} = \{year \leftrightarrow Pop, GDP\} \quad (1)$$

$$\{X(2) \leftrightarrow T(2)\} = \{Pop, GDP \leftrightarrow F_1, F_2, F_3, F_4\} \quad (2)$$

Tangent hyperbolic and linear activation functions are chosen for the hidden layer neurons and output neurons respectively for both ANN models. The training of both ANN models can be performed by the DFO routine with an objective of minimizing the respective MSE values.

$$\text{Minimize } MSE(1) = \frac{1}{2N} \sum_{n=1}^N \sum_{i=1}^2 (O_i^n(1) - y_i^n(1))^2 \quad (3)$$

$$\text{Minimize } MSE(2) = \frac{1}{2N} \sum_{n=1}^N \sum_{i=1}^4 (O_i^n(2) - y_i^n(2))^2 \quad (4)$$

The DFO works directly on decision variables through nature-inspired operators and enables the searching in multi-dimensional solution space with an objective of reaching the desired goal. As there are no restrictions during search, it explores the entire solution space and offers the global best solution. It can be effectively used for training the ANNs of the proposed forecasting model with an objective of minimizing the MSE. Each dragonfly [8] in the population of both the ANN models is defined to represent the connection weights between input, hidden and output layers as

$$\text{dragonfly (1)} = [W_{ih}(1), W_{bias\_h}(1), W_{ho}(1), W_{bias\_o}(1)] \quad (5)$$

$$\text{dragonfly (2)} = [W_{ih}(2), W_{bias\_h}(2), W_{ho}(2), W_{bias\_o}(2)] \quad (6)$$

Both the ANN models are independently trained with an objective of minimizing the respective MSE of Eq. 3 or 4 using the appropriate data set and combined as depicted in Fig. 1 to form the PM.

For training both the ANN models, an initial population of dragonflies is obtained by generating random values to every individual in the population. The ANN, after setting the connection weights from the dragonfly, is actuated by

presenting the training data and the corresponding  $MSE$  is evaluated. and the exploration and exploitation phases, which represent social interaction of dragonflies in navigating and searching for foods and avoiding enemies, are performed for all the dragonflies in the swarm with a view of minimizing their  $MSEs$ . The iterative process is continued till convergence [8].

After training, the combined ANN models that represent PM are used to forecast the future demand of primary fuels.

### 3. RESULTS AND DISCUSSIONS

The PM is built using ANNs and trained by DFO. The historical data comprising India's primary fuels such as coal, lignite, crude oil and natural gas, the per capita GDP and the population data during the period of 1980-2012, are taken from [2,11,12]. The results of the PM are compared with that of the regression model (RM) [10] with a view to illustrate the superiority of the PM. The accuracy of the forecast is evaluated through the following mean absolute percent error (MAPE) [10].

$$MAPE = \frac{\sum_{i=1}^{no} \left| \frac{AFD_i - FFD_i}{AFD_i} \times 100 \right|}{no} \quad (7)$$

Where

$AFD_i$  represents the actual  $i$ -th fuel demand

$FFD_i$  denotes the forecasted  $i$ -th fuel demand

The results of the PM along with RM for the chosen yester years are given in Table 1. This table comprises the intermediate results in terms of per capita GDP and population, the actual and forecasted primary fuel demands and their calculated MAPEs. The net MAPE of the PM is 1.8508, which is lower than that of RM. The lowered MAPE validates the model and ensures the accuracy of the results for future years. The forecasted primary fuels along with intermediate predictions of per capita GDP and population during the years 2013- 2025 are presented in Tables 2 and 3 respectively for RM and PM. It is to be noted from the results that the demand for fuels given by the PM is higher than that of RM. While comparing the MAPE values of all the

methods, the MAPE of PM is the lowest, indicating a robust forecast, than that of the existing RM, thereby allowing the policy makers to allocate appropriate funds for procurement of the primary fuels for meeting the future demand.

**Table 1 Comparison of Results for Yesteryears**

Year	Method	Intermediate Results/Data		Primary Fuels Actual/Forecasted values				MAPE	
		Per Capita GDP	Population (Millions)	Coal	Lignite	Crude Oil	Natural Gas	RM	PM
1985	Actual	627.10	749.18	154.30	8.04	42.91	4.95		
	RM	635.80	759.13	159.84	8.34	40.46	5.94	8.2578	
	PM	630.25	751.39	154.64	8.19	41.68	5.12		2.0967
1990	Actual	883.04	833.93	214.06	14.07	51.77	12.77		
	RM	892.94	838.22	206.59	15.21	50.07	11.94	5.3438	
	PM	882.87	836.39	214.56	13.65	51.38	12.82		1.0909
1995	Actual	1150.19	934.23	273.42	22.15	58.74	18.09		
	RM	1126.23	927.13	279.50	21.55	58.13	21.13	5.6940	
	PM	1142.00	932.59	271.84	22.56	59.28	18.17		0.9476
2005	Actual	2190.27	1080.26	407.04	30.23	130.11	26.86		
	RM	2206.65	1086.40	431.00	29.59	129.50	30.94	5.9156	
	PM	2200.24	1079.12	417.85	29.23	135.87	27.58		3.2678
Average of MAPE (ANN Model-2)								6.3028	1.8508

