



## Forecasting the Ranks of Sites suitable for Power Plant Installations

Kalyani Sambhoo(Salla)<sup>1</sup>, Sanjay Kadam<sup>2</sup>

<sup>1</sup>Department of Computer Science, Modern College, Pune -5, Maharashtra, India,  
Email: profkalyanis@gmail.com

<sup>2</sup>ECIP, Centre for Development of Advanced Computing, Pune – 7, Maharashtra, India, Email: sskadam@cdac.in

### ABSTRACT

An increase in the number of decision parameters used for ranking of sites for a power plant installation using the soft computing techniques leads to complex formulations that are computationally expensive[41]. Amongst a large number of decision parameters, if some of the parameters do not significantly contribute towards the ranking process, then we need not consider these for decision making. Moreover, it is very tedious to form fuzzy sets for all the 87 decision parameters from several environmental experts, which serve as inputs to certain soft computing techniques used for ranking. The decision parameters comprise of some parameters used to describe air quality, water quality, land suitability, socioeconomic and ecological suitability. We have made an attempt to reduce the number of input decision parameters so that the processing is computationally fast without significantly degrading the accuracy of the end results. We have also attempted to predict futuristic values of some of the relevant parameters to infer site suitability and/or ranking, futuristically (subsequent five years) which can act as a planning tool.

### Keywords:

Prediction, Decision support, Fuzzy Soft Set, Linguistic Fuzzy Soft set, Parameter reduction, Futuristic ranking, Planning tool

### Academic Discipline

Computer Science

### SUBJECT CLASSIFICATION

Soft Computing, Decision Support System, Prediction, Parameter Reduction

### TYPE (METHOD/APPROACH)

Fuzzy Soft Set, Linguistic Fuzzy Soft Set, Parameter Reduction, Futuristic Ranking, Planning Tool

## 1. INTRODUCTION

Decision making has been a challenging task since the inception of human race. Taking an accurate and acceptable decision is even more challenging. For example, deciding a place for setting up a new airport in city, urbanization of a landscape (e.g., Lavasa, Pune, India), selecting a site for setting up an industry at a certain location, building of urban structures or planning parks in a city all demand major decision making efforts and require several environmental clearances from the respective government agencies. One of the major decision making that has an impact on the environment and overall ecosystem of a place is about setting up of major industry such as a power plant at a particular location. This would involve whether to allow this industry to be developed or suggest appropriate alternatives[35] with suitability index or ranks.

Globally, energy needs are on an increase due to the rapid industrialization and urbanization, this calls for the installation of power plants in the country. Traditionally, the location of a power plant is decided and then Environment Impact Assessment (EIA) [35] is carried out. In our work, we consider the relevant parameters which depict air quality, water quality, land suitability, ecological quality, socioeconomic aspect and seismic zone information. With increasing number of parameters, it becomes difficult to compare sites and identify their suitability for power plant installation.

The environmental impact of setting power plants for generation of electricity may pose several environmental concerns[35]. The storage and usage of water by these power plants may lead to water shortage and may lead to drought, resulting in disturbance of the overall ecosystem. On the other hand the thermal power plants typically release CO<sub>2</sub>, SO<sub>2</sub>, Ash, Particulate matter and Nitrogen oxides that harm public health. Similarly, the radioactive waste generated by nuclear power plants can pose a threat to the overall ecosystem. Hence, if a power plant industry is set in an already polluted location or in a location that is densely populated, it may further worsen the environment or cause harm to the public health. Therefore, in order to reduce the effect of pollution caused due to power plant installations, finding site suitability and ranking of feasible sites is of outmost importance. The model developed in this work is useful as a planning tool for the environmentalist.

This research work also deals with parameter reduction, which is about finding a subset of important parameters (from a larger set) that can be used for calculating site suitability index and subsequently ranking of a given set of power plant sites without sacrificing the accuracy of the methods used. That is, instead of using a complete set of parameters, we use only a subset of these parameters to get the same results. We also predict futuristic values of some relevant parameters to infer site suitability (and/or ranking), futuristically. We have used past ten-years data to predict the futuristic values of these parameters based on non-linear regression analysis. The parameter reduction techniques help in addressing limitations of certain soft computing techniques, where the absence of Fuzzy sets[29] for all of the attributes together with lengthy computation time make the fuzzy formalism time-consuming and tedious for decision making.



Parameter reduction has been attempted by several researchers in the past. A modified feature selection method[35] based on fuzzy rough set theory and differential evolution was proposed to reduce parameters out of the original 99 features for Pentagon dataset and order the relevant features[10, 33]. A new efficient normal parameter reduction algorithm using soft sets was proposed, which was based on oriented parameter sum[43]. A fuzzy mutual information measure based algorithm for feature selection within imprecise data and vague problems was proposed by Grande et al.[9]. This algorithm was applied on the dataset of chemical analysis of wines grown. The data comprised of 13 features and 3 class values. Swarm intelligence has been applied for parameter reduction on real valued data sets by Dorigo et al.[20]. The capability of ants to identify the shortest path to the food source has been explored for identifying the most relevant parameters in the dataset.

However, none of these methods have been applied for parameter reduction for site suitability which also embeds some linguistic information such as accessibility to community services, willingness for resettlement, availability of adequate infrastructure, response to public involvement programs etc. We believe that research into newer computing methods and techniques could alleviate some of the above-mentioned problems in construction of Fuzzy sets with increasing number of parameters. Although, the techniques like Ant colony optimization, Principal component analysis and Fuzzy soft sets have been used in the past for parameter reduction, it has not been applied to optimize parameters considered to rank industrial sites. All the techniques except fuzzy soft sets are supervised parameter reduction techniques. Pre-defined class labels have been assigned by the existing BEES method[4]. In our opinion, the use of soft computing methods like Fuzzy soft sets[6, 12] would reduce the subjectivity aspect and also address the imprecise nature of the governing parameters and impact indicators.

In the prediction part several researchers have proposed prediction of futuristic values based on past data. The prediction of impacts is a systematic way of anticipating future values on the basis of the past data. A genetic algorithm(GA) was applied in the refinement of input data selection for the purpose of air temperature prediction [42]. This GA based approach to determine the duration and resolution of prior input data resulted in more accurate ANN models than the existing ones for predicting air temperature. A fuzzy time series prediction of hourly particulate matter PM10 parameter, which is a top priority pollutant considering public health was done by Sfetos et al.[33, 34]. Artificial neural networks were applied for prediction of dissolved oxygen level in water, which is the best indicator of the health of a water ecosystem[2].

However, the above approaches were applied to specific problem domains like linking air pollution to chronic illness, air temperature prediction, prediction of hourly particulate matter, dissolved oxygen level in water etc. In our work, we use the predicted values to find the futuristic ranks of sites suitable for power plant installation. Soft computing techniques represent a significant change in both the approach and outcome of environmental evaluations in comparison to the earlier work done[18]. Our proposed techniques, namely, Ant colony optimization[17], Principal components, Latent semantic analysis, Particle swarm optimization and Fuzzy soft sets[40] were applied for attribute reduction of data of sites for consideration for an upcoming Power Plant Installation. As far as our knowledge is concerned, no such attempt has been made for parameter reduction and prediction of certain environmental parameters for site selection of an upcoming power plant.

The paper is organized as follows: Section 1, begins with an introduction of the problem domain and various techniques used for parameter reduction and prediction. It briefly discusses the existing work done and the scope of the existing techniques. Section 2, describes the Methodology, where an overview of the methods used is discussed, Section 3, presents a Case study, Section 4, presents Results and Section 5, presents Conclusions.

## 2. METHODOLOGY

### A. Attribute Reduction

Data reduction could be applied to the given data set using the following strategies[12]:

- Attribute subset selection, where irrelevant, weakly relevant or redundant attributes or dimensions may be detected and removed
- Dimensionality reduction, where encoding mechanisms are used to reduce the size of the data

The generation procedure implements a search method that can generate subsets of features. It may start with no features, all features, a selected feature set or some random feature subset. Those methods that start with an initial subset usually select these features heuristically beforehand. Features can be added (forward selection) or removed (backward elimination) iteratively. An alternative selection strategy is to select instances and examine differences in their features[12]. The evaluation function calculates the suitability of a feature subset produced by the generation procedure and compares this with the previous best feature and replacing it if found to be a better one[17]. Attribute reduction techniques are useful in order to improve the efficiency of algorithms and to analyze the results in a better way.

