Available online at www.joac.info

ISSN: 2278-1862

Journal of Applicable Chemistry

2014, 3 (1): 6-29

(International Peer Reviewed Journal)

State-of-Art-Review (SAR-Invited)
Mathematical Neural Network (MaNN) Models
Part I#: Data-driven Soft-models for ozone in air quality#

K RamaKrishna[1], Ch. V. Kameswara Rao[1, 3], V. Anantha Ramam[2] and R. Sambasiva Rao[2*]

1. Department of Chemistry, Gitam Institute of Science, Gitam University, Visakhapatnam, 530 017, **INDIA**
2. School of Chemistry, Andhra University, Visakhapatnam 530 003, **INDIA**
3. GMRIT, Rajam, Srikakulam District, **INDIA**

Email: karipeddirk@gmail.com, rsr.chem@gmail.com

Accepted on 05th January 2014

Dedicated to Malladi Narasimha Sastry (popularly known as M N Sastry), former professor of Nuclear chemistry, Andhra University, on his *sahasra chandra darsanam* (thousand lunar months of life on the lap of mother earth)

ABSTRACT
The quality of environment especially of air and water has a key role on human health. The effect is pronounced on children and vulnerable groups. The efforts of governmental agencies are to forecast and inform the alarming situations of the day in advance. The monitoring, modelling and now short term and long term prediction is pivotal in environmental research. And, passing the skills down to routine maintaining stations are in the direction of avoiding the ill-effects of pollutants. The chemical, biological and meteorological micro processes and their interaction are complex, non-stationary and thus cannot be handled from first principles. Artificial neural networks (ANN), later preferably called neural networks are in fact Mathematical Neural Network Models (MNNM) partially resembling mega biological neural networks. These data driven NNs, a subset of evolutionary models revolutionalised the modeling practices in environmental science since late 1990s. The earlier model driven linear/non-linear models are integrated as a priori knowledge in sequential NNs. An increase of 0 to 1.7K and 1.4 to 2.4K of temperature in select US cities is forecasted during the periods 2020-2050 and 2051-2099. This long term forecast paves way in the management and planning of eco-balance. Cause and effect models fail for ozone due to non-linear, dynamic processes occurring in the formation, decomposition and transport of ozone. Tropospheric ozone is modelled by SLP-NN using 13 variables. The prediction of ozone at ground level in UK, Europe, Athens and Dallas and surface level in Austria, UK, Chile and Korea with NNs is more reliable compared to non-linear multiple regression and generalized additive model. The excedence predictions in Athens as well as rural places in UK are successful experiments. NEUROZONE is an automatic software prediction tool and it is successful in six out of seven incidences of ozone excedence. The factors for failure of ozone predictive models are described.

Keywords: Air pollution, Ozone excedence, Human health, Neural network model, Forecast, Nature mimicking algorithms.

#: Presented at National Seminar on Climate change – Impact on Environment and Human health, 11th and 12th February, 2011, Gayatri Vidya Parishad, Visakhapatnam

Contents

1. Introduction
   1.1 Complexity of Environmental data
      1.1.1 Targets of environment quality
      1.1.2 Prediction of temperature by 2099
   1.2 Air quality
2. Computational models
   2.1 Time series modeling
   2.2 Nature inspired algorithms
   2.3 Neural Networks
      2.3.1 Internal representation of input in hidden layer
      2.3.2 Advances in SLP
      2.3.3 M(ultiple hidden) LP-NN
   2.4 Hybrid models
3. Results and discussion
   3.1 Formation and destruction of ozone
      3.1.1 Surface level O_3
      3.1.2 Troposphere ozone
      3.1.3 Global scenario of ozone
      3.1.4 Ill effects of ozone on human health and plant kingdom
         ■ Remedial measures for ill effects of ozone
   3.2 Complexity of ozone modeling
   3.3 Modeling and forecast of ozone
      3.3.1 Ground level and tropospheric ozone
      3.3.2 Modeling high ozone event
         Case studies
   3.4 Automatic predictive computational tools
   3.5 Failure of forecast models for ozone
4. Future scope

INTRODUCTION

The constituents and composition of environment, water, soil and their interfaces are continually changing with human activity and rare natural calamities. The dynamics in spatio-temporal regime and many complex interactions restores eco-balance. But, when the balance is out of gear, the quality of environment progressively diminishes to a point of concern. At this juncture, human intervention is indispensable as it is only alternative for prevention and remedial measures to the possible extent. It led to upsurge of environment and health as a scientific discipline with the advent of stipulations by US-EPA, EC etc. to uphold the health of life on the globe. The number of even published reports exponentially increased and models are diverse from simple multiple linear regression (MLR), autoregressive (AR) moving average (MA) to hybrid neural networks (NNs) and other nature inspired procedures. On the experimental front, high frequency sampling, automatic analysis with hyphenated instruments increased the accuracy of data procurement. The high end on-board laboratory in satellites outpour precise primary data like of total ozone in spatio-temporal regime. In this review, how the forecast of ozone improved to an appreciable level using neural networks [1-49] is documented with select case studies globally.

Based on the exponential growth of publications in apparently diverse disciplines using advanced information extracting mathematical tools and hyphenated instruments, we reported earlier the results of NN models in multivariate calibration (MVC) [5], ceramics preparation [38], chemical kinetics [44] and bauxide ores [46]. In continuation of our efforts in application of NNs in fisheries [45] (piscimetrics), medicine (medicinometrics), nanoscience (nanometrics), quantum chemistry, pharmacy (pharmacometrics), gas chromatography [50], and chemistry (chemometrics) [51], the impact of NNs in air quality especially concerned with ozone are reviewed here.

1.1 Complexity of Environmental data : A large amount of data has been collected [47] in environmental chemistry. It is a prime source for deriving information and knowledge. The amount of data now available exceeds the scope for manual analysis. Environmental data are complex [41] due to inter relationship between numerous factors of different type viz., chemical, meteorological [42], biological etc. Further, the non-linearity, non-stationary noise in the measurement, missing data, wear and tear of the instrument etc. are the additional factors. The modeling is not trivial, as different combinations and interactions of variables explicitly not known are involved in micro and macro processes. The environmental data as a function of a time, the upward trend of pollution and day-night cycle patterns are introduced in modeling studies in recent past. Instead of stepwise elimination of the trend, periodicity, spikes and noise in the conventional time series analysis, all are simultaneously modeled in NNs.

1.1.1 *Targets of environment quality*: Total quality index is a prime requisite to check the progress and revisit the ongoing schedules as well as in new launching projects. The 2012 target program [12, 34] aimed at meeting the most difficult challenges in air quality. The governmental policies are for strict implementation of reduction of primary pollutants. The scheme includes defining baseline and specific reduced air emission as much as possible to meet national ambient air quality standards. Based on the Governmental laws and advices of the country and state stipulations, the objectives of air quality agencies are monitoring pollutants, forecasting the time/amounts of peaks in the pollutants, informing authorities/public the alarming situations and adapt (short term as well as long term) control measures to reduce emissions. This includes rescheduling the (traffic) activity and implementing stringent measures to reduce the emission at source. The long-term plans [23] include establishing air pollution prevention act operable within the state. It includes imposing restrictions on the sanctions, renewal and expansion of industries/commercial units and traffic. Sometimes, change over to a (new) green process on the manufacturing front is desirable. An integrated system is envisaged in developing a database of the meteorological and chemical information, estimating the warning levels for critical pollutants, predicting short term/long-term trends, proposing measures for immediate reduction of pollutant concentration below the alarming thresholds and long term plans. However, meteorological changes are the highly influential factors compared to pre-cursor concentrations in ozone trend. That is the sole reason of contradictory reports. Hence, an in depth monitoring of complex micro-atmosphere and its contribution to macro-atmosphere processes [15, 16] are mandatory.

The prime objectives of the environmental monitoring by Governmental and academic bodies are to diminish the activities contributing to lowering the quality of air, water and soil. The long-term focus is to bring down the concentrations of pollutants below the maximum tolerable limits even in labeled polluted zones.

The short term but indispensable activity is to alert the public about the time duration of maximum levels of ozone, SO_x, NO_x etc., to minimize the exposure of the children and high risk group of vulnerable individuals. This surge resulted in low and high frequency sampling and need-based analysis of varying levels of sophistication in local and global scenarios. In spite of the concerted efforts of researchers, technical staff over the last four decades with committed large budget, the accumulated data is hoch poch. However, with the limited reliable data in select pockets, the predictive modeling resulted in the strategic planning for cleaner (and reduced polluted) environment. The continual studies in this direction are for utopian living conditions with unperturbed health of the human beings on the planet by 2099 through 2050.