**Table 2 Results of the RM**

Year	Intermediate Results		Forecasted Results			
	Per Capita GDP	Population (Millions)	Coal	Lignite	Crude Oil	Natural Gas
2013	5125.28	1209.75	604.78	49.09	234.28	44.61
2014	5690.23	1226.25	631.96	53.33	248.48	47.72
2015	6295.78	1248.56	659.72	58.00	262.88	51.05
2016	6938.99	1267.88	687.94	63.02	277.39	54.53
2017	7615.92	1287.75	716.40	68.33	291.90	58.14
2018	8321.45	1309.94	744.90	73.85	306.34	61.83
2019	9049.44	1330.69	773.23	79.45	320.56	65.51
2020	9792.53	1351.38	801.09	84.99	334.46	69.06
2021	10542.00	1372.50	828.10	90.31	347.90	72.37
2022	11287.82	1390.13	853.97	95.18	360.76	75.37
2023	12018.59	1406.69	878.20	99.34	372.90	77.81
2024	12721.35	1418.81	900.35	102.48	384.16	79.56
2025	13381.58	1431.25	919.88	104.25	394.40	80.41

**Table 3 Results of the PM**

Year	Intermediate Results		Forecasted Results			
	Per Capita GDP	Population (Millions)	Coal	Lignite	Crude Oil	Natural Gas
2013	5621.78	1209.26	593.38	43.25	228.87	40.72
2014	6097.08	1222.35	616.31	44.91	241.26	42.18
2015	6592.88	1235.20	639.43	46.60	253.80	43.63
2016	7109.15	1247.82	662.66	48.32	266.44	45.06
2017	7645.91	1260.26	685.95	50.08	279.14	46.47
2018	8203.15	1272.55	709.24	51.87	291.83	47.85
2019	8780.88	1284.75	732.46	53.70	304.48	49.21
2020	9379.09	1296.89	755.53	55.57	317.02	50.53
2021	9997.79	1309.04	778.39	57.48	329.39	51.82
2022	10636.97	1321.26	800.96	59.42	341.54	53.06
2023	11296.63	1333.60	823.15	61.41	353.39	54.26
2024	11976.78	1346.13	844.89	63.44	364.88	55.40
2025	12677.41	1358.92	866.08	65.51	375.94	56.49

#### **4. CONCLUSIONS**

DFO is a population based optimization algorithm that simulates the static and dynamic swarming behaviors of dragonflies. A DFO based ANN model for predicting India's primary fuel demand has been suggested. Two ANNs have been trained through DFO. Primary fuel demands during the years 1990-2012 have been considered for training and validating the model. The proposed model requires an input, the year of the forecast, and predicts the primary fuels' demand. The forecasts up to the year 2025 have been compared with that of the RM and illustrated that the PM is more accurate than the existing RM. The suggested method can be modified to include the other fuels.

#### **5. ACKNOWLEDGMENTS**

The authors gratefully acknowledge the authorities of Pondicherry Engineering College and Annamalai University for the facilities offered to carry out this work.

#### **6. REFERENCES**

- [1] Assareh.E, Behrang.M.A, Assareh.R, Hedayat.N. (2011). Global Electricity Consumption Estimation Using Particle Swarm Optimization (PSO), *World Academy of Science, Engineering and Technology*, 55: 200-204.
- [2] Energy Statistics available at [www.mospi.gov.in](http://www.mospi.gov.in)
- [3] Brown. R.G. (1963). *Smoothing, Forecasting and Predictions*, Prentice-Hall, Englewood Cliffs, NJ.
- [4] Shiskin. J. (1957). *Electronic computers and business indicators*, National bureau of economic research, occasional paper no. 56.
- [5] Shiskin. J. (1961). *Tests and revisions of bureau of the census methods of seasonal adjustments*, Bureau of the Census, Technical paper no. 5
- [6] Armstrong. J.S. (1978). *Forecasting with econometric methods: Folklore versus Facts*, *Journal of Business*, S1: 549-600.
- [7] Box. G.E.P and Jenkins.G.M. (1976). *Time series analysis: forecasting and control*, rev. ed., Holden-Day, San Francisco.
- [8] Seyedali Mirjalili. (2015). *Dragonfly algorithm: A new meta-heuristic optimization technique for solving single-objective, discrete and multi-objective problems*, *Neural Comput and Applic.* DOI. 10.1007/s00521-015-1920-1.
- [9] Rakoth Kandan Sambandam, Sasikala Jayaraman. (2016). *Self-Adaptive Dragonfly Based Optimal Thresholding for Multilevel Segmentation of Digital Images*, *Journal of King Saud University - Computer and Information Sciences*, DOI: <http://dx.doi.org/10.1016/j.jksuci.2016.11.002>.
- [10] Jayaraman Kumraran, Govindasamy Ravi. (2015). *Long-term Sector-wise Electrical Energy Forecasting Using Artificial Neural Network and Biogeography-based Optimization*, *Electric Power Components and Systems*, 43(11): <http://dx.doi.org/10.1080/15325008.2015.1028115>.
- [11] Data for per capita GDP available at [http://www.indexmundi.com/india/gdp\\_per\\_capita\\_\(ppt\).html](http://www.indexmundi.com/india/gdp_per_capita_(ppt).html)
- [12] Data for population available at <http://www.populstat.info/Asia/indiac.htm>