## A.1 Ant Colony Optimization (ACO)

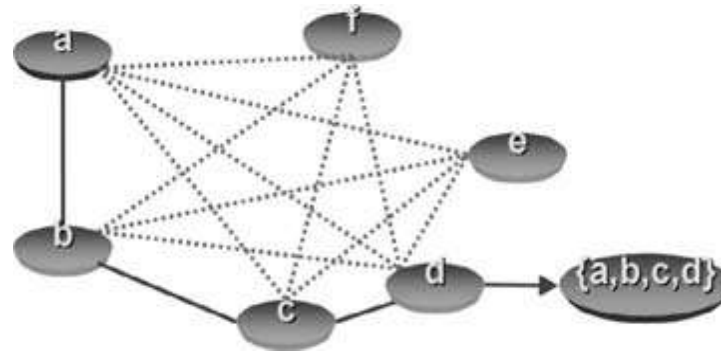


Fig 1: ACO problem representation [17]

ACO represents the problem as a graph where nodes represent features, with the edges between them denoting the choice of the next feature. The optimal feature selection is an ant traversal through the graph where a minimum number of nodes are visited that satisfies the traversal stopping criterion. In figure 1, the ant is currently at node 'a' and has a choice of which feature to add next to its path (represented with dotted lines). It chooses feature 'b' next based on the transition rule, then 'c' and then 'd'. On arrival at 'd', the current subset {a, b, c, d} is determined to satisfy the traversal stopping criterion (e.g., a suitably high classification accuracy has been achieved with this subset). The ant terminates its traversal and outputs this feature subset as a candidate for attribute reduction.

The ACO process begins by generating a number of ants,  $k$ ; the ants are then placed randomly on the graph (i.e., for each ant)[13]. Alternatively, the number of ants to place on the graph may be set equal to the number of features within the data; each ant starts path construction at a different feature[13]. From the predefined initial positions, they traverse edges probabilistically until a traversal stopping criterion is satisfied. The resulting subsets are gathered and then evaluated. If an optimal subset has been found or the algorithm has executed a certain number of times, then the process halts and outputs the best feature subset encountered. If neither condition holds, then the pheromone is updated, a new set of ants are created and the process iterates once more. The technique also tries to learn the relation between attributes and how they affect the output class.

ACO algorithm works as follows[7]:

Step 1: Description of the problem as a graph with a set of nodes and edges between nodes

Step 2: Heuristic desirability of edges. A suitable heuristic measure of the "goodness" of paths from one node to every other connected node in the graph

Step 3: Construction of feasible solutions

Step 4: Pheromone updating rule. A suitable method of updating the pheromone levels on edges with a corresponding evaporation rule. Typical methods involve selecting the 'n' best ants and updating the paths they chose[16]

Step 5: Aggregation of results which is a subset of the original attributes (nodes). Each ant in the artificial colony maintains a history i.e. the path it has chosen so far in the construction of a solution.

A total number of 87 attributes were considered by the methods but the computations became more tedious with increasing number of sub-attributes and dynamic nature of environmental parameters.

## A.2 Particle Swarm Optimization

Particle swarm optimization (PSO) is a population-based stochastic optimization technique. In PSO the system is initialized with a population of random solutions, called particles. Optima are searched by updating generations, with particles moving through the parameter space towards the current local and global optimum particles[17]. At each time step the velocities of all particles are changed depending on the current optima. Although there are similarities with GAs, PSO systems tend to require fewer design choices, such as the choice of evolutionary operators [17, 39].

For this application the particles will represent potential membership function definitions defined by sets of parameters. The initial population of particles could be generated by random parameter deviations from the original membership functions. Extra constraints will need to be enforced in order to restrict search to meaningful fuzzifications. Particles are rated according to the fuzzy-rough dependency degree in order to provide a measure of fitness. From this, the local and global optima can be determined and used to adjust particle velocities.



### A.3 Principal Component Analysis

Principal Component Analysis(PCA) is a dimensionality reduction tool in common use, perhaps due to its conceptual simplicity and the existence of relatively efficient algorithms for its computation[27]. PCA transforms the original features of a dataset to a reduced set with parameters which are uncorrelated, also termed as principal components. The method works on the hypothesis that a large feature variance corresponds to useful information, with small variance equating to information that is less useful. Data are transformed in such a way as to allow the removal of those transformed features with small variance. This is achieved by finding the eigenvectors of the covariance matrix of data points and constructing a transformation matrix from the ordered eigenvectors which transforms the original data by matrix multiplication.

### A.4 Latent Semantic Analysis(LSA)

Unlike other models, LSA treats attributes as if they are not independent of each other; it attempts to automatically derive and model interrelationships between them[14].

Similarity is computed in the following way[14]:

1. Choose  $k$  (not greater than the number of terms or documents)
2. Add the weighted vectors for each term—multiply each vector by term weight—sum each element separately
3. Repeat for query or second document
4. Compute inner product—multiply corresponding elements and add

### A.5 Fuzzy soft sets

For a soft set  $(F, E)$ ,  $E = \{e_1, e_2, \dots, e_m\}$ , if there exists a subset  $A = \{e'_1, e'_2, \dots, e'_p\} \subset E$  satisfying

$$f_A(S_1) = f_A(S_2) = \dots = f_A(S_n) \quad (1)$$

then  $A$  is dispensable, otherwise,  $A$  is indispensable.  $B \subset E$  is a normal parameter reduction of  $E$ , if the two conditions as follows are satisfied[43]

(1)  $B$  is indispensable

$$(2) f_{E-B}(S_1) = f_{E-B}(S_2) = \dots = f_{E-B}(S_n) \quad (2)$$

Suppose  $A = \{e'_1, e'_2, \dots, e'_p\} \subset E$ . If  $B \subset E$  is defined as a normal parameter reduction of  $E$  then

$$f_{E-B}(S_1) = f_{E-B}(S_2) = \dots = f_{E-B}(S_n) \quad (3)$$

In other words, if  $A = E - B$  can be reduced, then

$$f_A(S_1) = f_A(S_2) = \dots = f_A(S_n) \quad (4)$$

Therefore, the following oriented parameter sum must be satisfied

$$S'_{11} + S'_{12} + \dots + S'_{1p} = q$$

$$S'_{21} + S'_{22} + \dots + S'_{2p} = q$$

$$S'_{n1} + S'_{n2} + \dots + S'_{np} = q$$

The set ' $B$ ' acquired in the above is the reduced attribute set

## 3. CASE STUDY

### A. Parameter Reduction

The parameter reduction techniques work on identifying the most relevant features in comparison to the remaining and form a subset of such attributes with respect to the original attributes. This has proved to be useful in order to address the limitations of certain soft computing techniques used for ranking in the case of increasing number of parameters and absence of Fuzzy sets for all of the attributes. With increasing number of parameters the computations became tedious. Also to overcome limitations of certain soft computing techniques there was a need for parameter reduction techniques. We have used predefined class labels to reduce the number of attributes in such a way that the accuracy of classification is maintained within some predefined limits. The existing method namely, BEES was used to assign weightages for each of the sub-attributes.

The guidelines provided by Central Pollution Control Board(CPCB)[5], India were used to verify the appropriate range of values for the sub-attributes. Based on the score obtained by BEES method and guidelines provided by CPCB, the sites were classified as 'V.Good', 'Good', 'Fair' and 'Poor'. In our work, ACO with WEKA random forest, PSO, PCA and Fuzzy soft sets techniques have been applied for parameter reduction. The procedure to improve the accuracy of the multi-class classification process can be summarized as follows:



## A.1 Ant Colony Optimization

Real ants are able to find the shortest path between their nest and the food sources because of the chemical substance called as pheromone they deposit on their way[1, 28]. The pheromone evaporates over time. The shortest paths will contain more pheromone(as the rate of pheromone deposition is relatively greater than the rate of evaporation for such paths) and will consequently attract a greater number of ants in comparison to the longest paths(which would be taken by only few ants). ACO is a class of algorithms which was initially proposed by Dorigo[34, 37].

