Artificial Neural Network Analysis of the Solar-Assisted Heat Pumps Performance for Domestic Hot Water Production

Alireza Zendehboudi, Xianting Li*, Siyuan Ran

Department of Building Science, School of Architecture, Tsinghua University, Beijing 100084, China
xtingli@tsinghua.edu.cn

Abstract. The solar-assisted heat pump devices have attracted ample attention as potential alternatives to the conventional technologies for domestic hot water production. Precise COP estimation of the hybrid system is a prerequisite prior to conducting any effort to design and install as well as a necessity for market support and optimal system assurance. Here, to develop a model along that has a high precision and reliability, but less complexity and computational time, the ability of soft computing approaches for predicting the performance of a solar-assisted heat pump system for domestic hot water production is reported. The experimental data gathered from a real project in China and different intelligent models are developed based on the measurements. First, owing to the complexity of our studied system and the high number of influential parameters, a novel and unique multi-objective optimization technique based on NSGA-II optimization method is proposed to determine a set of optimum variables with the highest influence on the desired output. Based on the Pareto frontier, seven input variables out of forty nine are considered for the desired output. The reliability of the developed models is evaluated via statistical and graphical error analyses. It was inferred that integration of the MLP-ANN with the suggested variable selectin algorithm outperformed the other methods by introducing an R^2=0.9951 and RMSE=0.0917. The current investigation can aid as a gear in the direction of improving the precision in estimation of performance of solar-assisted heat pump devices.

Keywords: Domestic hot water system, solar-assisted heat pump, artificial neural network, estimation.

1 Introduction

Nowadays, environmental pollution and energy shortage are two main challenges in the world. Therefore, saving energy in heating, ventilation, and air-conditioning devices plays a momentous role in the worldwide energy usage. In China, the domestic hot water production is responsible for 23.4% of the total residential building energy consumption in urban areas [1]. Currently, boilers and electric water heaters are

commonly utilized in residential and commercial buildings to provide the domestic hot water. One of the main concerns of prevalent utilization of these systems is their low primary energy efficiency as well as contributing to air pollution, climate change, and global warming. Consequently, to avoid the fossil fuels waste and the nettlesome problems that they bring about, renewable energy sources like solar energy have attracted much attention from industries and researchers in recent decades as a substantial resource. Solar energy is able to offer a novel alternative solution due to its clean nature and abundance in the world.

Recently, new-generation hot water devices such as the solar water heater have attracted ample attention as potential alternatives to the conventional technologies. However, the relatively low efficiency makes some limitations to exploit the solar water heaters via various technologies, especially for locations with low solar potential. In this context, the solar-assisted heat pumps were developed and presented based on the integration of solar water heaters and air source heat pumps, which are using the benefits of each single device. In an indirectly connected system, water or a solution is utilized to transfer heat from the solar collector to the evaporator of the heat pump. In recent years, there have been several researches on the experimental and analytical studies of solar-assisted heat pumps [2,3,4].

Owing to variability of solar radiation, high number of influential parameters, and the considerable complexity that is involved in the solar-assisted heat pumps, the reliable information on the hybrid systems performance is a prerequisite prior to conducting any effort to design, install, and evaluate the aforementioned systems. Generally, the superior way of knowing the performance of the systems under different operating conditions is to establish a highly equipped laboratory and conduct sufficient and nettlesome experiments. Nevertheless, due to a series of restrictions like requiring considerable investments and proper instruments the experimental approaches may not assess performance of these types of complex systems. Therefore, it is of great importance to develop the numerical approaches and computer technologies to tackle the limitations and alternatively evaluate performance of such systems. Soft computing techniques are an alternative, which have the high ability to model complex relationships between input-output pairs. The major merits of artificial neural algorithms are their capability to map nonlinear functions, to forecast with high generalization capability, to deal with multivariable problems, and to process large amount of data, which make them practical for modeling applications [5].

The primary objective of this investigation is to present a reliable, accurate, and practical model to simulate and monitor the performance of the solar-assisted heat pumps. This provides us the possibility to select the correct value of each operating variable to maximize the hybrid systems performance. To fulfill this objective properly, experimental data samples from a real project in China are collected and a Multilayer Perceptron-Artificial Neural Network (MLP-ANN) is developed. Firstly, a novel method for determining the most dominated input variables is introduced to elucidate the influence of each input variable on the output. Ultimately, a comparative study is carried out to appraise the reliability of the proposed MLP-ANN model against the other well-known intelligent models, such as Hybrid-Adaptive Neuro Fuzzy Inference System (Hybrid-ANFIS), Genetic Algorithm-Least Square Support Vector Machine

(GA-LSSVM), and Group Method of Data Handling (GMDH) methods. The results obtained are presented in further detail in following sections.

