Research Article
Artificial Intelligence based Solver for Governing Model of Radioactivity Cooling, Self-gravitating Clouds and Clusters of Galaxies

[1,2]Junaid Ali Khan and [1,2]Muhammad Asif Zahoor Raja
[1]Center of Computational Intelligence, P.O. Box, 2300, Islamabad, Pakistan
[2]Department of Electrical Engineering, COMSATS Institute of Information Technology, Attock, Pakistan

Abstract: In this study, a reliable alternate platform is developed based on artificial neural network optimized with soft computing technique for a non-linear singular system that can model complex physical phenomenas of the nature like radioactivity cooling, self-gravitating clouds and clusters of galaxies. The trial solution is mathematically represented by feed-forward neural network. A cost function is defined in an unsupervised manner that is optimized by a probabilistic meta-heuristic global search technique based on annealing in metallurgy. The results of the designed scheme are evaluated by comparing with the desired response of the system. The applicability, stability and reliability of the proposed method is validated by Monte Carlo simulations.

Keywords: Artificial neural networks, Monte Carlo simulation, simulated annealing, singular non-linear system

INTRODUCTION

In the recent decades, the rapid development in merging artificial intelligence with nonlinear science has made ever increasing interest of physicists, scientists and engineers (Koumboulis and Mertzios, 2005; Mertzios, 2000). In particular the systems that contain singularity and nonlinearity simultaneously are considered to be the challenges for the researchers. These systems arise in the modeling of extensive applications including physical structures, random processes, control theory and real dynamical systems (Costin and Costin, 2001; Parand and Taghavi, 2008) etc. The singular non-linear system under consideration is used in modeling of radioactivity cooling, self-gravitating clouds and in the clusters of galaxies (Herisanu *et al.*, 2008) and governing mathematical expression in the form of second order Initial Value Problem (IVP) is given as:

$$\begin{cases} \dfrac{d^2 y(t)}{dt^2} = -\dfrac{2}{t}\left(\dfrac{dy(t)}{dt}\right) - y^\lambda & 0 \le t \le T \\ y(0) = 1, \ \dot{y}(0) = 0, \end{cases} \tag{1}$$

where λ is the polytrophic index and its value is taken as $\lambda = 5$, the variation of input span $t \in \Re$ between (0,T).

A substantial work (Liao, 1992; Marinca *et al.*, 2008) has been done to provide the analytic as well as numerical solution of the given system (1) in a limited domain. The solutions provided by the traditional methods has limitations such as the validity of the

results on predefined discrete grid of inputs and for small perturbed response of the system requires complex iterative procedure. Due to this these classical techniques are not effective for models that based on real time processing. In this regard, the real benefit of the ANN can be exploited as they provide the results on the continuous grid of time, results can be obtained readily on any input point without repeating the whole chembersome procedure and obviously the time and space complexity is incredibly low (Saloma, 1993; Fogel, 2006). Thus the weaknesses of the classical method can be the strengths for the artificial intelligence techniques, although lots of work has been done in the area of traditional techniques (Dehgham and Shakeri, 2008; Chowdhury and Hashim, 2007; Shawagfeh, 1993; Richardson, 1921; Monterola and Saloma, 1998) as well as computational intelligent techniques (Khan *et al.*, 2011a; Raja *et al.*, 2011a; Fogel and Ghozi, 1997; Khan *et al.*, 2011b; Raja *et al.*, 2011b) but no one has yet tried to optimized the given singular non-linear system via ANN optimized with Simulated Annealing (SA).

The study presents a machine learning procedure based on SA incorporated with ANN to optimize the non-linear singular systems of radioactivity cooling, self-gravitating clouds and clusters of galaxies. The ANN mathematical modeling for singular non-linear system is performed by the linear combination feed forward neural network with log-sigmoid as the activation function in the hidden layers. The designed weights of the ANN are optimized by a meta-heuristic probabilistic algorithm so called SA that belongs to the class of an efficient constraint optimization problem

Corresponding Author: Junaid Ali Khan, Center of Computational Intelligence, P.O. Box, 2300, Islamabad, Pakistan

solving algorithms. The designed scheme is applied on the given IVP and the results are compared with exact response of the system. A sufficient large number of independent runs of the proposed solver are used to get a comprehensive statistical analysis describing the reliability, applicability and effectiveness of the given scheme.