The main underlying idea, loosely inspired by the behavior of real ants, is that of a parallel solution search using several concurrent process or ants. These processes or ants attempt to find solutions to a given problem based on the local problem data and on dynamic memory structure containing information(eg., amount of pheromone depositions) on the quality of previously obtained results. The collective behavior emerging from the interaction of the different search processes has proved effective in solving combinatorial optimization problems.

Each and every ant in the artificial colony maintains a memory which stores the paths it has chosen so far in constructing the solution[16]. This information can be used in the construction of the solution.

### ACO algorithm works as follows[16, 17]:

Step 1: Description of the problem as a graph with a set of nodes and edges between nodes

Step 2: Heuristic desirability ( $\eta$ ) of edges. A suitable heuristic measure of the "goodness" of paths from one node to every other connected node in the graph

Step 3: Construction of feasible solutions

Step 4: Pheromone updating rule. A suitable method of updating the pheromone levels on edges with a corresponding evaporation rule. Typical methods involve selecting the  $n$  best ants and updating the paths they chose

Step 5: Aggregation of results which is a subset of the original attributes (nodes) Each ant in the artificial colony maintains a history i.e. the path it has chosen so far in the construction of a solution.

This history can be used in the evaluation of the newly created solution and may also contribute to the decision process at each stage of solution construction. Two types of information are available to ants during their graph traversal, local and global, named as ' $\beta$ ' and ' $\alpha$ ' respectively. A total number of 87 attributes were considered by the ANN methods but the computations became more tedious with increasing number of attributes. For Fuzzy soft sets based method we made an attempt to compare the results obtained with the ACO based feature selection technique and without feature selection. ACO works by describing the problem as a graph with a set of nodes and edges between nodes.

The nodes represent the 87 sub-attributes, which describe the main attributes like Air, Water, Land, Socioeconomic and Ecological quality for evaluating sites. The edges represent the choice of the next feature. Ant traversal over edges and selection of nodes is mapped to identifying relevant features, which have significant impact on assignment of the output class. An ant in the colony maintains a history, which is the path it has chosen in constructing the solution or identifying a subset of features. The process begins by generating a number of ants and places these ants randomly on the graph comprising 87 nodes representing the sub-attributes for assessment of environmental quality. Local heuristic ' $\beta$ ' is obtained through problem specific measure[15]. Global information represented as ' $\alpha$ ' is available to ants through the deposition of artificial pheromone on the graph edges[10]. The outcome of ACO were the following sub-attributes: CO, SO<sub>x</sub>, PM<sub>2.5</sub>, DO, BOD, Sodium absorption and sulphate. The values of the parameters giving desired classification accuracy were  $\alpha = 1$ ,  $\beta = 2$ , number of ants = 50, and classification accuracy = 85%.

Table 1 and Table 2, depict the reduced parameter set obtained with ACO technique. The ranking process with reduced parameter set was repeated with Fuzzy soft sets method[3] as shown in Table 3, 4 and Table 5.

**Table 1: Fuzzy soft set for Air parameters**

Air	CO	PM10
Jaitapur	1	1
Kundankullam	1	0.2





**Table 2: Fuzzy soft set for Water parameters**

Water	DO	BOD	Sodium Absorption	Sulphate	pH
Jaitapur	1	0.8	1	0	1
Kundankullam	1	0.7	0.9	0	1

**Table 3: Resultant Fuzzy soft set of Air and Water parameters**

U	e <sub>11</sub>	e <sub>13</sub>	e <sub>14</sub>	e <sub>21</sub>	e <sub>22</sub>	e <sub>23</sub>	e <sub>24</sub>
Jaitapur	1	1	1	1	1	1	1
Kundankullam	1	1	1	1	0.7	0.9	1

**Table 4: Comparison Table**

U	Jaitapur	Kundankullam
Jaitapur	7	7
Kundankullam	5	7

**Table 5: Score and Rank table**

U	Row sum	Column sum	Score(S <sub>i</sub> )	Rank
Jaitapur	14	12	2	1
Kundankullam	12	14	-2	2

Resultant fuzzy soft sets were obtained by computing the max membership between the membership grades of selected columns(attributes)[44].

### A.2 Attribute Reduction with Fuzzy Soft sets

Fuzzy soft sets have been applied to obtain a reduced parameter set which is indispensable[8, 40]. After applying the algorithm, the final reduced parameter set turns out to be {NO<sub>x</sub>, SO<sub>2</sub>, PM<sub>10</sub>, PM<sub>2.5</sub>, DO, ThickPOP, Effects on Economy, Community Services, chances of fire}. Fuzzy soft set based ranking technique was applied to repeat ranking with the reduced parameter set.

**Table 6: Comparison table**

ThermalPP	Badarpur	Amreli	Bharuch	Tuticorin	Khammam	Row sum	Column sum	Score
Badarpur	5	2	2	2	2	13	21	-8
Amreli	4	5	4	4	4	25	19	6
Bharuch	4	4	5	4	4	21	19	3
Tutocorin	4	4	4	5	4	21	19	3
Khammam	4	4	4	4	5	21	19	3

Here, Amreli is ranked as first, followed by Bharuch, Tuticorin, Khammam and lastly Badarpur. After obtaining the predicted values for year 2016, ranking was again repeated with reduced parameter set and Fuzzy soft set approach as follows.

**Table 7: Ranking for the year 2016**

Site	Row Sum	Column Sum	Score (Si)	Rank
Site1(Badarpur)	12	16	-4	3
Site 2(Bharuch)	16	14	2	1
Site 3(Tuticorin)	17	14	3	2

Therefore the reduced parameter set gives the same results as complete parameter set.

### A. 3 Latent Semantic Analysis(LSA) with Ranker

LSA treats attributes as if they are not independent of each other; it attempts to automatically derive and model interrelationships between them. Out of the 87 attributes some of the higher ranked attributes by this technique were: CO, Lead, NOx, Ozone, SO<sub>2</sub>, PM<sub>10</sub>, pH, Total Coliform, DO, BOD, Ammonia in water, Electrical conductivity, sodium absorption, boron, re-settlement, proximity to resources, unloading space, transportation facility, away from sanctuary, infrastructure available, away from thick population and follows seacoast regulation. Table 8, shows the attribute reduction techniques applied.

**Table 8: Attribute Reduction techniques applied**

Sr. No	Attribute Evaluator	Search Method	Number of Attributes
1.	FuzzyRoughsubsetEval	AntSearch	8
2.	LatentSemanticAnalysis	Ranker	Ranked Attributes
3.	Principal Components	Ranker	18
4.	FuzzyRoughSubsetEval	PSO	8
5.	Fuzzy Soft set based approach		Subset pairs

**Table 9: Ranking with Principal Component optimization(PCO)**

Sr. No	Site Name	Ranking with PCO	Ranking without PCO
1.	Jaitapur	1	1
2.	Kundallam	2	2

LSA was applied for 5 Thermal Power plants. The attributes were ranked in the order of higher importance in the order: CO, Lead, NOx, Ozone, SO<sub>2</sub>, PM<sub>10</sub>, PM<sub>2.5</sub>, pH, Total coliform, DO, BOD, Ammonia in water, Electrical conductivity, sodium absorption, boron, re-settlement, suitability of land, proximity to market, unloading space, transport facility, away from sanctuary, suitable infrastructure, away from Thick population and follows seacoast regulations. The application of Fuzzy soft sets is as follows for given five Thermal Power plants:

**Table 10: Fuzzy soft set for Air quality parameters of Thermal Power plant**

Air	CO=a1	SO <sub>2</sub> =a2	PM <sub>10</sub> =a3	PM <sub>2.5</sub> =a4	Lead=a5	NOx=a6	Ozone=a7
Badarpur	1	0.5	0	0	0	1	0
Amreli	0.4	0.3	0	0	1	1	0
Bharuch	0.5	0.4	1	0.8	0.6	1	0
Tutocorin	0.9	0.4	0.3	0.3	0	0.2	0
Khammam	1	0.4	1	1	0	1	0