1.1.2 Prediction of temperature by 2099: The change in climate, which reflects the effects of global warming is of utmost importance. Global climate models attempt to simulate the behavior of the climate system. In-depth studies probes into the macro-/micro-/nano-(molecular level)-/single_cell-/single_molecule-physical, chemical, geological and biological processes in pure-phases and interfaces responsible. This paves way to understand the past as well as future climate changes. The earlier equilibrium model is not valid because the components of the climate are not in a steady state. The recent results from transient climate change experiments are of use to translate coarse scale outputs into finer details.
Using a hybrid model for time series data of minimum and maximum daily temperatures in 26 stations of mid western of USA is generated. The time periods are 1990-2001, 2020-2099 and 2050-2099. This approach was found to give better results than a transfer function method. An increase of 0 to 1.7K and 1.4 to 2.4K are predicted for 2020-2029 and 2050-2099 respectively. These long term forecast of global warming indicates an average maximum increase of 3°K and already high polluted pockets with no hope of meeting even earlier EPA standards of good air quality further complicates the already complex scenario. SMAURL was a European project on environmental monitoring of Villa San Giovanni with an objective of designing mobile terminals to monitor air quality. The gathered real time data at urban centers using NN modeling is instrumental for administrative decisions in environmental protection.

1.2 Air quality: The perturbation in the composition and quality of air as a result of industrial revolution no doubt partly compensated by natural adjustments (eco-balance), but human intervention hastens capping of black hole. The human health and survival period are complex to the core, but alarming in polluted stripes. The risk increases and may lead to a fatal condition for vulnerable (and/or aged) people and children. The utopian goal of reduction of pollution to zero level over the globe is a herculean task now. Yet, one of the priority targets of the governments is to take measures to arrest the increase of harmful gases/vapours, solids and aerosols in the environment.
The air monitoring is one of the prime concerns to reduce health risk and to sustain eco-balance. Monitoring of a pollutant, or a benign chemical [35] is equally important to study the state of environment and human health. The governmental, non-governmental, local bodies and research institutes are indulged in developing the databases almost in continuous time scale at urban as well as rural locations. The other objective is to arrive at predictive models with robustness and applicability within the achievable accuracy. The targeted objectives mostly in practice are pollutant monitoring and one-/two-day forecast of peaks in O_3, SO_x, NO_x etc. in densely populated cities. The consequent useful information is to take measures to regulate sanctions for industrial operations and alert the vulnerable public to safe guard their health. US-EPA designated San Diego (CA), Yuma (AZ), EL Passo (Tx) as cities of no scope for attainment of high air quality**.**
Vehicular emission consists of gases and particles. The accumulation leads to high levels of air born pollutants especially when inversion occurs in the lower layer of the atmosphere [25]. The geography with surrounded mountains, light winds and atmospheric temperature are conducive for occurrence of inversion. The ambient particles are primary cause of decrease in quality of life resulting in morbidity/mortality and Los Angeles is one of the highly polluted areas. CO, NO_2, particulate matter (PM10), O_3 and SO_2 are some of the pollutants of utmost concern. The European parliament passed the directive, stipulating the ceilings for emission of the said pollutants. The measures like ‘Sunday on foot’ and ‘odd/even number plates’ for cars etc. are in the direction of reducing the pollution resulting from vehicular traffic. Under this scenario short term (six to 24 hour ahead) forecasting is sufficient. Further, the probability of alarming situations alerts the governmental agencies to control the man made activity diminishing the ill effects of catastrophes. This will prevent the annoyances to the residents of the city. Already many agencies accepted NNs to be superior information technology compared to the well-nurtured statistical methods.

COMPUTATIONAL MODELS

Traditional mathematics uses abstract/symbolic variables combined through arithmetic/logical/Boolean operators. The solution is obtained by deductive and inductive methods. Geometry adds visualization of transformation of problem space to solution space. Algebra (linear/tensor) simplifies the code. Transformation of explanatory/response variables into mathematical/chemical/physical spaces and (first, second) derivatives of object function facilitates the application specialist comfortable with solution of the task. Mathematics is deemed to be a clear thinking as it is expressed by a series of unambiguous steps. Even understanding a phenomenon completely leads to microscopic set of processes. Thus mathematics or biology both being a subset of nature have the intrinsic characteristic of clarity i.e., what is happening. The opaqueness/difficulty is with the perceiver, his tools, comprehension, retention, integration, recapitulation and expression.

2.1 Time series modeling : The response of a mega to micro-processes with the lapse of time (on different scales viz. minutes to years) found a niche in the annals of data as 'Time Series' (TS). In time series models, time itself is considered as independent (explanatory) variable unlike in regression and optimization procedures. The objective of time series modeling is to develop a relationship between the current observations with the previous ones in the time domain. Input values are only lagged values of response and do not require explanatory variables. Classical time series analysis considers (linear) trend, seasonal component and spikes as components. The modeling is performed in a sequential manner after removing the spikes. The sum of responses of trend-model and season-model must account for the observed response. In this case, the residuals between observed and calculated values are only a random (or white) noise. But, the necessary condition is that data should be stationary. Otherwise, the data is made stationary by de trending, calculating first and second order differences and Fourier transformation.
The time series data of real-world phenomenon are rarely pure linear or nonlinear, but often contain both. The approximations of ARMA models to complex nonlinear problems are inadequate. NNs for pure linear problems have yielded mixed results. The human expert/analyst does not get a priori information and thus a hybrid system outperforms/superior to the component models even in prediction.
The current response sometimes depends upon m-previous values and the data set is termed as auto-correlated. When, there is no trend after finite differences of autocorrelation function (ACF), the time series is considered as stationary. But, in many real life problems non-stationary time series prevail. As the complexity of the task grows, multiple time series arise. Persistent model advocating the same behavior in the next time step fails, in all real life situations of weather forecasting. The subtle interest of governmental agencies in time series data in forecasting (predicting) m-steps ahead in time for the retrospective inspection and development of theories/hind-cast is a novel tool. The term 'now cast' is also referred to as forecast, for instance ozone or pollution level within an hour/eight-hour period.
Moving average (MA), auto regression (AR), ARMA, ARMAX, ARIMA are the linear time-series models accounting for the effect of mean, first order processes with different lags and exogenous variables. Recent methods viz. VEC, GARCH tackle complicated time series trends. The periodicity is modeled with spectral analysis techniques like Fourier transform (FT), peridiogram, Kalman filter, its advance and Voltera series. Many parametric and non-parametric methods [24] have been in vogue extensively for short and long term time series forecasting. System identification is an integrated software paradigm to automatically analyze time series data, of course with intermittent analyst's options.

2.2 Nature inspired algorithms : Mathematicians and/or application scientists are inspired by the tiny as well as complex macro processes in nature and exploited the titbits to formulate new mathematical approaches. Although, the understanding of nature’s path is incomplete, translation into algorithm is biased/restricted by constrains. Further, the implementation is again confined by mathematical framework. But, astonishingly, the end result miraculously excelled the so called structured procedures. This stimulus arose interest to pursue nature mimicking algorithms (E-man, Evolution of Mimics of Algorithms of Nature or Name, Nature’s Algorithms Mimics Evolution) as a distinct/separate paradigm [11, 48]. By now, around 30 basic modules are popular. Also, the modifications in each one of them, hybridization up to quaternary systems and implementing hierarchical approaches resulted in an exponential number of computer software implements. At this juncture, it is a herculean job to say which set of algorithms evolves to be the best for a task. The only way at disposal now is to progressively apply the classical and E-man for tasks of increasing complexity and gain intelligence and render computational intelligent tools of the day super/hyper intelligent. Evolution plays a key role in striking balance in the lifecycle of the species in the flora and fauna, animal kingdom and environment. The changes to combat against predators, survival with rapid fluctuations in harsh environment and to improve survival periods are affected after a passage of large number of generations. The timescale appears to be short in species with huge offspring production and very long in higher animals with low offspring. The mutation in long life species is a boon as well as a curse, but it is a natural way of radical spur in traits.

Computer scientists of 1960s attempted to implement the natural human traits (viz. vision, speech, hearing, theorem proving and decision) on a computer (machine)/robot. This was the start of a new era, artificial intelligence (AI) which mimics the then believed natural intelligence. With an in depth research in neuro-biology/physiology/brain chemistry/surgery over the last half a century opens the eyes that the function/impairment/dysfunction of each component of a natural sense organ involves coordinate distributed synchronous and asynchronous billions of physical/chemical processes at the core. The consequences of these processes mediate through biomolecules and in vivo environment at the organ level. It is definitely astonishment within the common sense frame of human race. But, many amazing outcomes of processes right from a honeybee/ant to bats/birds through fish/frogs, rats/cats in their foraging, defense/reproduction activity follow coolly without any tag of pride of intelligence. A bird's eye view of retrospective survival (save extinct species) of animated kingdom is through evolution, combating against hardships of the other life and drastic changes of environment from ice ages to the present day polluted scenario.