PROPOSED METHODOLOGY

In this section, we describe the designed mathematical modeling based on neural networks along with formulated fitness evaluation function so called cost function. The learning procedure based on hybrid computation is also narrated to get optimum weights of the model.

Artificial neural network mathematical modeling: The feed-forward ANNs are the most popular architectures due to their universal function approximation capabilities (Rarisi *et al.*, 2003; Raja *et al.*, 2010; Tsoulos *et al.*, 2009), structural flexibility and availability of a large number of training methods (Khan *et al.*, 2011c). An arbitrary continuous function and its derivatives on a compact set can be arbitrarily approximated by feed-forward ANN with multiple inputs, single output, single hidden layer with a linear output layer having no bias (Aarts and Veer, 2001). An approximate mathematical model has been developed using feed-forward ANN. For this, following continuous mapping is employed based on the solution $y(t)$ and its first $dy(t)/dt$ and second order derivative $d^2y(t)/dt^2$, respectively:

$$\begin{cases} \hat{y}(t) = \sum_{i=1}^{m} \alpha_i f(w_i t + b_i) \\ \dfrac{d\hat{y}(t)}{dt} = \sum_{i=1}^{m} \alpha_i \dfrac{d}{dt} f(w_i t + b_i) \\ \dfrac{d^2 \hat{y}(t)}{dt^2} = \sum_{i=1}^{m} \alpha_i \dfrac{d^2}{dt^2} f(w_i t + b_i) \end{cases} \qquad (2)$$

where α_i, w_i and b_i are real-valued bounded of ith adaptive parameters, m is the number of neurons and f being the transfer function taken as log-sigmoid $f(t)=1/(1+e^{-t})$. The differential equation neural networks (DE-NN) architecture of system in (1) is modeled with the linear combination of the networks provided in the Eq. (2). The generic form of DE-NN architecture is shown in Fig. 1.

The Cost Function (CF) is formulized in an unsupervised manner as the sum of error of the governing DE of the subsequent model and its initial conditions.

$$CF = \frac{1}{N+1} \sum_{m=0}^{N} \left(\frac{d^2 \hat{y}(t_m)}{dt^2} + \frac{2}{t_m} \frac{d\hat{y}(t_m)}{dt} + (\hat{y}(t_m))^5 \right)^2$$
$$+ \frac{1}{2} \left\{ (\hat{y}(0) - 1)^2 + \left(\frac{d}{dt} \hat{y}(0) \right)^2 \right\}, \qquad t \in (0,1) \qquad (3)$$

Here inputs $t \in (t_0 = 0, t_1 = 0.1,\ldots,t_N = 1)$ with a N number of steps, $\hat{y}$, $d\hat{y}/dt$ and $\hat{d}^2 y/dt^2$ are the NN models as given by the equations in set (2). As the value of the CF approaches zero that is subject to finding the optimal weights i.e., α_i, w_i and b_i, it means the unique solution of the problem has been observed that not only satisfies the equation but also the initial conditions as well.

Adaptive learning procedure: The description of learning methodology adopted for finding the weights of the DE-NN architecture is provided here, which is based on annealing of metallurgy. The brief introduction of simulated annealing, flow diagram and step by step procedure for execution of the algorithm is also described.

Metropolis invents SA technique in 1950s (Khan *et al.*, 2011d), as it can get the global minimum in a low computational time unlike Golden search, steepest descent method, Conjugate-Gradient method, Qusai Newton method (Schmitt, 2002), Davidon-Fletcher-Powell methods (Roger and Powell, 1963) etc which are likely to get stuck in the local minimum. The real advantage of SA is its robustness, simplicity of concept, ease in implementation and is a generic probabilistic metaheuristic for the global optimization problems. The converging capabilities of SA are remarkable in pattern analysis, intelligent system designs and adaptive signal processing (Mantawy *et al.*, 1999; Gallego *et al.*, 1997). The initialization of the solution space is to be generated on the basis of scattered neighborhood points. The size of the population of neighborhood points is dependent on the nature of the problem to be optimized. The generic flow diagram of the algorithm used to get optimized response of the singular non-linear system is provided in Fig. 2.

The logical step of the machine learning procedure is provided below:

Step 1: Generation: A well scattered neighborhood sample size of 240 has been generated with randomly bound real numbers.

Step 2: Initialization: Initialize the values of the start point size, number of iterations, number of maximum function evaluation and other parameters such as Tol Fun, Tol Con etc.