**Table 11: Fuzzy soft set for Water quality parameters of Thermal Power plant**

Water	pH =b1	TotColi =b2	DO= b3	BOD= b4	Ammonia =b5	ElecConduct =b6	SodiumAbs =b7	Boron =b8
Badarpur	0	1	0.6	1	0	1	1	1
Amreli	0	1	0.6	1	1	1	1	1
Bharuch	0	1	0.4	1	1	1	1	1
Tutocorin	0	1	0.5	1	0	0	0	0
Khammam	0	1	1	0	0	0	0	0

**Table 12: Fuzzy soft set for Socioeconomic quality parameters of Thermal Power plant**

SocioEco	Reset=c1	Suit =c2	Prox =c3	Unloading =c4	Transport =c5	Sanct =c6	Infra =c7	Pop =c8	Sea Regulation =c9
Badarpur	1	0	0	0	1	1	1	1	1
Amreli	1	1	0	0	1	1	1	0	1
Bharuch	1	1	1	1	1	1	1	1	1
Tutocorin	1	1	1	0	1	1	1	1	1
Khammam	1	1	0	0	0	1	0	1	1

**Table 13: Resultant Fuzzy soft set for Air, Water, Land parameters of Thermal Power plant**

A/W/L/S	d <sub>11</sub>	d <sub>12</sub>	d <sub>23</sub>	d <sub>33</sub>	d <sub>34</sub>	d <sub>45</sub>	d <sub>66</sub>	d <sub>78</sub>
Badarpur	1	1	0.6	0.6	1	0	1	1
Amreli	0.4	1	0.6	0.6	1	1	1	1
Bharuch	0.5	1	0.4	1	1	1	1	1
Tutocorin	0.9	1	0.5	0.5	1	0.3	0.2	0
Khammam	1	1	1	1	1	1	1	0

**Table 14: Resultant Fuzzy soft set with all parameters of Thermal Power Plant**

A/W/L/S	d <sub>11</sub> ^ C <sub>1</sub>	d <sub>12</sub> ^ C <sub>2</sub>	d <sub>23</sub> ^ C <sub>3</sub>	d <sub>33</sub> ^ C <sub>3</sub>	d <sub>34</sub> ^ C <sub>4</sub>	d <sub>45</sub> ^ C <sub>5</sub>	d <sub>66</sub> ^ C <sub>6</sub>	d <sub>78</sub> ^ C <sub>7</sub>
Badarpur	1	0	0	0	0	0	1	1
Amreli	0.4	1	0	0	0	1	1	1
Bharuch	0.5	1	0.4	1	1	1	1	1
Tutocorin	0.9	1	0.5	0.5	0	0.3	0.2	1
Khammam	1	1	0	0	0	0	1	0

**Table 15: Comparison Table**

	Badarpur	Amreli	Bharuch	Tutocorin	Khammam
Badarpur	8	6	3	4	7





Amreli	7	8	4	5	7
Bharuch	7	8	8	7	7
Tutocorin	6	6	4	8	6
Khammam	7	6	3	3	8

**Table 16: Score and Rank for Thermal Power plants**

	Row sum	Col sum	Score(S <sub>i</sub> )	Rank
Badarpur	28	35	-7	4
Amreli	31	34	-3	3
Bharuch	37	22	15	1
Tutocorin	30	27	3	2
Khammam	27	35	-8	5

The above computations were performed for Nuclear and Oil based power plants.

Results were as listed in the following table

**Table 17: Ranks after application of parameter reduction by LSA**

Site No/Name	Rank	Rank with ACO
<b>Thermal</b>		
Site 3 (Bharuch)	1	1
Site 4 (Tutocorin)	2	2
Site 2 (Amreli)	3	3
Site 1 (Badarpur)	4	4
Site 5 (Khammam)	5	5
<b>Hydroelectric</b>		
Site 3 (Sikkim)	1	1
Site 1 (Arunachal Pradesh)	2	2
Site 2 (Himachal Pradesh)	3	3
<b>Nuclear</b>		
Site 1 (Jaitapur)	1	1



Site 2 (Kundallam)	2	2
<b>Gas</b>		
Site 2 (Shadol)	1	1
Site 1 (Patan)	2	2

The results were also tested with ANN for reduced parameters obtained with ACO, PSO and LSA.

## B. Prediction

There is a growing evidence linking air pollution to acute and chronic illnesses amongst all age groups[25]. Therefore prediction of important air pollutant concentrations as well as quantification becomes very important. The effects of air pollution on health are very complex, as there are many different sources and their individual effects vary from one to the other. Air pollutants that are inhaled affect human health severely by way of damaging the lungs and respiratory system.

Last ten years data was acquired from Central pollution control board, India. Only the important parameters like SO<sub>x</sub>, NO<sub>2</sub> and PM<sub>10</sub> depicting Air Quality and DO, pH, BOD and Temperature depicting water quality were considered for prediction.

### B.1 Time series prediction

The techniques used for prediction are: Time series prediction and Multi layer perceptron ANN. There are two standard approaches to forecast a time series. The first one uses Trend analysis.

The second approach consists in the so called decomposition of a time series where each element  $x_t$  of the time series is decomposed into the following components

$$x_t = Tr_t + St + Ct + Et$$

where  $Tr_t$ ,  $St$ ,  $Ct$ ,  $Et$  are the t-th trend, seasonal, cyclic and error components, respectively.

**Table 18: Badarpur Thermal PP Prediction of water quality, station code 1375, Yamuna river, Okhala Barrage , Badarpur Thermal PP**

Year	Temp	DO	pH	BOD	Nitrate	FC
2008	22.6	0.2	7.2	2.9	0.1	40
2009	24.2	0.1	7.4	1.7	0.8	46
2010	23.3	0.3	7.6	3	1.86	86
2011	23.7	0.5	7.8	3.3	2.16	78
2012	23.8	2.0	7.9	4.6	2.32	76
2013	24.05	1.82	7.89	5.73	2.4	75
2014	24.29	2.22	7.94	6.52	2.52	74.33
2015	24.53	2.62	7.99	7.31	2.6	64
2016	24.77	3.02	8.04	8.10	2.70	54.32
2017	25.01	3.42	8.09	8.89	2.98	44.33

### Prediction of pH value for station code 1375, Yamuna river, Okhala Barrage, Badarpur Thermal PP

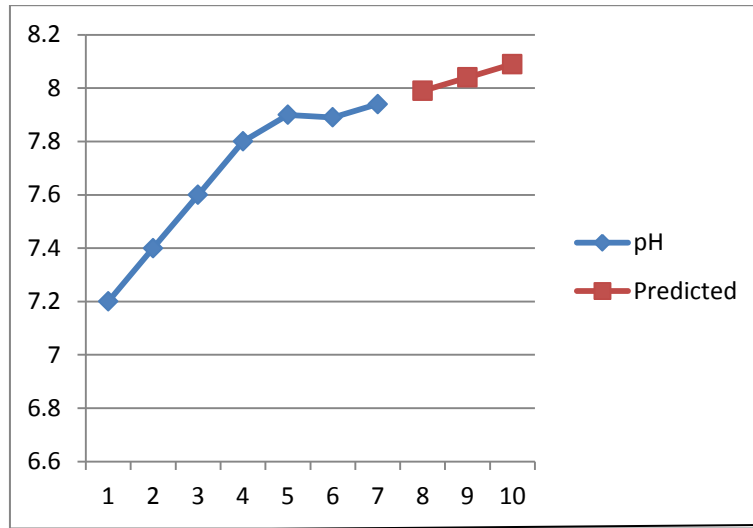


Fig 2: Prediction of pH values

As shown in Figure 2, a steep rise in pH value indicates hazardous sign for under water life. Fish can survive in a pH of 6.5 and upto 9. Guidelines for pre-treatment of waste water to curb this hazard could be issued. Similarly prediction of Air Quality parameters at certain locations was carried out.