The word intelligence became popular and accepted for anything above the common (average) level at that point of time and in that spatial region. Thus common-/ nature- swarm-/computational- intelligence is in vogue. In fact the tools of extracting knowledge are now better compared to those of writing down the heuristics manually and implementing in software. The jargon of first two decades of expert systems is a major chunk of first generation AI. The traits/habits/external and internal biological changes of any species, ant to higher animals through evolutionary processes (skills at different time) is mainly for survival. The foraging, defense from predators and reproduction are bare minimum activities in the life cycle. In addition, home building and colony shifting in ants/honeybees and birds add one more dimension. The very high voltage production in electric fish with the activity of a single neuron, vomiting digestive system on the invading enemy and death of honeybee after biting the enemy during comb protection are a few among thousands of defense operations. The evolution approach of algorithms from simple to complex depending upon task/ sub goals/data structures/ accuracy/CPU cost crossed teething problems. Artificial neural network models are first software products from inspiration of brain in general and that of human being in particular. The credit goes to McCulloch and Pitts' [30] seminal contribution in nineteen forties.

2.3 Neural Networks: The application of neural network (NN) models in time series (TS) analysis is multifold. The advances in NN and the imbibing character paved way for dedicated NNs implementing all types of AR models viz. ARMA, ARIMA, ARMAX models. Further, MLP and SLP NNs have been used with lagged inputs to forecast the future trend. Recurrent NNs (Rec. NN) had a niche in the complex TS tasks.

In 1986, Rumelhart and McClelland [36] proposed a single hidden layer perceptron (SLP)-NN affecting the data flow from input, hidden to output layers. The layers are successively connected and there are no feed back or direct connections. Each unit in any layer has fan-in connections from all units in the preceding (immediately below) and fan-out connections to all units in succeeding (immediately above) layer (Fig 1). The word hidden is used as an end user is not interested in the details of operations in it. The strengths of connections are termed as weights (W) in analogy with synaptic strength in neurobiology. A simple nonlinear (sigmoid) function is used as TF in the hidden layer. Further, it is neither a rule nor exception of restricting a single transfer function in a PE and/or in each layer.

Fig.1 : Object oriented (OO-) format of NN-architectures

(a) RBF

```
%
% oo_rbf.m  (R S Ra0) 14-9-10
%
clean
   dimx = 3; dimy = 1; NP = 11;
   dimRL = 3;

% Network Name
  NN.name ='RBF'

% Layer Names

 RBF.Layer(1,:).name = 'IL'
 RBF.Layer(2,:).name = 'RL'
 RBF.Layer(3,:).name = 'OL'

 % Layer dimensions --Data
structure

 RBF.Layer(1,1).neurons = dimx
 RBF.Layer(2,1).neurons = dimRL
 RBF.Layer(3,1).neurons = dimy

 % Weight

 RBF.weights(:,1,2) = 'WIR'
 RBF.weights(:,2,3)  = 'WRO'

 % Number of patterns
 RBF.NP = NP

 %  --Architecture ----

 RBF.connections =
'[Full,Forward]'
 RBF.confluence = '[dot product]'

 % Layer TFs

 RBF.Layer(1,:).TF = 'Identity'
 RBF.Layer(2,:).TF = 'Kernel'
 RBF.Layer(3,:).TF = 'Identity'
```

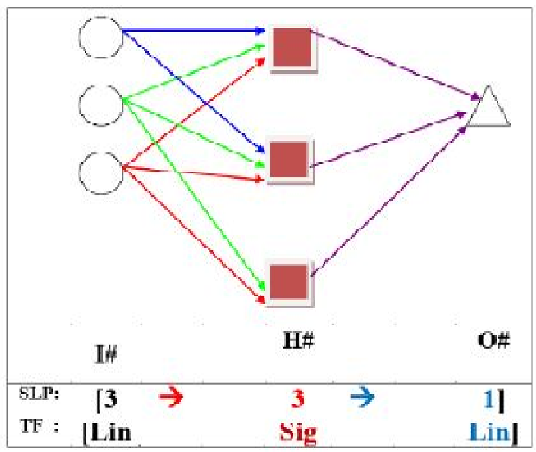
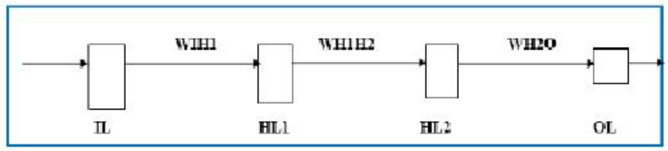


Fig. 1(b) MLP

```
%
% oo_mlp.m  (R S Ra0) 14-9-10
%
clean
   dimx = 2; dimy = 1; NP = 11;
   dimHL1 = 5; dimHL2 = 4;

% Network Name
  NN.name ='MLP'

% Layer Names

 MLP.Layer(1,:).name = 'IL'
 MLP.Layer(2,:).name = 'HL1'
 MLP.Layer(3,:).name = 'HL2'
 MLP.Layer(4,:).name = 'OL'

 % Layer dimensions --Data structure
*****

 MLP.Layer(1,1).neurons = dimx
 MLP.Layer(2,1).neurons = dimHL1
 MLP.Layer(3,1).neurons = dimHL2
 MLP.Layer(4,1).neurons = dimy

 % Weight

 MLP.weights(:,1,2) = 'WIH1 '
 MLP.weights(:,2,3) = 'WH1H2'
 MLP.weights(:,3,4) = 'WH2O '

 % Number of patterns
 MLP.NP = NP
 %
************************************

 %  --Architecture -------------

 MLP.connections = '[Full,Forward]'
 MLP.confluence  = '[dot product]'

 % Layer TFs

 MLP.Layer(1,:).TF = 'Identity'
 MLP.Layer(2,:).TF = 'TFH1'
 MLP.Layer(3,:).TF = 'TFH2'
 MLP.Layer(4,:).TF = 'Identity'
 % ------------------------------
```



$$MLP:\left[\begin{bmatrix}[Input,\ [I\#,H\#,O\#]],\\ [[Confl\_HL1,Confl\_HL2,Confl\_OL],[WIH1,WH1H2,WH2O],[TFHL1,TFHL2,TFOL],]\\ Output\end{bmatrix}\right]$$

The weights (WIH, weight matrix of connections from input to hidden layer, WHO, weight matrix of connections from hidden to output layer) are learnt (refined) by back-propagation algorithm, which is steepest gradient procedure tailor made for on-line learning. A transformation of input with a non-linear (sigmoid) TF and back-propagation learning algorithm not only successfully implemented XOR gate, but has become a laudable architecture for a reborn and revitalized NN paradigm.

It has only a little more intelligence in solving pattern recognition task. But, SLP imbibes a galaxy of hither to matured mathematical/statistical procedures. This paradigm opened new vistas in non-collapsing learning of even odd data structures with unknown complex functional relationships. The input to output (I/O) mapping in NNs $y\_SLP = \left[Unscal \left[\left\{ TFHL \left[\left(\left\{ scal \left(x^T \right) \right\} * WIH \right) \right]^T \right\} * WHO \right] \right]$ is affected by choosing architecture (i.e. number of hidden layers, type of connections between neurons), TFs and training sets.

2.3.1 Internal representation of input in hidden layer : The output of sigmoid neurons of the hidden layer produces vectors that are linearly independent of input vectors [21]. In other words, the dot product of WIH and X passed through sigmoid transfer function (SGTF (X^T*WIH)) results in neuron-unit wise (not pattern wise) internal representation. The output from each of hidden layers is referred as internal representation of I/O mapping in NNs. The linear independence of internal representation of SLP neurons is an essential property for exact learning of a function. Yet, the hurdles are noise and outliers in data structure to obtain a solution of desired quality. The optimisation method, object function (performance, error, parameter reliability) and ranges of data play a major role in solution, generalizability, reliability, predictability, etc. A detailed account of the calculations for SLP (1-2-1) for a simulated function $y = x^2$ is given in Fig 2. The development of output from individual hidden neurons is in Fig. 2c and 2d, while Fig. 2e and 2f depict cumulated one and after scale transformation to the data domain respectively.