Step 3: Fitness Evaluation: Calculate the fitness, by using the cost function given in Eq. (3).

Step 4: Termination Criteria: The algorithm is terminated on the basis of the following criteria:

- Predefines fitness value i.e., *MSE 10^{-10}* is achieved
- Predefines number of iterations executed
- Predefine functions evaluations met (Max FunEvals)
- Constraints tolerance (TolCon) limit crosses
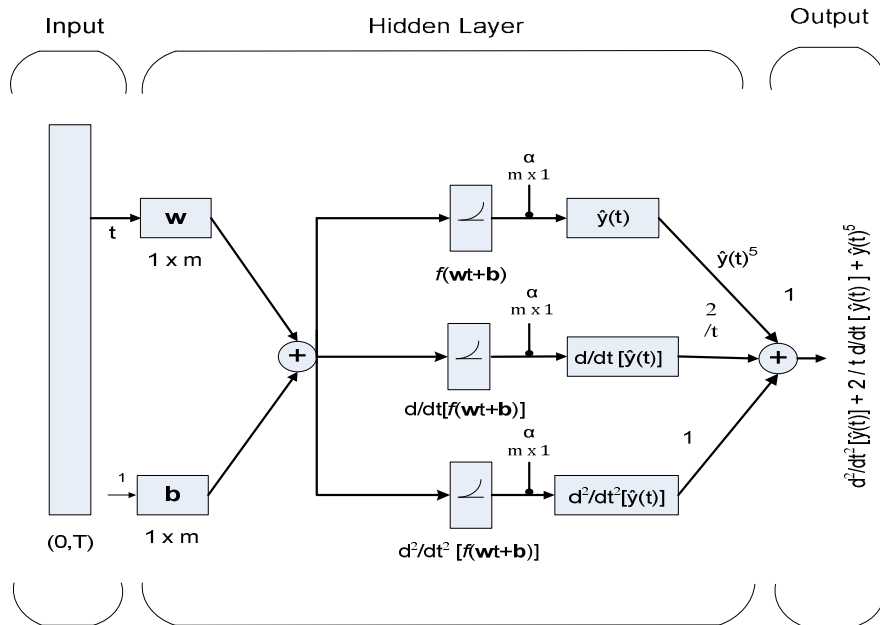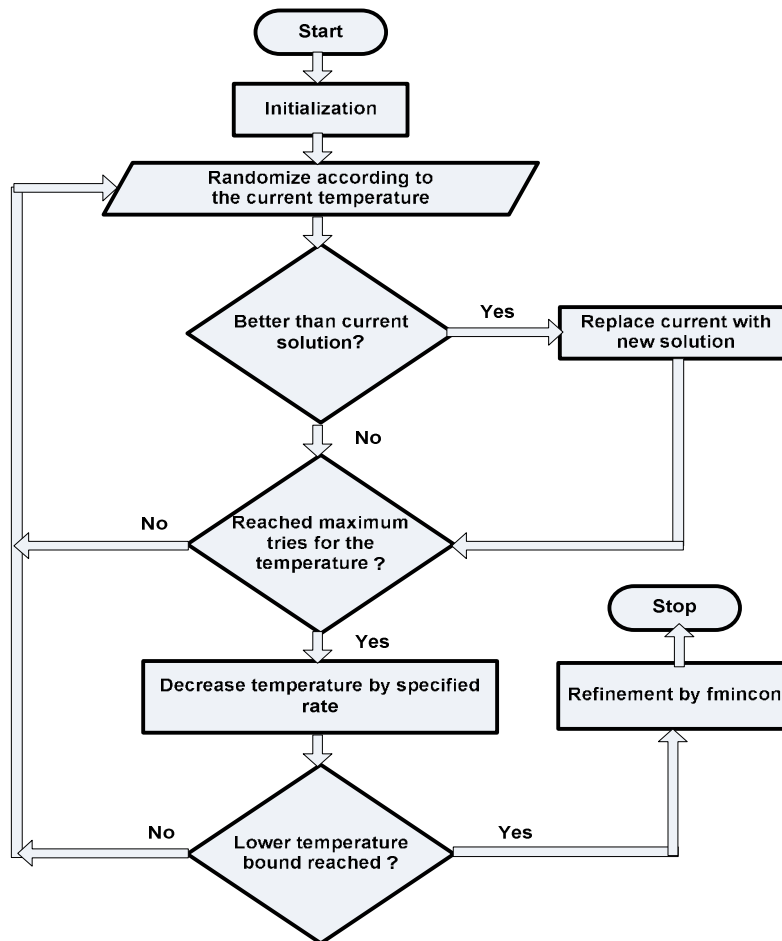- Function tolerance (TolFun) limit crosses