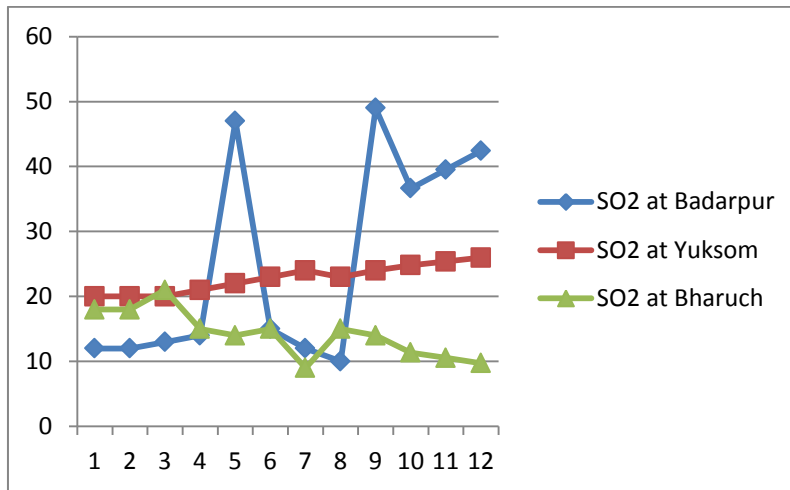


Fig 3: Prediction of sulphur dioxide in three main locations

Figure 3, shows the SO2 levels at Badarpur, Yuksom and Bharuch. If SO2 exceeds the maximum range of 80  $\mu\text{g}/\text{m}^3$  washing of coal or installation of precipitators can be suggested. Pollution from power stations can be reduced by several techniques. Significant improvements in emission levels can be achieved through cleaner power stations. Treatment can be before, during or after combustion. Fuel switching: replacement or blending of high sulphur coal with low-sulphur coal, or switching to oil or gas or renewable sources can reduce SO2 levels.

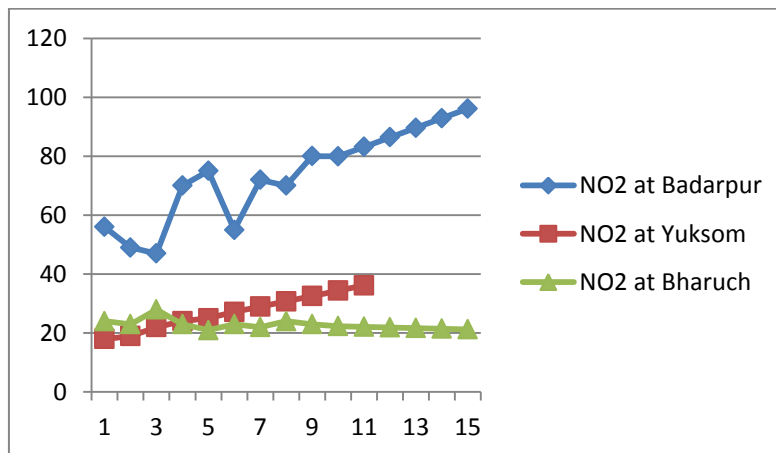


Fig 4: Prediction of Nitrogen dioxide in three main locations

If NO<sub>2</sub> exceeds the maximum limit of 80 µg/m<sup>3</sup> certain measures like installation of ESP can be suggested. Lime Injection Multi-Stage Burning (LIMB) is the injection of fine lime into combustion chamber, fixing sulphur, reducing emissions by 35-50%; Fluidized Bed Combustion (FBC): air under pressure is injected into fine mixture of coal, lime and sand, until whole mass behaves like a boiling fluid. Very efficient combustion, reducing sulphur emissions by 90%. Also reduces NO<sub>x</sub> because combustion temps. are relatively low.

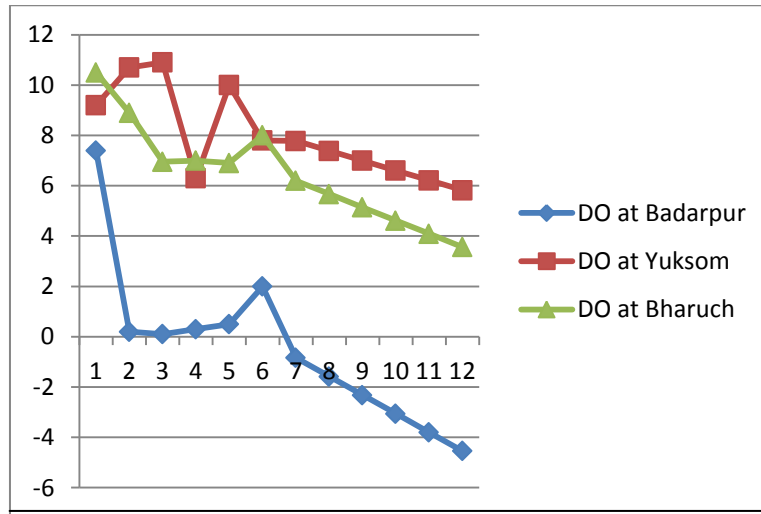


Fig 5: Prediction of dissolved oxygen in three main locations

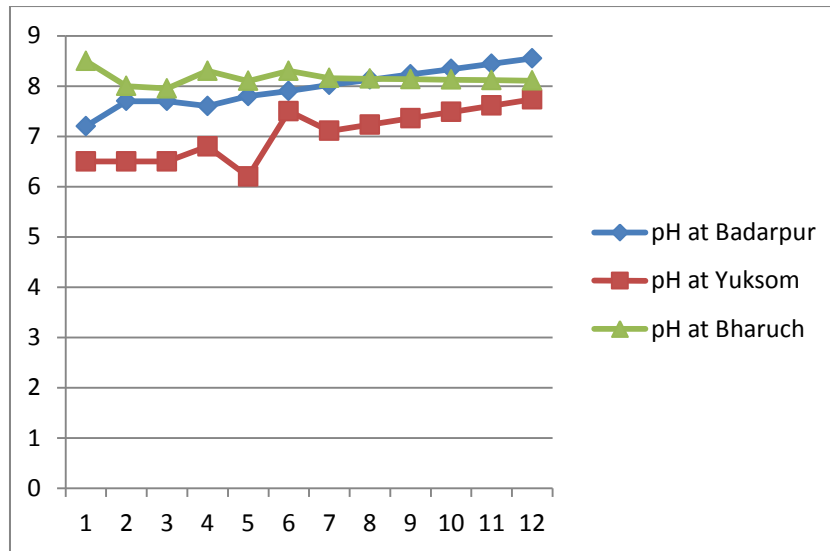


Fig 6: Prediction of pH in three main locations

If the pH level ranges outside 6.5 to 9 and DO is less than 6mg/l, certain pre-treatment measures can be suggested.

**Table 19: Prediction for year 2016, 2017 and Ranking with Fuzzy Soft sets method**

Site	Row-Sum	Column-Sum	Score (Si)	Rank
<b>2016</b>				
Site1(Badarpur)	12	16	-4	3
Site 2(Bharuch)	16	14	2	1
Site 3 (Tutocorin)	17	14	3	2
<b>2017</b>				
Site1(Badarpur)	14	17	-3	3



Site 2(Bharuch)	18	15	3	1
Site 3 (Tutocorin)	20	16	4	2

## B.2 Multilayer Perceptron ANN

This type of neural network is the most common supervised neural network. It consists of multiple layers of processing elements connected in a feed forward manner. This type uses the backpropagation of errors to train the Multi Layer Perceptron.

### Output Layer ( Prediction)

These components make up the output layer of the neural network. The synapse makes the connection between each second hidden layer PE and each output processing element (PE)[32]. The synapse contains the connections and the trainable weights for each connection. The second component is the bias axon. This component has the processing elements for the output layer, each of which sums the weighted connections from the second hidden layer. For classification problems, the output is a tan h axon that saturates at +/- 1 this is ideal for classification.

### Controllers

These three controllers contain the global control parameters for the network[32]. It contains parameters such as the number of epochs per run, the number of exemplars per epoch, the data sets to use, etc. The most important aspect of ANN in time series forecasting is "generality", which refers to their ability to produce reasonable forecasts on data sets other than those used for the estimation of the model parameters.

## B3. Learning Vector Quantization Neural Network Model

The original learning vector quantization neural network model (LVQ) used in our work has 87-10-4 architecture, that is, 87 neurons in the input layer, 10 neurons in the hidden layer and 4 neurons in the output layer. This architecture was fixed after performing several experimentations of training and testing, starting with just one hidden layer and varying the number of neurons in the hidden layer(s), and finally selecting the best configuration that gave the best results.