Fig 2. : Function approximation with SLP

(a) DATA and output

x	y	ynn	ynn-y
-4	16		
-3	9	8.999663	-0.000337
-2	4	4.001684	0.001684
-1	1	0.9959792	-0.004021
0	0	0.00531	0.00531
1	1	0.9959907	-0.004009
2	4	4.001671	0.001671
3	9	8.999667	-0.000333

Weight matrix

MLP	h1#01	h1#02	y
Threshold	0.8566991	4.208081	0.8888181
X	-3.351389	3.3511	
h1#01		6.040439	
h1#02		6.041658	

Architecture

	I-H1-O	PSP	:	Lin
architecture	1-2-1	Error fn	:	SS
TFs	lin-logis-lin	Tr	:	Marquardt

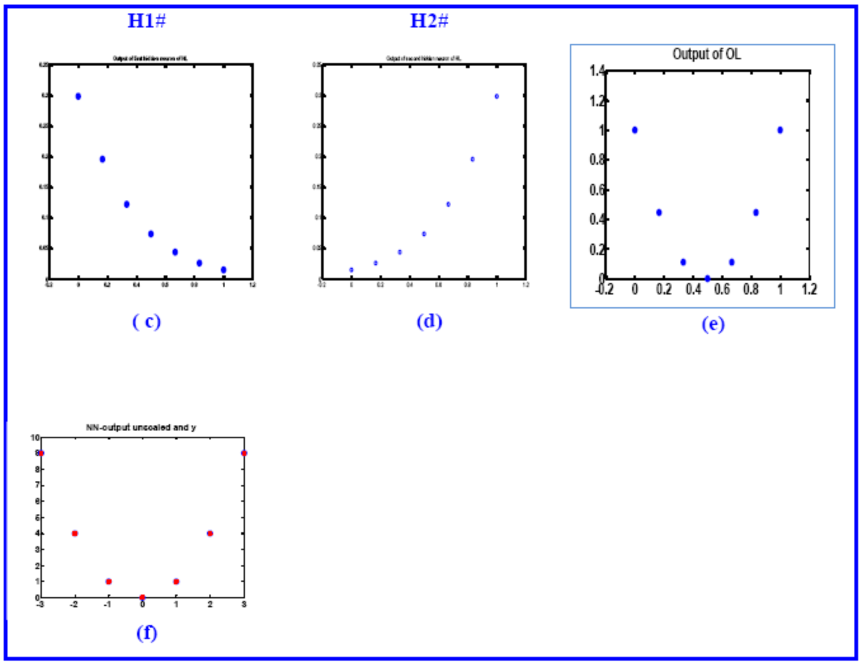
Scaling

	Shift	Scale
X	0.5	0.1666667
Y	0	0.1111111

(b) Scaling

$$\begin{bmatrix} \text{shiftx} & \text{scalex} \\ \text{shifty} & \text{scaley} \end{bmatrix} = \begin{bmatrix} 0.5 & 0.1666667 \\ 0.0 & 0.1111111 \end{bmatrix}$$

x = xraw*scalex + shiftx;
y = yraw*scaley + shifty;



(c) (d) (e) (f)

2.3.2 Advances in SLP : With basic architecture of single layer perceptron (SLP) (I# →H# →O#), different (Radial basis function, complex-SLP, fuzzy-SLP, Boolean-SLP, binary-SLP etc.) have emerged depending upon the type of input data (integer, real, complex, fuzzy, quaternion) (Chart 1).

Chart 1: Popular SLP/MLP NNs based on number of hidden layers and real/complex TFs/data

(a) Different SLPs based TF in hidden layer

NN	TF in H1	output	NN-$$
SLP	sigmoid	Linear	SLP-linear
	sigmoid	sigmoid	SLP-sigmoid
SLP	RBF	Linear	RBF
	RBF	Sigmoid	RBF-sigmoid
SLP	Rectangular	Linear	SLP-rectangular
SLP	Complex	Linear	SLP-Complex
SLP	Clifford	Linear	SLP-Clifford

(b) Different M(2) LPs based hidden layer

$$-MLP-NN	H1	H2
Centroid based-	Adaptive centroid	Sigmoid
Compositional-fuzzy-NN	TF-compositional($)	TF-compositional ($$)
Confluence operators : $ - inf(t,w) ; $$- sup(t,w)		

SLP	Quaternion	Linear or Quaternion	SLP-Quaternion(@@)
@@: Input can also be Quaternion			

(c) MLPs with different types of hidden layers			
MLP-NN-	H1	H2	output
	sigmoid	Sigmoid	sigmoid
With two hidden layers	RBF	Sigmoid	Linear
	RBF	Sigmoid	sigmoid
	RBF	RBF	Linear
	RBF	RBF	Sigmoid

The weight, TF and output data type also play a role. In fact, SLP with real values of input, Ws, TF and output is to be referred as real_valued_SLP. But, even if the TF alone is complex, while other components being real, this NN is called complex-SLP. But, all components being complex is an ideal complex SLP. The data flow in NN depends on the direction of connections between neurons in the layers. Fig. 1(a) incorporates the connectivity matrix for SLP (or RBF) of architecture 3-3-2. Each neuron in a hidden layer is connected to all the neurons in immediate preceding and succeeding layers. These connections are only in the forward direction. There are no inter-connections between the neurons in a layer. In 1958, Rosenblatt proposed inhibitory connection. The transfer function (TF) in all hidden layers is same usually. But a four layer RBF is recently proposed with radial basis transfer function in the first hidden layer and sigmoid in the second hidden layer. The architecture of RBF is arrived at automatically [29] with evolutionary methods viz. GA and genetic programming. Elman [13] proposed neural network architecture with partial reverse connections to deal with time-dependent tasks. The hybrid self-organizing-RBF-NNs [26] have been successfully employed to model data dynamic in time and space.

2.3.3 M(ultiple hidden) LP-NN : MLP-NN consists of one input layer, one output layer and several (two to four or more) hidden layers. The first layer with I# neurons and a bias neuron of constant (output) is the input layer and last layer with O# neurons is the output layer [18]. The direction of connections in MLP are described in Fig 1b with resulting NNs thereof. However, no study has been made for I - H1(Sigmoid-TF) – H2(RBF-TF) - O(Lin), I - H1(RBF-TF) – H2(RBF-TF) - O(Lin) NN. In some cases the TF in OL is sigmoid instead of a linear one in RBF-NN. In each layer, for every neuron an inner product of the incoming signals matrix and corresponding weight matrix is computed. The operation of TF on it results in the output of the neuron. It will be transmitted to all neurons in the succeeding layer. Generally the TFs of input and output layers are identity and linear respectively. The neurons in MLP are trained in a cooperative manner. In other words the weights of the neurons are adjusted at the same time with a goal of arriving at a network satisfying the convergence criteria.

The purpose of hidden layers is to develop heuristic rules, extract knowledge and affect non-linear transformation. A large variety of non-linear transfer functions are used in each neuron of the hidden layers. But 2-D, 3-D plots from each hidden layer are instrumental in probing into micro details of input to output (I/O) mapping. The number of PEs in each of the hidden layers is determined by trial and error, heuristic rules and semi/full automatic strategies.

$$y\_MLP = \left(Unscal\left[TFO\left(\left\{ \left(TFH2\left[\left\{ TFH1\left[\left(\left\{ scal\left(x^{T} \right) \right\} *WIH1 \right) \right]^{T} \right\} *WH1H2 \right] \right) * \ WH2O \right\} \right) \right] \right)$$

16

But, a variety of TFs (hard_limitor, sigmoid, radial basis function etc.) are implemented in output layer depending upon the task viz. classification, pattern recognition and function approximation. MLP-NNs are efficient tools for supervised classification and function approximation in addition to development of knowledge extraction systems in hybrid fuzzy-NNs. Fig 3 describes OOP format for representation of MLP and its implementation and GUI produces the figurative form of NN. The research mode module with typical architectures, TFs and neurons awaits software implementation and a part of futuristic integrated circuit (IC).

Fig 3. State-of-the-art-mode- MLP-NN

Input to Professional II

MLP

File	
Train	quad_trn.asc
Test	quad_tst.asc

	#PEs	**Lrn coefficient**
Input	2	0.300
Output	1	0.150

# Hidden layers	**2**

Bias	
Connected	yes

If #HL > 0

#HL	**#PEs**	**TF**	**Lrn coefficient**	
1	2	SG	0.300	
2	1	SG	0.150	

Learning rule- Output Layer
Back propagation
Conjugate gradient
Delta-Rule
…

Transfer function
Linear
TanH
…..