Fig. 1: DE-NN architecture

Fig. 2: Flow diagram of simulated annealing

Table 1: Parameters setting of the algorithm

Parameters	Value/setting	Parameters	Value/setting
Start point size	30	Re-annealing interval	100
Creation	Randomly between (-1, 1)	Temperature update	Exponential
Neighborhood sample size	240	Initial temperature	100
Annealing function	Fast	Bounds	(-10, 10)
Iteration	10000	Hybrid function	Call for FMINCON
Max funevals	90000	Hybrid function Call Interval	End
TolFun	10^{-12}	Time limit	Info
ToLCon	10^{-12}	Others	Defaults

If any of the above criteria met, then go to step (6)

Step 5: Renewal: The each point is ranked on the basis of the minimum of the fitness values. Update the current point with best neighboring point depending upon the calculation of temperature and energy and then go to step 3.

Step 6: Rapid Local Convergence: The global best point is given as a start point to FMINCON algorithm of the MATLAB optimization tool box for further fine tuning.

Step 7: Statistical Analysis: Store the global best point for the current run and then repeat step (3) to step (6) to have a sufficient number of global best points for a better statistical analysis.

RESULTS AND DISCUSSION

In this section, singular non-linear system provided in Eq. (1) has been simulated using the given scheme to see the response of the system and applicability of the proposed scheme. The desired response [32] of the model is represented by the relation $y(t) = (1 + t^2/3)^{-0.5}$ and is calculated in an entire finite domain between 0 and 1. This will work as a comparative analyzer with the approximated response. The desired response has been taken by a fiddly iterative method which is not worthy in the real time applications. Another problem is that by increasing the domain of the solution, the time and space complexity increases incredibly. Now the same model is approximated by the proposed scheme, the number of neurons in each hidden layer of the DE-NN network are taken to be $m = 10$ that results in 30 unknown adaptive weights (α_i, w_i and b_i). During the experimentation, it has been noticed that a better approximated response has been observed i.e the error of the desired and approximated response is really small, if we restrict the adaptive weights interval between [-10, 10]. The optimization of these weights is carried out using built-in function for SA in MATLAB and the generic and specific parameter setting/values used for the execution of the algorithm is listed in Table 1. The specific parameters are consist of annealing function, temperature updating technique, re-annealing interval and hybridization with any local search technique so that the fine tuning can be achieved to expedite the learning procedure for a better convergence.
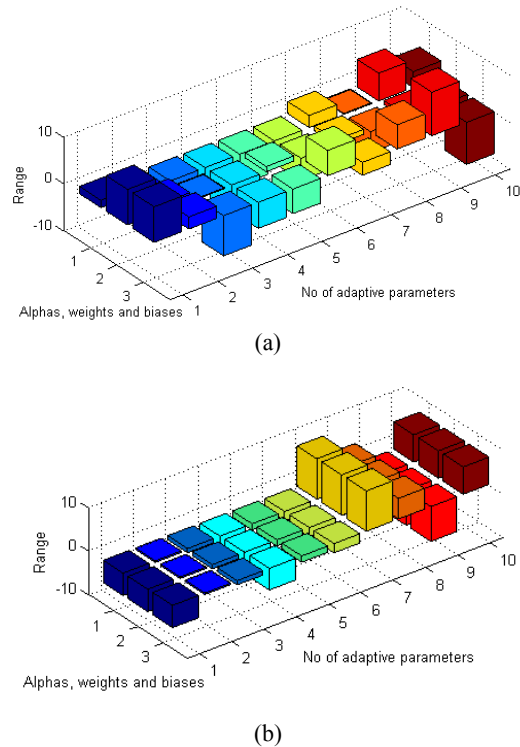

(a)

(b)

Fig. 3: Adaptive parameter obtained by the proposed scheme

A Monti Carlo simulation has been performed for the inputs of the training set from $t \in (0, 1)$ with a step size of 0.1 for 100 independents. Each run of the algorithm is performed for 10000 iterations so that enough data set should be achieved to see the reliability, applicability effectiveness of the given scheme. The one of the best and worst weights achieved in 100 independent runs are drawn in the form of 3D bars in Fig. 3 based upon the value of the fitness achieved. It is quite evident from the Fig. 3a and b that values of the weights in both cases are found to be in the restricted range of [10, 10] and the number of the neuron in the hidden layers are 30.