The above model was repeated with a reduced parameter set obtained with the parameter reduction techniques discussed in this work. The output for the corresponding input vector consists of a four-valued numeric vector representing the class to which a site belongs (i.e., V. Good, Good, Fair and Poor) using BEES method. For example, (1, 0, 0, 0) represents 'V.Good' class, (0, 1, 0, 0) represents 'Good' class, and so on. In fact the output class for each pattern has been assigned by referring to the guidelines provided by the Central Pollution Control Board, Ministry of Environment and Forest, Government of India. For some patterns, expert opinion was also considered to decide the output class of these patterns.

A total of 50 patterns were used in the LVQ simulations, out of which 70% (35 patterns) were used as training data and the remaining 30% (15 patterns) were used as the test data. We have used MATLAB software for our simulations. The patterns were created by referring to various EIA reports on setting up hazardous industrial power plants at different sites across the nation. Each pattern comprises values for the above-mentioned sub-attributes. During the training process, the output units are positioned by adjusting their weights through supervised learning. A supervised version of Kohonen learning rule was applied. The LVQ network consists of a single layer (competitive layer) of neurons, where each neuron is fed with the values of the input vector. The software simulation of a LVQ network involves introduction of an input layer to feed the inputs to the competitive layer and an output layer to get the user-defined output classes.

The learning vector quantization neural network model was trained with the following attributes or parameter values (in bracket): learning (supervised), learning rate (0.01), learning function (learning vector quantization 1), network performance or mean squared error (0.01), distance function (Euclidean), transfer function in the hidden layer (compet), and transfer function in the output layer (purelin). Before training, the input patterns were normalized in each of the 87 sub-attributes so that their values ranged in the interval [0, 1]. The error in the network predictions and the observed values gradually decreases with execution of epochs; where one epoch represents a training cycle that feeds each pattern in the training set to the LVQ network. The network required 10 epochs (10 training cycles, each comprising 35 patterns). The trained LVQ network was then used to predict the classes (later converted to ranks based on transfer function) for the 15 test samples. The results obtained are more or less same as those obtained by back propagation artificial neural network.

## B.4 Classifier Evaluation

In our work, the ANN classifiers such as back propagation neural network model, learning vector quantization neural network model have been used for the classification of the sites. In order to evaluate the performance of these classifiers we have used Naive Bayes and Decision Tree (Random forest) techniques. We converted our data files in excel to CSV (comma separated values) and then we imported this data file into the WEKA software and later saved it as .ARFF files(attribute relation file format). Weka is a software, which has a collection of machine learning algorithms for data mining tasks and it contains tools for classification, regression, clustering and visualization.

The Naive bayesian classifier is capable of calculating the most probable output depending on the input. This classifier estimates the class conditional probability by assuming that the attributes are conditionally independent, given the class label. With this classification, prior knowledge and observed data can be combined. The Bayesian type of classification

provides a useful perspective for understanding and evaluating many learning algorithms. The Naive Bayes classifier is a probabilistic learning method based on the Bayes theorem. It is a simple technique for constructing classifiers that assign class labels to problem instances which are represented as vectors of feature values, where the class labels are drawn from some finite set in this work as four class labels namely, A-V.Good, B-Good, C-Fair, and D-Poor. A Naive Bayes classifier considers each of these features to contribute independently to the probability, say this class is 'A-V.Good' regardless of any possible correlations between features. One major advantage of using Naive Bayes is that it only requires a small amount of training data to estimate the parameters necessary for classification.

Decision tree learning uses a decision tree as a predictive model which maps observations about an item to conclusions about the item's target value or class. Tree models where the target variable can take a finite set of values are called classification trees. Advantages of using decision tree classifiers are they are easy to implement, require small amount of data to train model and provides good results in most of the cases. A Random Forest classifier uses a number of decision trees in order to improve the classification rate.

**Table 20: Naïve Bayes classification summary**

Description	Value
Kappa statistic	0.5979
Mean absolute error	0.1511
Root mean squared error	0.3780
Relative absolute error	40.2404
Root relative squared error	87.0672
Coverage of cases (0.95 level)	74 %
Mean relative region size (0.95 level)	27.5 %
Total number of instances	50



Fig 7: Receiver Operating Characteristic curve

The ROC curve is shown in Figure 7, for 'class A' - V.Good with back propagation ANN. It shows the substitution between the true positive rate or sensitivity (proportion of positive tuples that are correctly identified) and false positive rate (proportion of negative tuples that are incorrectly identified as positive) for a given classification problem. Any increase in the true positive rate occurs at the cost of an increase in the false positive rate. The area under the ROC curve is a measure of the accuracy of the model. The vertical axis of the ROC curve represents the true positive rate. The horizontal axis represents the false positive rate. As we start from the bottom left hand corner where true positive rate and false positive rate are both zero. The kappa statistic is a chance corrected measure of agreement between classification and the true classes. It is calculated by considering the agreement expected by chance away from the observed agreement and divide this by the maximum possible agreement. A value of greater than 0 means that the classifier is doing better than chance. The kappa statistic can be further enhanced by using the bagging and boosting techniques.



**Table 21: Receiver operating characteristic curve area with Naive Bayes**

	LVQ ANN	BP ANN
Class A	0.8359	0.8186
Class B	0.8565	0.8515
Class C	0.8190	0.8325
Class D	0.8683	0.8012

In the Table 21, the receiver operating characteristic(ROC) curve values have been shown. For a classifier, if the area is closer to 1, the classification is more accurate. If the area is closer to 0.5 the lesser is the accuracy of that classifier. Receiver operating characteristic(ROC) curves are very useful tool for visualizing and evaluating classifiers. They are able to provide a better measure of classification performance than scalar measures such as accuracy, error rate or error cost.

#### 4. RESULTS

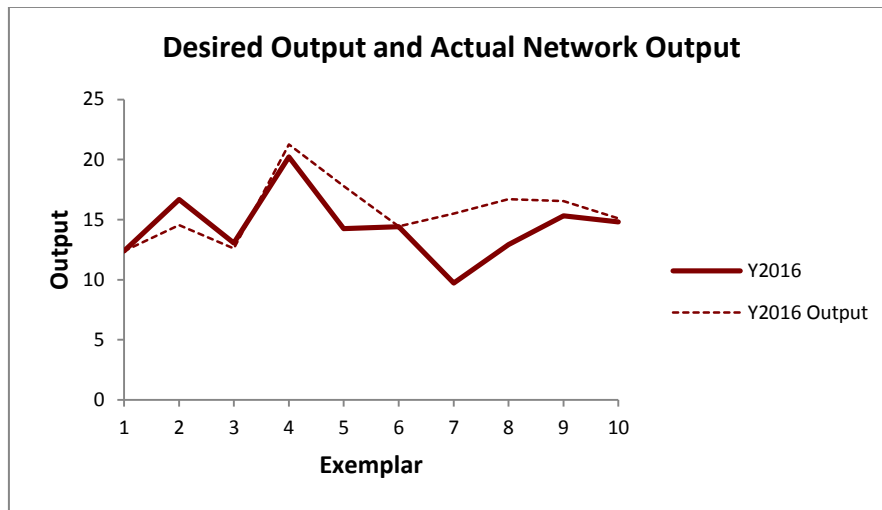


Fig 8: Desired and Actual Network Output obtained by MLP for SO2

Figure 8, shows actual and desired output of SO2 levels for ten locations with MLP. The values predicted by Time series have been passed as desired output values to MLP. The values predicted are for the year 2016. The plot indicates that the values are much within range and do not exceed the upper limit of 80 µg/m3. Therefore deterrent measures are not suggested for these locations. The prediction was done for important air polluting sub-attributes like SO2, NOx, PM10 and also important water polluting sub-attributes like DO, pH and BOD. In this ANN, 50% data was tagged as training data, 25% was tagged as cross validation data and 25% data was tagged as testing data. Therefore data of twenty sites was passed as training data, data of ten sites was passed as cross validation and ten sites was passed as test data.