Research mode FF-MLP

No of hidden layers
#

Object fun
Kullback-Liebler information distance ESS

Transfer functions

NN	#layers
IO	0HL-NN
SLP	sHL-NN
M(2)LP	2HL-NN

Neurons	
Real	Non-linear
	Higher order
Complex	Quaternion
Quantum	

Architectures
Centroid-based MLP-NN
Dynamic architecture-2 (DA2-)-MLP-NN
Ito's method
Funahashi method
A bottleneck network
cascade correlation

Training algorithms **Training** Layer wise Subset W initialization Weight pruning	➢ Rule generation ➢ Incorporation of a priori knowledge ➢ Intelligent evolution ➢ Automation ➢ Self evolving
	SLP : Single-hidden layer-NN M(2)LP : 2-hidden layer-NN 0HL : zero-hidden layer NN [Neuron and PU or PE are synonyms]

The variations in time or space are similar in mathematical sense. Spatio temporal NNs tackle simple as well as complex response resulted in variation of both time and space. Recurrent NNs (Rec-NNs) on the other hand, excelled many of the hither to available non-linear time series models. Omlin and Giles [33] circumvented the short coming of NNs that they are opaque by proposing a way of extracting rules from trained recurrent model of discrete time series.

2.4 Hybrid models: Reliable and possibly robust one-step and multi-step- ahead forecast algorithms are the need of the hour. The hierarchical hybrid algorithms meet the requirement of course with the limitations of hardware, software expertise and cost at the disposal for a task. The residual time series after accounting for ARMA type models is analyzed with black box NNs. This approach is a hybrid model retaining the transparency of the classical models. At this juncture, experts intelligently use distribution/information/fuzzy characteristics of data/noise and methods from classical/advanced statistics and NNs in prediction. One of the trend setting packages is 'PREDICT' from neural ware corporation.

RESULTS AND DISCUSSION

3.1 Formation and destruction of ozone : The number of sources of primary pollutants viz. oxides of nitrogen (NO_x), volatile organic carbons (VOCs), non-methane hydrocarbons (NMHCs) is multifold [2]. The formation, transport and interaction with other gaseous pollutants in presence of UV-radiation are well understood processes. But, these processes are too complex to have a road map of the dynamics in time and vertical air column over different geo-regions and inter-phases.

VOCs escape into the atmosphere by evaporation from bulk use of organic liquids as well as in the process of distillation [49], which is rampant in countries with higher average temperature [3]. In urban areas, more than 50% of VOCs and NO_x, which are precursors of ozone, are from highway traffic —buses, cars, and trucks---- and industrial emissions. Ozone is a secondary pollutant formed by the chemical interactions of NO_x with VOCs, in presence of UV-radiation of sunlight under suitable conditions [2] [20]. This is an extremely reactive pollutant and scavenged by precursors, for example, in the thermal oxidation of NO to NO_2. Land-sea breezes in coastal areas also influence amount of ozone.

The study of ozone is categorized into surface level and troposphere with emphasis on global scenario as well as typical cities spread over different continents. In the case of ozone levels much care need to be exercised in situ in tackling incidences of very high ozone even for a brief period. The long term planning brings down the average ozone pollution level to meet first the local law and finally the global ideal air quality standards.

3.1.1 Surface level O_3: Surface level O_3 is the primary constituent of photo-oxidative smog. There is a gradual increase of the levels of surface ozone in the past two decades and more significant amounts in Korea [20]. Natural, anthropogenic emissions and meteorological conditions have non-linear effect on content of ozone. The effect of ambient ozone at 14 sites in eight European countries was studied on White Clover. Twenty-one input parameters were considered to predict biomass change which has an influence on ozone. Even complex MLR models are inferior in prediction compared to NN model with 5 inputs. The success of NNs in predicting hourly ozone [10] in Spain is clear from table 1.

Table 1: Performance measures for hourly forecasting of ozone in Bilbao

Location	Forecast	O_3 (t+1) R	FA2	FB	FV
Elorrieta	MLP1	0.890	0.954	0.029	0.106
	MLP2	0.905	0.972	0.017	0.051 check
	LR	0.880	0.948	0.017	0.080
	Persistence	0.890	0.965	0.000	0.004
Mazarredo	MLP1	0.854	0.961	0.022	0.067
	MLP2	0.889	0.972	0.022	0.077
	LR	0.838	0.938	0.022	0.072
	Persistence	0.865	0.966	0.001	0.001

Inputs			
2	Ozone conditions		r^2
2	Temperature	Tr	0.84
1	Semi-urban/rural	Te	0.71

3.1.2 Troposphere ozone : The concentration of troposphere ozone depends primarily upon natural and anthropogenic processes and meteorological conditions [27]. It is in fact, a result of several complex processes including regional transport of O_3 as well as its precursors. Over the last forty years, there is a two-fold increase of troposphere ozone and NNs predicted well [37].

3.1.3 Global scenario of ozone : A perusal of historical data reveals [15, 16] that there is a reduction of precursor emissions in the last thirty years reflecting the reduction of ozone by 12 ppb. In Ireland (Mace Head), there was a control of CO, NO and non-VOC (non-methane) during 1987 to 1995. It reflected in a downward trend (-0.39 ppb/year) of surface ozone. Increase of the ground level ozone, a constituent of photochemical smog, is one of the primary air quality issues in developing as well as developed countries of the world.

The standard for national ambient air quality prescription of ozone (235 $\mu g/m^3$) in an area in USA is said to be attained, only if the ratio (of number of days on which maximum hourly average exceeds 235 $\mu g/m^3$ to the expected number forty three of days per calendar year) is equal to or less than one, for a three-year period continuously. The attainment (compliance) or non-attainment is crucial in getting support from U S government. A labeled non-attainment area is levied with high costs for auto-inspections. The funds for highway will be decreased and sanctions will be critical on industries and commercial expansion.

3.1.4 Ill effects of ozone on human health and plant kingdom : Ozone is one of the most influential pollutants of the air diminishing the quality with a consequent effect on human health and adverse effects on plant kingdom. It increases chronic respiratory infection, globe ocular damage and lung inflammation.

Further, ozone aggravates asthma, impairs lung defense mechanism and reduces immunity in the human body. High levels of surface as well as tropospheric ozone have potential hazard for human health, advances the death of vulnerable people by some days (called harvesting effect), especially those suffering from acute respiratory problems including asthma [9]. The effect on children is tremendous particularly during high pollution days in summer time [20]. The ill effects are explained as a result of visible foliar injury and physiological impairment. Ground level ozone has adverse effect on respiratory health of children of age between 5 and 7. The forecast announcement in the city is of vital importance for those at risk of respiratory distress. However, if care is taken for the reduction of SO2 and PM, a drastic improvement of health conditions and considerable decrease in morbidity and mortality was observed.
Ozone affects vegetation and ecosystem adversely. The ill effects of tropospheric ozone are decline of many forest trees, reduction in growth and yield of crops**.** In vegetation, it reduces growth/yield/bio-mass production, acute foliar/leaf injuries, and shifts the competitive advantages of plant species in mixed population. The leaf senescence is considered as one of the causes. The plants are more sensitive to low NO_x, medium range temperature and high 24 hrs mean ozone [31]. The reduction in the growth/yield of agricultural crops and forest is remarkable. NNs are better predictors when plant-environment interactions are non-linear.

- Remedial measures for ill effects of ozone : In fact remedial measures are not simple, since amount of ozone depends on geographic location, meteorological profile, concentrations of precursors/primary pollutants, industrial activity and traffic [17]. Thus, ozone pollution cannot be alleviated independently without regulating many parameters with an integrated approach [3].

3.2 Complexity of ozone modeling : In general, interactions of pollutant with atmosphere are complex; much more so for ozone. The formation steps of ozone/precursors, emission, atmospheric transport, mixing, photochemical path, decomposition and reactions of ozone/its precursors [3] are all dynamic/non-linear/non-stationary processes. Thus, the complete picture of ozone cycle is complex leading to collapse of the direct cause-effect models. The relationships between smog duration and ozone levels/meteorological factors are highly involved processes [49], which are composite in concentration, time and geological-space. This renders even the modeling (leave alone prediction) from first principles is out of reach, leaving the only option of non-collapsing data driven paradigm with maximum possible robust performance.
The information of interest to governmental agencies regarding ozone is numerical values (index) forecasted daily maximum, especially in summer/winter and false positive/negative alarms of the predicted model. The former is for the purpose of alerting the local public in advance and to take up short-/long-term control (if not prevention) measures. The governess announces the excedence limits of pollutant species based on scientific consolidated stipulations of EPAs and informs forecasted alarming situations of the day in advance to civic public. Keeping in view of the ill effects on human health, one-, two day-ahead forecast of possible excedences in levels of ozone, SO_x, NO_x, etc. especially in cities with high population and traffic. It helps to divert vehicular traffic, announce local holiday for schools and alert the old/disease prone/low- immunity individuals to avoid long exposure to high polluted atmosphere outside the home. The later is for the scientific personal to redesign the frequency of spatio-temporal monitoring schedules and implementing advances in analysis and modeling. The progress in monitoring schedules/methods, pollution indices, control measures during the last two decades are noteworthy The thrive at the moment is to enhance the frequency of sampling, precision and accuracy of observed data, picking up of relevant variables and implementing the latest advances from cradle to prediction of modeling.