The weights of the Fig. 3a and b are used the approximate the model and the results are drawn on the log scale in order to have a clear look. It is quite evident from the Fig. 4a that the exact and approximated results are matched accurately on the other hand the results in Fig. 4b obtained via worst of the weights show the good convergence but a small fraction of the error from the
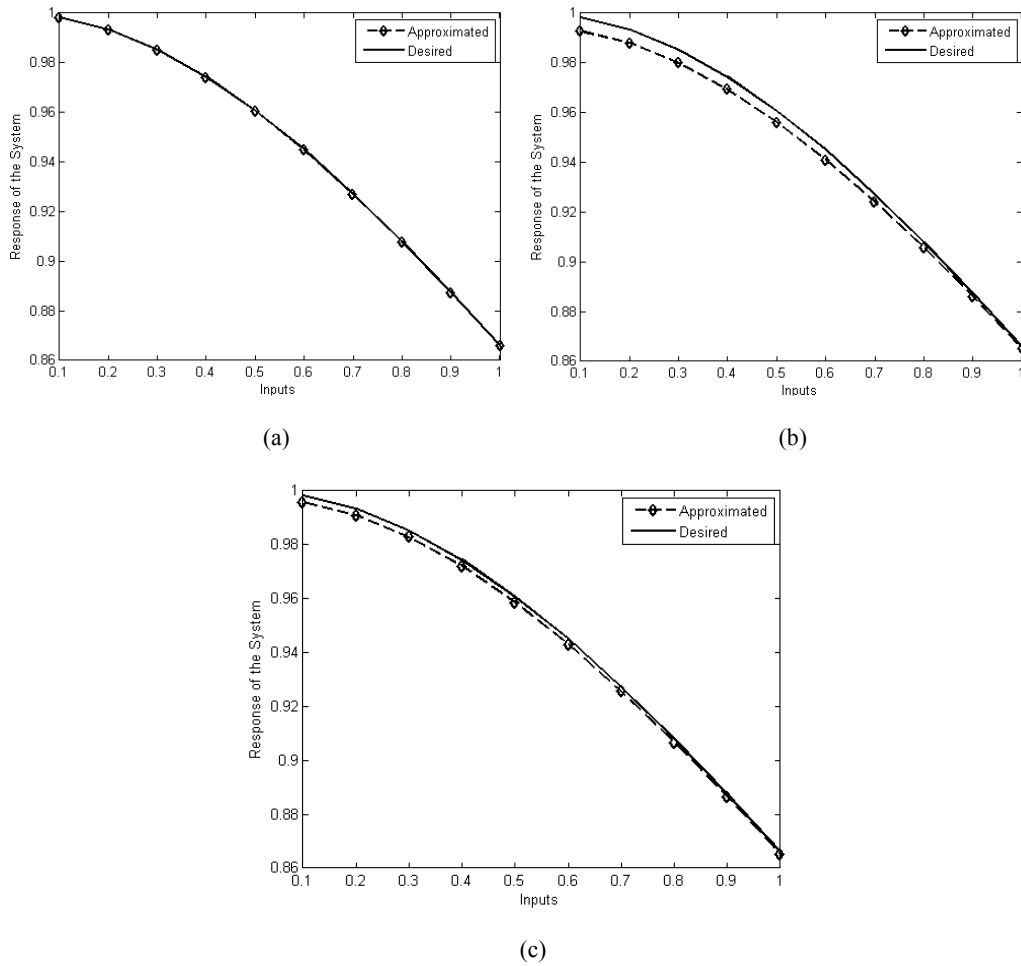
(a)

(b)

(c)