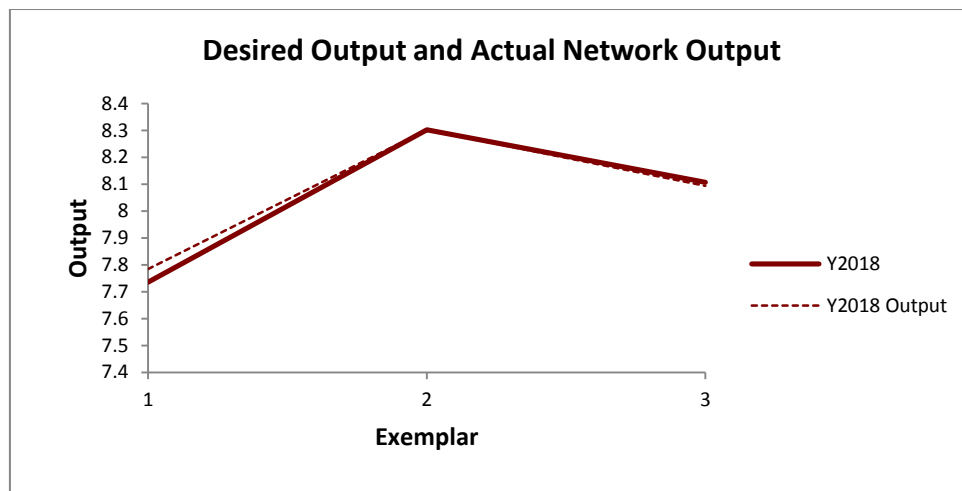


Fig 9: Desired and Actual Network Output obtained by MLP for pH

Figure 9, shows actual and desired output of pH levels with MLP for three important locations of Hydro electric Power plant namely Dibrugarh, Arunachal Pradesh; Chamba, Himachal pradesh; Yuksom, Sikkim. The values predicted by Time

series have been passed as desired output values to MLP. The values predicted are for the year 2018. The plot indicates that the values are much within range and do not exceed the upper limit of 8.5. Therefore, deterrent measures are not suggested for these locations. In this type of ANN 50 % data was tagged as training data, 25% was tagged as cross validation and 25% data was tagged as testing data.

**Table 22: Performance Table**

<b>Performance</b>	<b>Y2018</b>
MSE	0.000869146
NMSE	0.015705463
MAE	0.020876809
Min Abs Error	0
Max Abs Error	0.049289524
R	0.99780568

Performance table shows a mean square error of 0.0008 which is considered as acceptable and training was adequate.

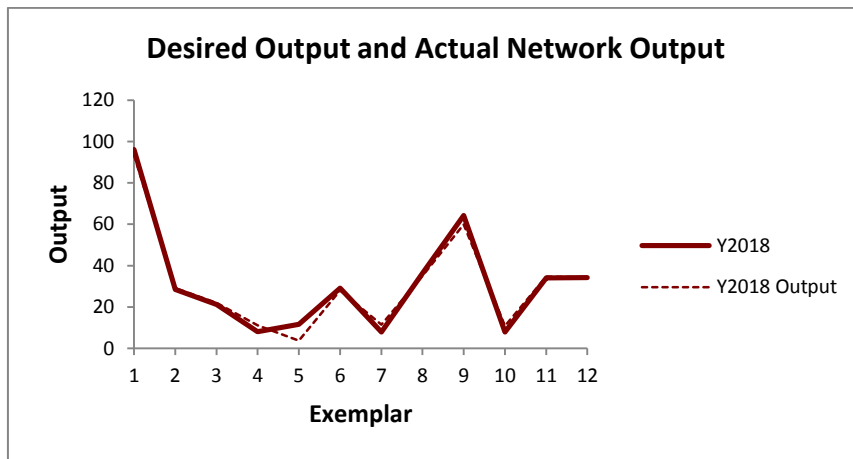


Fig 10: Desired and Actual Network Output obtained by MLP for NO2

Figure 10, shows actual and desired output of NO2 levels with MLP for twelve important locations of Power plants in India. Five Power plants located at Badarpur, Amreli, Bharuch, Tutocorin, Khammam. Three Hydro electric power plants located in Arunachal Pradesh, Chamba, Himachal pradesh; Yuksom, Sikkim. Two Nuclear Power plants located at Jaitapur and Kudankullam. Two Oil/ Gas based power plants in Patan and Shadol.

The values predicted by time series have been passed as desired output values to MLP. The values predicted are for the year 2018. The plot indicates that the NO2 levels have a steep rise for the first Thermal Power Plant site located in Badarpur and exceeds the permissible limit of 80 µg/m3. Therefore, Environmental Experts can analyze the trend of important parameters and decide if it is essential to issue deterrent measures like staged combustion, fluidized bed combustion, treatment of flue gas with systems like selective catalytic reduction and activated carbon process.

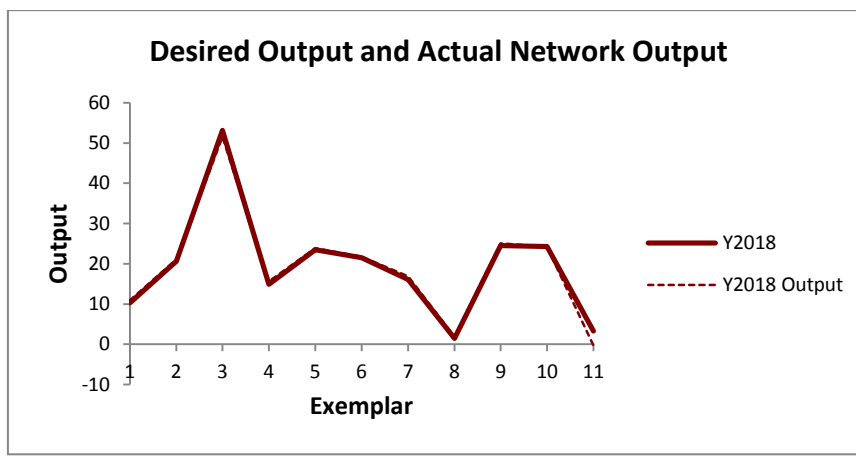


Fig 11: Desired and Actual Network Output obtained by MLP for SO2



Figure 11, shows the desired verses actual network output obtained by Multi-layer preceptron of SO2 for a given location.

**Table 23: Performance table**

<i>Performance</i>	<i>Y2018</i>
MSE	1.653917108
NMSE	0.009534173
MAE	0.86972002
Min Abs Error	0.08975206
Max Abs Error	3.602092118
R	0.995259005

The above table shows absolute error which is within acceptable range.

## 5. CONCLUSION

The main motivation for parameter reduction is to reduce computational complexity of certain soft computing methods, where the number of computations can become infeasible with increase in the number of decision parameters. It also helps to overcome the limitations of certain techniques. The research work explores various methods to identify a subset of important parameters, which can efficiently perform ranking without significant loss of accuracy. That is, the ranking of a given set of potential sites with a reduced set of parameters is almost same as that got by using the whole (larger) set of parameters.

The research work also explores methods to predict the futuristic values of some important decision parameters in order to study their trend over the coming years, followed by ranking of the sites using these futuristic values. This work is relevant as the site suitability calculated is based on the data available in the Environmental Impact Assessment (EIA)[11, 17, 19, 20, 21, 22, 23, 24, 26, 27, 28, 30, 31, 33, 35, 36, 38] reports and with Central pollution control board. It is important to study the past trends of the decision parameters and predict their futuristic values and subsequently find site suitability and the corresponding ranks over the coming years (five years). This aspect, to the best of our knowledge, has not been attempted before. Environmental Experts can suggest certain mitigation measures when the values of decision parameters for a given location do not fall into permissible range as per Central Pollution Control Board (CPCB) guidelines.

We have observed that the soft computing techniques when combined with experts' knowledge base can provide alternate or better methods for finding site suitability index and/or ranking of the sites for planning of a power plant installation. The model helps in overcoming some of the limitations of existing manual techniques, especially the computational and decision making complexity with the increase in the number of decision parameters. The decision results obtained using soft computing methods are free from any bias and ensure fair evaluation of the decision parameters. In our future work, we plan to explore the use of 'stack data' of power plants and its adjoining industries. If such data is made openly available, it could be used to study and analyze the industrial development around a power plant and also to infer the impact of the power plant pollution followed by ranking of the sites.