3.3 Modeling and forecast of ozone: NN predictive models for ambient pollutant concentration and extreme have been compared with reports using classical statistical (ARIMA, MLR) and mechanistic models covering all over the world. This endorses the utility of the paradigm even under varying meteorological and terrain conditions. The size and complexity of the models differ on size of the data, input variables, performance measures and above all the goals set in the study. The case studies as well as literature reports described are selected with an objective of covering diverse situations. It is a reinforced learning strategy to enforce a long-term memory (LTM) impression only on sustaining methods automatically. This helps to have broader perception of the influencing factors (proven to be successful in specific case studies) and to be pruned for his/her study area. If a test is successful in more than 60% criteria, it can be accepted. The frequently used statistical performance measures are included.

3.3.1 Ground level and tropospheric ozone : Modeling, a statistical/mathematical looking at the results with only a fraction of the entirety of the scene generates complimentary information, endorses/refutes the lookup table picture. Model driven techniques widely available even in free to high end complex software packages have been in wide use during the last half a century. Many of the earlier studies in ozone modeling were confined to local data sets with trodden linear statistical/ARMA modeling techniques [40]. Recursive modeling was used for non-stationary time series, which allows variation in seasonal adjustment as well self-adaptive implementation of state space forecasting. Further, change of dynamics and mild non-linear behavior is taken care in the time variable parametric models [49]. Data driven-, distribution free-, automated-, self-organizing- and evolving- nature inspired algorithms changed the scenario especially in late 1990's. Although fuzzy expert systems and NNs generally improve the predictive accuracy, forecast of O_3 level where breeze has a profound influence is not up to the mark. Abdul-Wahab et al. [2] employed SLP-NN to predict tropospheric ozone concentration with the inputs, CH_4, NMHC, CO, CO_2, NO, NO_2, SO_2, WS, WD, temperature, RH, solar energy and suspended dust. Due to large CPU times required, different runs were performed with very low to large training set (Table 2).

Table 2: Comparison of different NN models for prediction of tropospheric ozone in Kuwait

Statistic		Run 1	Run 2	Run 3
	Tr	4078	1386	16
	Te	719	244	3
MSE		6.619	14.030	20.998
R^2		0.990	0.981	0.956
Epochs		13 427	19 384	1 054 327
CPU time (hr:min:sec)		9:30:01	8:30:05	1:22:33

Software : NeuroShell MSE: Mean Squared Error

3.3.2 Modeling high ozone event: The forecast of maximum and excedence incidents of ozone are of prime concern to governmental officials to diminish the morbidity of civilians especially of children, high risk/vulnerable adults and aged persons. Accurate prediction of ozone levels at least one day ahead are necessary for public warning and to control the consequences in a slow pace [17]. Summer time ground level ozone [8] exceeds one-hour threshold of 180 $\mu g/m^3$ [43]. The variables relevant to traffic were introduced to model ozone in urban area, considering a rural point as a background. But, the model failed due to the incorrect assumption that 'degree of cleanliness of rural site remained same'.

Case studies

Many of the earlier studies in ozone modeling are confined to local data sets with conventional modeling techniques [40]. Further, only maximum ozone concentration in a day was documented, but duration of time at the peak concentration is the critical parameter to alleviate health hazards. The successful literature reports in modeling and forecasting [14] of air pollutants and ongoing automated real time implementations follow.

In yesteryears, the results of NN were compared [9] with only simple models viz., persistence, ARIMA. The latter methods however do not function well for O_3 prediction. The prediction of daily maximum was studied [28] with LR and NNs for 8 cities in U.S. during May to September. It is reported that both models under predict the high O_3 events.

The models of ozone in Athens, UK, Europe, Dallas at ground level, U.K, Austria, Chile, Korea at surface level and Finland in troposphere with NNs have good predictive ability. Daily maximum ozone [22] is correlated with wind speed, temperature (T) and pH. The models used are MLR, non-linear-multivariate regression, generalized additive model and NNs, which imbibe non-linear relationship involved in O_3 formation. Gómez-Sanchis et al [17] (2006) envisaged the need of employing more relevant variables for futuristic studies. NNs were employed for k-hours ahead (k = 1, 2 …8) forecast of O_3 and NO_2 levels in Bilpao area. Chaloulakou et.al [8] (2003) made an exhaustive comparison of forecasting daily maximum O_3 (Table 3a) in different countries over a long period of time, endorsing the application of NNs in any meteorological zone. The modeling in different locations of Athens is targeted to predict the excedence days (Table 3b) as well as quantitative evaluation of ozone concentration (Table 3c). Abdullah [1] (2009) compared Fisher (LDF, Linear Discriminant Function, Quadratic DF), REGF (REGression Function), MLP-BP-NN and RBF-NN to forecast possible shifts in the prediction of ozone in Houston for a 12 year period (1990 to 2002 daily date) air quality data from Texas commission on environmental quality. This is an extensive comparative study while the earlier studies used ARIMA and regression models.

Table 3a: Prediction of probability of detection of daily maximum ozone in different European counties and USA using NNs

Country	Sites	Time	Input		RMSE %	POD	Year of publication
			M	P			
USA	8	1991-1995	4	1	16-30	NA	1997
USA	7	1998-1999	6		19	0.42	2000
UK	5	1993-1996	2	1	16-34	0.13	1999
Italy	3	2000	3	1	13-24	0.43-0.61	2001
Athens	1	1998	5	3	23	0.67	2000
Athens	4	1992-1999	8	3	17-25	0.7-0.76	2003

POD: probability of detection M(eteorological) P(ersistence)

Table 3b: Prediction of incidence of high (>180 $\mu g / ml^3$) ozone in Athens

Site	Number of days	Prob of detection		Success Index	
		NN	LS	NN	LS
Liossia	56	0.75	0.64	0.63	0.54
Maroussi	82	0.67	0.44	0.50	0.33
N.Smirni	18	0.00	0.13	0.00	0.11

Table 3c: *Comparison of performance of NN with LS in prediction of ozone at different sites in Athens*

Site	*RMSE T3*

	Va		*Te*	
	NN	*LS*	*NN*	*LS*
Liossia	29.6	35.2	29.6	0.9
Maroussi	29.0	31.9	29.0	28.1
N.Smirni	29.5	31.3	39.4	42.7

Gardner and Dorling [15, 16] modeled maximum surface ozone concentration in five rural sites of UK (Table 4) over the period 1979 to 1997. The objective was to pinpoint non-local trends in controlling urban vehicular emissions. The high values of RMSE and corresponding low magnitudes of R^2 for Yarner Wood and Sibton stations are explained due to missing data and larger distance between ozone monitoring and meteorological station respectively.

Table 4 : MLP NN for Maximum ozone concentration at different locations

Statistic	Eskdalemuir		Harwell		Lullington Heath		Sibton		Yarner Wood	
	Te	Va	Te	Va	Te	Va	Te	Va	Te	Va
RMSE	7.31	8.32	10.42	13.58	9.32	8.40	12.15	11.11	9.43	10.34
*R*2	0.57	0.67	0.63	0.63	0.70	0.62	0.52	0.66	0.61	0.27
Input variables										
Daily maximum temperature	Total daily sunshine (a surrogate for daily maximum solar radiation which was unavailable)				Mean daily wind speed		Vapor pressure		Total cloud cover	

Gardner and Dorling [15, 16] emphasized MLP-NNs are indispensable to model hourly ozone concentration at surface level although regression trees are physically interpretable (Table 5). The interactions and non-linear trends of meteorological and temporary predictors are captured accurately by the two sigmoid hidden layers of MLP.