Fig. 4: The desired and approximated response of the singular non-linear system
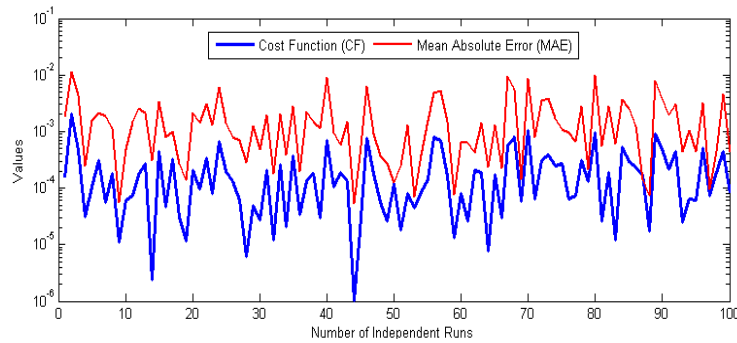

Fig. 5: Values of fitness and mean absolute error for 100 independent runs

desired response is observed. The absolute error from the desired is 10^{-05} to 10^{-06} and 10^{-03} to 10^{-04} for the best and worst weights, respectively. On the other hand the mean error of the approximated response is found to be in the range 10^{-05} to 10^{-06} and the results are plotted in Fig. 4c. There is always a trade of between the level of accuracy and computational complexity (space and time) of an algorithm. Therefore, even better results can

be achieved on the cost of machine learning and computational budget. The execution time is also calculated to see the real time efficiency of the proposed scheme and it is found that the average time taken for best and worst response is 31.1424 and 79.9004 seconds, respectively which is indeed a good projection of the proposed scheme.

Moreover, the reliability of the proposed stochastic algorithm is being represented by providing the fitness achieved in each independent run and is plotted in Fig. 5. It is quite clear from the Fig. 5 that fitness of the cost function lie from *1e-03* to *1e-06* and mean absolute error from the desired response lie in the range *1e-03* to *1e-06*. The accuracy of the scheme is based on statistical analysis of 100 independent runs and it is found that the mean and standard deviation in the values of the fitness are found to be $2.3161e^{-04}$ and $2.9848e^{-04}$, respectively. Similarly the mean and standard deviation in the execution time of 100 independent runs is to be 45.2562 and 16.4926 seconds, respectively.

It can be seen that no divergence is observed, so 100% convergence is achieved for the entire input domain. Finally, it can be concluded that the proposed scheme provides consistent accuracy and convergence. The whole analysis carried out for this study is based on LATITUDE D630, with Intel(R) Core (TM) 2 CPU T7600@2.33GHz processor, 2.00 GB RAM and running with MATLAB version 2011b.

CONCLUSION

On the basis of simulation and results provided in the last section it can be concluded that:

The stochastic solver based on DE-NN optimized with hybridized SA can effectively provides the solution of the Non-Linear Singular System with mean of the absolute error lies in range of 10^{-05}.

The reliability and effectiveness of given scheme is validated from statistical analysis and is found that the confidence interval for the convergence of the given approach is 100% to get an approximate solution of the model in a acceptable error range.

The proposed scheme can readily provide the solution on the continuous grid of the inputs unlike classical methods. Thus this provides an alternate approach to researchers to apply the machine learning procedures to complex real life problems.

The real advantage of the proposed scheme is its robustness, simplicity of concept, ease in implementation and is its generic probabilistic metaheuristic global optimization capabilities.

In future, one can look for applications of other artificial intelligence techniques optimized with ant/bee colony optimization, genetic programming and differential evolution etc. for solving such a vast applications.

ACKNOWLEDGMENT

REFERENCES

Aarts, L.P. and P.V.D. Veer, 2001. Neural network method for solving the partial differential equations. Neural Process. Lett., 14: 261-271.

Chowdhury, M.S.H. and I. Hashim, 2007. Solutions of time-dependent Emden fowler type equations by homotopy-perturbation method. Phys. Lett. A., 368: 305-313.

Costin O. and R.D. Costin, 2001. On the formation of singularities of solutions of nonlinear differential systems in antistokes directions. Invent. Math., 145(3): 425-485.

Dehgham, M. and F. Shakeri, 2008. Solution of an integro-differential equation arising in oscillating magnetic field using He's Homtopy perturbation method. Progr. Electromag. Res., 78: 361-376.

Fogel, D.B., 2006. Evolutionary Computation: Toward A New Philosophy of Machine Intelligence. 3rd Edn., John Wiley and Sons, Inc., Hoboken, New Jersey.

Fogel, D.B. and A. Ghozi, 1997. Schema processing under proportional selection in the presence of random effects. IEEE T. Evol. Comput., 1(4): 290-293.