## REFERENCES

- [1] Abd-alsabour, Nadia, and Marcus Randall. "Feature Selection for Classification Using an Ant colony System", 2010 Sixth IEEE International Conference on e-Science Workshops, 2010.
- [2] Areerachakul S., Junsawang P., Pomsathit A., Prediction of Dissolved Oxygen using Artificial Neural Network, Computer communication and Management, pp 524-528, 2011
- [3] A.R. Roy, P.K. Maji, A fuzzy soft set theoretic approach to decision making problems, J.Comput. Appl. Math. 203 (2007), pp 412-418
- [4] Battelle Columbus Laboratories, <http://www.scopenvironment.org/downloadpubs/scope5/chapter04.html> (accessed October 2012)
- [5] Central Pollution Control Board, Government of India, 2012, <http://www.cpcb.nic.in/>
- [6] D. Molodtsov, A soft set theory – first results, Comput. Math. Appl. 37 (4) (1999)19–31
- [7] E. Bonabeau, M. Dorigo, G. Theraulez, Swarm Intelligence: From Natural to Artificial Systems, Oxford University Press, New York, 1999, pp. 73–78
- [8] Feng, F.. "Application of level soft sets in decision making based on interval-valued fuzzy soft sets", Computers and Mathematics with Applications, 201009



- [9] Grande J. and Suarez M. and Villar J., A Feature Selection method using a Fuzzy Mutual Information Measure, Innovations in Hybrid Intelligent Systems, pp 56-63, 2007
- [10] Goodarzi, Mohammad, Matheus P. Freitas, and Richard Jensen. "Feature Selection and Linear/Nonlinear Regression Methods for the Accurate Prediction of Glycogen Synthase Kinase-3 $\beta$  Inhibitory Activities", Journal of Chemical Information and Modeling, 2009
- [11] EIA, Nuclear Power Project Jaitapur, National Environmental Engineering Research Institute, Nagpur, 2010
- [12] Jensen R., "Supplementary Developments and Investigations", Computational Intelligence and Feature Selection, 09/09/2008
- [13] Jensen R., "Further Advanced FS Methods", Computational Intelligence and Feature Selection, 09/09/2008
- [14] Jensen R., "Applications II: Web Content Categorization", Computational Intelligence and Feature Selection, 09/09/2008
- [15] Jensen R., Studies in Computational Intelligence, 2006
- [16] Jensen R., Combining rough and fuzzy sets for feature selection- Ph.D. Thesis, 2005, University of Edinburg
- [17] Jensen R., "Performing feature selection with ACO", Studies in Computational Intelligence, 2006
- [18] Kalyani Sambhoo, Sanjay Kadam, Ashok Deshpande, Rule based fuzzy indexing for grading of proposed industrial sites for power plant installation, Int. J. Computer Technology. 10 (7), (2013)
- [19] Kudankulam Nuclear Power Project, Environment Impact Assessment Report, Nuclear Power Corporation Ltd., India, 2011
- [20] Marco Dorigo, Mauro Birattari, Thomas Stützle, Ant Colony Optimization –Artificial Ants as a Computational Intelligence Technique, 2006
- [21] Matrix, Environment Impact Assessment of Nuclear Power plant in Jaitapur,
- [22] Matrix Thermal Power Pvt Ltd, 2010, Khammam, Andhra Pradesh
- [23] Narmada Thermal Power Limited, Draft Environmental Impact Assessment Report for Thermal Power Plant, Bharuch, Gujarat, 2011
- [24] NTPC Limited, Badarpur Combined Cycle Power Project, Badarpur, Delhi, 2010
- [25] Pandey, J. and Rakesh Kumar and Devotta S., Health risks of NO<sub>2</sub>, SPM and SO<sub>2</sub> in Delhi (India), Atmospheric Environment, pp 6868–6874, 2005
- [26] Patel Energy Limited, Draft Environmental Impact Assessment for Coal Based Thermal Power Plant, Amreli, Gujarat, 2011
- [27] Principal Component Analysis,  
[https://en.wikipedia.org/wiki/Principal\\_component\\_analysis](https://en.wikipedia.org/wiki/Principal_component_analysis), [Online; Accessed 20 Dec, 2014]
- [28] R. Jensen, Q. Shen, Computational Intelligence and Feature Selection Rough and Fuzzy Approaches, John Wiley and Sons, Inc. Pub, Hoboken, 2008
- [29] Ross, T., Fuzzy Logic with Engineering Application, 1995, isbn 0-470-86075-8, Mc Graw Hill, New Mexico
- [30] R.S. Envirolink Technologies, GMR Bajoli Holi Hydro Power Plant, Environmental Impact Assessment Report, 2010
- [31] R.S. Envirolink Technologies, TT Energy, Environmental Impact Assessment of H.E. Project, 2010
- [32] Saikia, L. C., S. M. Borah, and S. Pait. "Detection and classification of power quality

- disturbances using wavelet transform, fuzzy logic and neural network", 2010 Annual (INDICON), 2010. IEEE India Conference
- [33] Sangeetha R., B. Kalpana, Enhanced Fuzzy Rough set based Feature selection strategy using Differential Evolution, International Journal of Computer Science and Applications, pp 13-20, 2013
- [34] Sfetsos A., Vlachogiannis D., Time Series Forecasting of Hourly PM10 using Localized Linear Models, Journal of Software Engineering, pp 374-383, 2010,
- [35] Shepherd R., Quantifying Environmental Impact Assessments using Fuzzy Logic, 2005, ISBN 978-0-387-28098-1, Springer
- [36] Smec India, Environmental Management Plan for NAFRA Hydro Electric Power Project, Smec India, SEW Nafra Corporation, 2009, Kameng, Arunachal Pradesh
- [37] S.N. Sivanandam. "Introduction to Particle Swarm Optimization and Ant Colony Optimization", Introduction to Genetic Algorithms, 2008
- [38] SPIC Electric Power corporation Pvt., Ltd., HW Tuticorin Thermal Power Project, 2010, Tuticorin, Tamilnadu
- [39] Sun, Yuanchang, and Jack Xin. "Nonnegative Sparse Blind Source Separation for NMR Spectroscopy by Data Clustering, Model Reduction, and  $\ell_1$  Minimization", SIAM Journal on Imaging Sciences, 2012.
- [40] Tanli Kuang. "A method to calculate importance weights of fuzzy soft sets with unequal importance weights", 2010 Seventh International Conference on Fuzzy Systems and Knowledge Discovery, 08/2010
- [41] Tao, Zhifu, Huayou Chen, Xia Song, Ligang Zhou, and Jinpei Liu. "Uncertain linguistic fuzzy soft sets and their applications in group decision making", Applied Soft Computing, 2015.
- [42] Venkadesh S., Hoogenboom G., Potter W., McClendon, R., A genetic algorithm to refine input data selection for air temperature prediction using artificial neural networks, Applied soft computing, pp 2253-2260, 2013
- [43] Xiuqin M., Sulaiman N., Hongwu Q., Herawan T., Zain, J., A new efficient normal parameter reduction algorithms of soft sets, Computers and Mathematics with Applications, Vol 62, pp 588-598, 2011
- [44] Zhi Kong. "Changing entries in soft sets", 2010 International Conference On Computer Design and Applications, 06/2010

### Author's biography with Photo



**Dr. Kalyani Sambhoo (Salla)** has completed Ph.D. in Computer Science from the Department of Computer Science, Savitribai Phule Pune University, Maharashtra, INDIA. She has done M.Phil. (Computer Sc.), YCMOU and M.C.S. (Computer Science), Pune University and B.C.S. (Computer Science), Pune University. She is working as an Assistant Professor at the Department of Computer Science, Modern College, Pune. Her research interests include Fuzzy Logic and Neural Networks.



**Dr. Sanjay Kadam** works as a Joint Director-ECIP in the Evolutionary Computing and Image Processing Group at Centre for Development of Advanced Computing, Pune, India. He has a Ph.D. in Computer Science from the University of London, UK, an M.Tech in Computer Science from Indian Institute of Technology (IIT), New Delhi, and an M. Sc in Mathematics from Pune University. His research interests include Computer Vision and Image Processing, Neural Networks, Parallel Processing, Fuzzy Computing, Web Technology, and Software Engineering.