Table 5: Comparison of NNs with regression trees and linear regression at different stations in UK

Site	RMSE		
	MLP (12:20:20:1)	Reg tree	MLR
Bristol	6.60	8.42	7.70
Edinburgh	6.80	8.16	7.98
Eskdalemuir	6.93	8.04	7.98
Leeds	6.55	7.49	8.48
Southampton	6.83	8.61	7.80

Input

MET

$$\sin\left(\frac{2*\Pi*d}{365}\right) \qquad \cos\left(\frac{2*\Pi*d}{365}\right)$$

$$\sin\left(\frac{2*\Pi*h}{24}\right) \qquad \cos\left(\frac{2*\Pi*h}{24}\right)$$

d: day H: hour

3.4 Automatic predictive computational tools: The purposes of automatic predictive computational tools in the case of ozone pollution are

- Classifying the day as polluted or non-polluted and
- To predict maximum ozone on a given day based on previous observed ozone data and forecasted metrological factors of the day

In this decade, efforts have been made for the computerization without human intervention and now countable number of automated pollution centers is in operation all over the world. Dutot et al. [12] reported that NEUROZONE model was implemented in real time in Orleans region. The NN is retrained with recent past data with same architecture. The changes are only in Ws, assimilating most recent trend also. The model predicts 24-hour ahead ozone concentration and future software implementations include longer lead times.

Out of seven incidents of ozone level excedence, six have been predicted by NN model affectively. The success index is 78% against a false alarm of 8%. NEUROZONE, an NN model for ozone prediction has been implemented in real time in the Orleans region. At the end of September of each year, the validated data is introduced into the training data, which was already employed to develop the model. Every year the values of the weights are re-estimated, which leads to increased reliability of prediction as time progresses. It is a natural consequence of increase in data, which obviously increases the frequency of extreme ozone values. The initial teething problems giving an impression of the questionability of the model will fade away. It is but natural that within frequency of catastrophic episodes the functioning of the model is with low success rate and with significant false positives and false negatives. The present forecasting time is 24 hours. The system will be tuned for longer lead-time predictions. It is a powerful tool in pollution management systems.

3.5 Failure of forecast models for ozone : Some of the factors for the failure of forecast models [9] are described in the chart 2.

Chart 2 : Typical factors causing failure of ozone (short term) forecast

⇒ *Some of the input parameters cannot be predicted in the morning Ex. NO*

⇒ *Training data are from limited or selected spatio-temporal grid. Thus, it may not be the representative of real trends over a larger scale.*

⇒ *The predictive model uses the previous day's observed metrological data. But, the forecast values function better*

⇒ *Most of the results are from ozone hind casts rather than true forecast*

⇒ *The performance is judged on overall statistics, but not extreme value statistic. But, the extreme values of ozone alone are crux of the problem.*

A comparison of NN models with other classical ones are briefly tabulated (Tables 6-9).

Table 6: Comparision of NN models in predicting ozone at different stations Tr set : (years 1998~2001)

Test set year 2002	Model	R2	RMSE	IA
Cutin	SOM + MLP	0.69	0.30	0.90
	MLP	0.66	0.32	0.89
	Clustering + MLR	0.58	0.35	0.85
	MLR	0.37	0.42	0.70
Chungming	SOM + MLP	0.63	0.31	0.87

MLP	0.59	0.32	0.85
Clustering + MLR	0.55	0.32	0.83
MLR	0.50	0.35	0.79

Table 7: Performance of NNs in comparison with LR for O_3 in different locations [Bilbao (Spain)]

Location	Forecast	O_3			
		MLP1	MLP2	LR	Persistence
Elorrieta	t+1	0.0006	0.0005	0.11	0.12
	t+4	0.000	0.002	0.48	0.54
Mazarredo	t+1	0.0002	0.0004	0.12	0.13
	t+4	0.04	0.01	0.53	0.70
Txurdinaga	t+1	0.00004	0.00008	0.09	0.10
	t+4	0.01	0.03	0.45	0.58
Deusto	t+1	0.20	0.003	0.11	0.12
	t+4	0.04	0.04	0.43	0.51

Table 8: Quarterly and annual Performance of ozone model in the validation step

Model	RMSE $(\mu g/m^3)$				
	I Trimester	II Trimester	III Trimester	IV Trimester	Annual
Time series (TS)	19.51	10.48	10.46	22.84	14.47
MLR	21.19	10.16	7.65	12.11	18.15
SLP	17.40	9.33	11.04	16.93	12.74

SLP : 12-10-1 TF: sigmoid
SLP >> MLR for prediction; SLP = MLR for training

Table 9: Comparison of accuracy of ozone prediction by NNs with chemical transportation model

Method	RMSE	d	FAR	SI
MLP1	15	0.92	0.06	0.65
MLP2	18	0.92	0.17	0.69
MLP3	18	0.92	0.14	0.72
CHIMERE	27	0.87	0.02	0.55
Persistence	20	0.88	0.07	0.64

Schlink et al. [40] employed a deterministic approach to model ozone formation, from hydrocarbon and NO_x amounts in a boundary layer of pollutant. Further, seasonal variation of hydroxyl radical, temperature and mixing heights are included. The dry depolarization velocity of ozone to the land surface was assumed to be constant. The primary objective of hybrid model of Cobourn et al. [9] is to forecast one hour peak ozone content approaching or exceeding 120 ppb. The adjunct model, a subset of the hybrid model performs better than the standard model on the days where the meteorological conditions are conducive to high ozone level. It is successful to forecast high ozone events in Louisville, KY. Heo and Kim [20] employed two types of forecast models. The first uses the fuzzy expert system to predict the appearance of ozone >80ppb, while, the second system forecasts the maximum ozone on the following day. Ozone models are reported for Lyon (France) in table 10, based on NO_2, O_3, and T (1 to 10 classes) and decision label (polluted/non-polluted day).

Table 10: Average max daily ozone (ppb) within meteorological regimes at different monitoring stations

Chiayi			Chianjin		
Code	Number	Average	Code	Number	Average
B1	420	73.9	H1	525	85.8
B7	452	68.7	H4	524	74.8
B2	255	56.7	H3	266	58.2
B6	151	48.1	H2	332	54.5
B5	276	45.8	H5	120	32.3
B3	152	30.0			
B4	87	26.7			
		57.6			
	1793	57.6		1767	68.9

Cutin			Chungming		
Code	Number	Average	Code	Number	Average
C3	390	78.0	A4	330	75.3
C2	210	61.7	A6	304	66.2
C7	247	54.6	A5	427	57.5
C4	355	38.6	A1	291	50.6
C1	305	36.4	A2	239	42.2
C6	191	30.1	A3	191	33.6
C5	92	29.9			
	1790	50.4		1782	56.5

Means are indistinguishable by Waller–Duncan k-ratio t-test (k =100, P=0.05)

Before the year 2001, the predictive process of photochemical pollutants (NO_2, O_3 etc.) was based on only human expertise in Athens (Greece) as was the practice in most other cities. A fast short-term (day-to-day) prediction is the basic need. When an air pollution episode is predicted, a complex prediction method in real time (outputting the results within few hours) has to follow for the same location that may be operated in a sophisticated central place. The urban air pollution in many metropolitan cities like Los Angeles, Mexico and Athens etc. [7] is high.

Statistical models function in cause and effect framework. Ozone conducive meteorological conditions are well known now. The photochemical smog is very high during summer. It is due to the high temperature, high insulation, high stability, low mixing heights and low mid day relative humidity. Many of the ozone episodes occur when the listed conditions prevail.

FUTURE SCOPE

During the last one-decade the MNNMs are compared mostly with MLR and rarely with non-linear models. Instead, comparison amongst different NNs (MLP, RBF, Fuzzy-ARTMAP, Kohonen etc.) with advanced training algorithms is a welcoming feature to uphold the superiority of data driven techniques over yester years model driven activity. The limitations of model driven mathematical/statistical procedures have given way to nature inspired algorithms, a paradigm shift in information extraction from huge spatio-temporal data. With E-man the data can be discrete/continuous and even multi-way. Currently the prospects of loose coupled sequential/hierarchical/hybridizing/cross disciplinary areas

realized the retention of the advantages of both the worlds (mathematical science, nature inspired paradigm) with synergism and minimized/nullified limitations.
Future ventures with FFNNs, Rec. NNs using the state-of-the-art algorithms and encapsulating the modules in sequential and hierarchical mode will start a new era in handling TS data retaining the desired characteristics and eliminating limitations of the component procedures. Simulated data sets planned with statistical experimental design, signal-to-noise (S/N) ratio and information content of various complexities in trend, seasonality, spikes and noise structure result in feasibility study of mega NN structures. The ensemble study offers robustness for perturbation.