Gallego, R., A. Alves, A. Monticelli and R. Romero, 1997. Parallel simulated annealing applied to long term transmission network expansion planning. IEEE T. Power Syst., 12(1): 181-188.

Herisanu, N., V. Marinca, T. Dordea and G. Madescu, 2008. A new analytical approach to nonlinear vibration of an electrical machine. Proc. Romanian Acad. Series A, 9(3): 229-236.

Khan, J.A., M.A.Z. Raja and I.M. Qureshi, 2011a. Stochastic computational approach for complex non-linear ordinary differential equations. Chin. Phys. Lett., 28(2): 020206-020209.

Khan, J.A., M.A.Z. Raja and I.M. Qureshi, 2011b. Numerical treatment of nonlinear Emden-fowler equation using stochastic technique. Ann. Math. Artif. Intell., 63(2): 185-207.

Khan, J.A., I.M. Qureshi and M.A.Z. Raja, 2011c. Hybrid evolutionary computational approach: Application to van der Pol oscillator. Int. J. Phys. Sci., 6(31): 7247-7261.

Khan, J.A., M.A.Z. Raja and I.M. Qureshi, 2011d. Novel approach for van der Pol oscillator on the continuous time domain. Chin. Phys. Lett., 28(1): 102-105.

Koumboulis, F.N. and B.G. Mertzios, 2005. Stability criteria for singular systems. J. Neural Parallel Sci. Comput., 13(2): 199-212.

Liao, S.J., 1992. On the proposed homotopy analysis teachnique for nonlinear problems and its applications. Ph.D. Thesis, Shanghai Jio Tong University, Shanghai, China.

Mantawy, A.H., Y.L. Abdel-Magid and S.Z. Selim, 1999. Integrating genetic algorithms, tabu search and simulated annealing for the unit commitment problem. IEEE T. Power Syst., 14(3): 829-836.

Marinca, V., N. Herisanu and I. Nemes, 2008. Optimal homotopy asymptotic method with application to thin film flow. Cen. Eur. J. Phys., 6(3): 648-653.

Mertzios, B.G., 2000. Direct solution of continuous time singular systems based on the fundamental matrix. Proceeding of the 8th IEEE Mediterrenean Conference on Control and Automation (Med 2000). Rio, Patras, Greece, pp: 17-19.

Monterola, C. and C. Saloma, 1998. Characterizing the dynamics of constrained physical systems with unsupervised neural network. Phys. Rev. E., 57: 1247-1250.

Parand, K. and A. Taghavi, 2008. Generalized laguerre polynomials collocation method for solving Lane-Emden equation. Appl. Math. Sci., 2(60): 2955-2961.

Raja, M.A.Z., J.A. Khan and I.M. Qureshi, 2010. A new stochastic approach for solution of Riccati differential equation of fractional order. Ann. Math. Artif. Intell., 60(3-4): 229-250.

Raja, M.A.Z., I.M. Qureshi and J.A. Khan, 2011a. Swarm Intelligent optimized neural networks for solving fractional differential equations. Int. J. Innov. Comput., Inform. Control, 7(11): 6301-6318.

Raja, M.A.Z., J.A. Khan and I.M. Qureshi, 2011b. Solution of fractional order system of Bagley-Torvik equation using evolutionary computational intelligence. Math. Probl. Eng., pp: 1-18, Doi: 10.1155/2011/675075.

Rarisi, D.R., M.C. Mariani and M.A. Laborde, 2003. Solving differential equations with unsupervised neural networks. J. Chem. Eng. Process., 42: 715-721.

Richardson, O.W., 1921. The Emission of Electricity from Hot Bodies. London.

Roger, F. and M.J.D. Powell, 1963. A rapidly convergent descent method for minimization. Comput. J., 6: 163-168.

Saloma, C., 1993. Computational complexity and observation of physical signals. J. Appl. Phys., 74: 5314-5319.

Schmitt, M., 2002. On the complexity of computing and learning with multiplicative neural networks. Neural Comput., 14(2): 241-301.

Shawagfeh, N.T., 1993. Nonperturbative approximate solution for Lane-Emden equation. J. Math. Phys., 34: 4364-4369.

Tsoulos, L.G., D. Gavrilis and E. Glavas, 2009. Solving differential equations with constructed neural networks. Neurocomputing, 72(10-12): 2385-2391.