ACKNOWLEDGEMENTS

The authors profusely thank editor-in-chief to invite for the series of state-of-art-reviews (SAR) on Mathematical Neural Network (**MaNN**) Models with a focus of research tutorials for future post-graduate education. In authors' laboratory, BrainMaker software was used in 1993 in the NN-model for calibration of antihypertensive drug labetelol by spectrophotometric data. The demonstration disk of Professional II from Neuralware Corporation kindled interest in the group. We thank DOD for financial support to purchase Professional II and Trajan for a project on modeling in fisheries during the period 1998-2001. RSR profusely expresses his gratitude to DRDO, Gwalior for NN modeling of gas chromatographic response with molecular descriptors (2000-2001) and DOEn for finances for NN modeling for bauxite ore benefaction in 2008. We thank our researchers, co-investigators and participants in training/lecture programs. This series of tutorial reviews are culmination of continuous efforts of our group to probe in to details of theory, software, applications of neural networks and their hybrids with statistical methods/soft-models and nature inspired algorithms over a period of two decades in chemical/environmental/pharmaceutical sciences and fisheries.

REFERENCES

[1] A.B.M. Abdullah, D. Mitchell, R. Pavur, Manage. *Environ. Qual.*, **2009**, 20, 73-81.
[2] S. A. Abdul-Wahab, S.M. Al-Alawi, *Environ. Model. & Software*, **2002**, 17, 219–228.
[3] E. Agirre-Basurko, G. Ibarra-Berastegi, I. Madariaga, *Environmental Modelling & Software,* **2006**, 21, 430–446.
[4] A. Braibanti, G. Nageswararao, S. B. Jonnalagadda, D. Sudarsan, R. Sambasivarao, *Ann. Chim. (Rome)*, **2001**, 91, 29-39.
[5] A. Braibanti, R. Sambasivarao, V. AnantaRamam, G.Nageswararao, V.V. Panakala Rao, *Ann Chim*, **2005**, 95, 291-301.
[6] A. Braibanti, R. Sambasivarao, A. Ravindrababu, G. Nageswararao, *Annali di Chimica,* **1995**, 85, 17-29.
[7] M. Breiling, J. Alcamo, *IIASA Technical report of Leader project, Laxenburg, Australia*, **1992**.
[8] A. Chaloulakou, M. Saisana, N. Spyrellis, *The Science of the Total Environment*, **2003**, 313, 1–13.
[9] W. G. Cobourn, L. Dolcine, M. French, M. C. Hubbard, *J. Air and waste manage. assoc.,* **2000**, 50, 1999-2009.
[10] A. Coman, A. Ionescu, Y. Candau, *Environmental Modelling & Software*, **2008**, 23, 1407-1421.
[11] S. Das, SayanMaity, B. Qu, P.N. Suganthan, *Swarm and Evolutionary Computation*, **2011**, 1, 71–88.
[12] A. Dutot, J. Rynkiewicz, F. E. Steiner, J. Rude, *Environmental Modelling & Software*, **2007**, 22, 1261-1269.
[13] J. L. Elman, *Cogn. Sci.*, **1990**, 14, 179–211.
[14] Yu. Feng, W. Zhang, D. Sun, L. Zhang, *Atmospheric Environment*, **2011**, 45, 1979-1985.

[15] M.W. Gardner, S.R. Dorling, *Atmospheric Environment*, **2000**, 34, 171-176.
[16] M.W. Gardner, S.R. Dorling, *Atmospheric Environment*, **2000**, 34, 21-34.
[17] J. Gomez-Sanchis, J. D. Martín-Guerrero, E. Soria-Olivas, J. Vila-Frances, J. L. Carrasco, S. Valle-Tascon, **2006**, *Atmospheric Environment*, 40, 6173-6180.
[18] P. Guo, M.R. Lyu, *Neurocomputing*, **2004**, 56, 101–121.
[19] P. Hájek, V. Olej, *Ecological Informatics*, **2012**, 12, 31-42.
[20] J.S. Heo, D.S. Kim, *Science of the Total Environment*, **2004**, 325, 221–237.
[21] V.S. Jagesh, C. Poon, *IEEE Trans. Neural Networks*, 10, 10-18.
[22] H. Jorquera, R. Perez, A. Cipriano, A. Espejo, M.V. Letelier, G. Acuna, *Atmospheric Environment*, **1998**, 32, 3415-3424.
[23] E. Kalapanidas, N. Avouris, *Environmental Modeling & Software*, **2001**, 16, 263–272.
[24] A.A.M. Khalaf, K. Nakayama, *IEICE Trans. Fundamentals*, **1999**, E82A (8), 1420-1427.
[25] D.R. Lawson, *J. Air & waste ManagAsso*, **1990**, 40, 156-165.
[26] J. Lian, Y. Lee, S. D. Sudhoff, S.H. Zak, *IEEE Trans. Neural Netw.*, **2008**, 19, 460-474.
[27] C. Libiseller, A. Grimvall, *Atmospheric Environment*, **2003**, 37, 3923-3931.
[28] H. Lu, J. Hsieh, T. Chang, *Atmospheric Research*, **2006**, 81, 124-139.
[29] D. Manrique, J. Rios, A. Rodriguez-Paton, *NeuroComputing*, **2006**, 69, 2268–2283.
[30] W.S. McCulloch, W. Pitts, *Math. Biophy*, **1943**, 5,115–133.
[31] G. Mills, G. Ball, F. Hayes, J. Fuhrer, L. Skarby, B. Gimeno, L. De. Temmerman, A. Heagle, *Environmental Pollution*, **2000**,109, 533-542.
[32] G. Nageswara Rao, K.Sambasiva Rao, R. Sambasivarao, H.A.R. Laila, *J. Ind. Council. Chemists*, **2000**, 17, 48-65.
[33] W. Omlin, Christian, C. Lee. Giles, *Neural netw.*, **1996**, 9, 41-52.
[34] J.B. Ordieres, E.P. Vergara, R.S. Capuz, R.E. Salazar, *Env. Model. & Software*, **2005**, 20, 547-559.
[35] S. Owega, G.J. Evans, R. E. Jervis, M. Fila, R. D'Souza, B. Khan, *Atmospheric Environment*, **2004**, 38, 5545–5553.
[36] D.E. Rumelhart, J.L. McClelland, *MIT Press, Cambridge, MA*.**1986**.
[37] E. Salazar-Ruiz, J.B. Ordieres, E.P. Vergara, S.F. Capuz-Rizo, *Environ. Model. & Software*, **2008**, 23, 1056-1069.
[38] K. Sambasiva Rao, R. Sambasiva Rao, *J. Korean Phys. Soc.*, **1998**, 32, S1850-S1851.
[39] R. Sambasiva Rao, G. Nageswara Rao, *Himalya Publisher*, **2005**, New Delhi (India).
[40] U. Schlink, S. Dorling, E. Pelikan, G. Nunnari, G. Cawley, H. Junninen, A. Greig, R. Foxall, K. Eben, T. Chatterton, J. Vondracek, M. Richter, M. Dostal, L. Bertucco, M. Kolehmainen, M. Doyle, *Atmospheric Environment*, **2003**, 37, 3237–3253.
[41] T. Slini, A. Kaprara, K. Karatzas, N. Moussiopoulos, *Environ. Model. & Software*, **2006**, 21, 559-565.
[42] G. Soja, A-M. Soja, *Atmospheric Environment*, **1999**, 33, 4299-4307.
[43] S.I.V. Sousa, F.G. Martins, M.C. Pereira, M.C.M. Alvim-Ferraz, **2006**, *Chemosphere*, 64, 1141-1149.
[44] M.V. Subba Rao, V. AnantaRamam, V. Muralidhara Rao, R. Sambasiva Rao, *Asian J Chem*, **2010**, 22, 5937-5950.
[45] I. Suryanarayana, A. Braibanti, R. Sambasiva Rao, V. AnantaRamam, D. Sudarsan, G. Nageswara Rao, *Fisheries Research*, **2008**, 92, 115-139.

[46] K. Viswanath, R. Sambasiva Rao, Ch V. Kameswara Rao, K. Rama Krishna, B. Rama Krishna, G E G. Santhosh, *J. Applicable Chemistry*, **2012**,1, 109-124.

[47] B. Yuan, X.Z Wang, T. Morris, *Waste Management*, **2000**, 20, 677-686.

[48] A. Zhou, B. Qu, H. Li, S. Zhao, P.N. Suganthan, Q. Zhang, *Swarm and Evolut. Comput.*, **2011**, 1, 32–49.

[49] A. Zolghadri, M. Monsion, D. Henry, C. Marchionini, O. Petrique, *Environ. Model. & Software,* **2004**, 19, 369–382.

[50] R Sambasiva Rao, *DRDE project report, Gwalior,* **2001.**

[51] L.H.A Rehman, K Sambasiva Rao, G. Nageswara Rao, R. Sambasiva Rao, *J. Ind. Council. Chemists,* **2001,** 17, 33-50.