2 Methods

2.1 Experimental set-up

The principle of the studied solar-assisted heat pump is shown in Fig. 1. The system is a practical project of domestic hot water for a hotel, which consists of three water tanks, i.e., solar water tank, heat storage water tank, and supply water tank. The cold water is pumped into the solar water tank and circulated between solar collector and solar water tank, which is installed at the roof of the hotel. After heated by solar energy, the water is then discharged to the heat storage water tank and is heated by air source heat pumps (ASHP1 and ASHP2). The hot water from heat storage water tank enters the supply water tank and is supplied to the hotel. Another air source heat pump (ASHP3) is set to ensure the temperature of the supply water. The heat storage water tank and supply water tank are located in a cabin beside the hotel, while the air source heat pumps are next to the cabin.

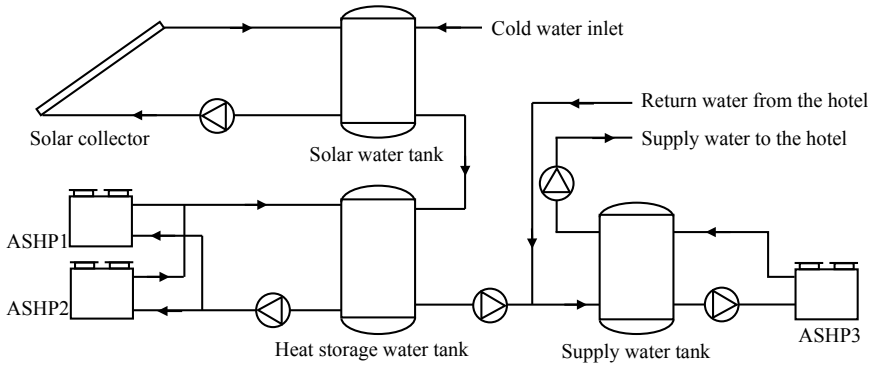


Fig. 1. Schematic diagram of the system.

The volume of the solar water tank, heat storage water tank, and supply water tank is 1.5m^3, 12m^3, and 3m^3, respectively; but the tanks are not always full of water. The solar collector is heat pipe evacuated tube collector with total area of 67.2 m^2. The intercept efficiency and efficiency slope of the collector is 0.760 and 2.284, respectively. The heating capacity of the three air source heat pumps is 28kW when ambient temperature is 15°C and electric power consumption is 7 kW. The maximum temperature of solar water tank is 50°C. When the temperature is above 50°C, the water is discharged to the heat storage water tank. The temperature range of heat storage water tank is 43-48°C, which is a bit lower than maximum temperature of solar water tank

Table 1. Recorded input parameters of the system.

Input parameters	
Date	Outlet water Temperature of ASHP2
Temperature of solar collector	Ambient temperature of ASHP2
Temperature of the bottom of solar water tank	On-off ratio of water inlet valve of the system
Temperature of the top of solar water tank	On-off ratio of defrosting mode of ASHP2
Temperature of the heat storage water tank	On-off ratio of compressor of ASHP2
Temperature of supply water tank	On-off ratio of fan 1 of ASHP2
Temperature of the water from the user	On-off ratio of fan 2 of ASHP2
On-off ratio of the pump for solar collector	Inlet water temperature of ASHP3
On-off ratio of the pump between heat storage water tank and supply water tank	On-off ratio of pump between heat pump and heat storage water tank of ASHP2
Outlet air temperature of the evaporator of ASHP2	Temperature of the supply water tank to the user
On-off ratio of the pump between solar water tank and heat storage water tank	Outlet air temperature of the evaporator of ASHP3
Outlet water temperature of ASHP3	Ambient temperature of ASHP3
Temperature of the electric heater	On-off ratio of defrosting mode of ASHP3
Inlet water Temperature of ASHP1	Inlet volume of supply water tank
Outlet water Temperature of ASHP1	On-off ratio of compressor of ASHP3
Outlet air Temperature of the evaporator of ASHP1	On-off ratio of pump between heat pump and heat storage water tank of ASHP3
Ambient temperature of ASHP3	On-off ratio of fan 1 of ASHP3
Temperature of the bottom of heat supply water tank	On-off ratio of pump between heat pump and heat storage water tank of ASHP1
On-off ratio of defrosting mode of ASHP1	Inlet water volume
On-off ratio of compressor of ASHP1	Volume of solar water tank
On-off ratio of fan 1 of ASHP1	Volume of heat storage water tank
On-off ratio of fan 2 of ASHP1	Volume of supply water tank
On-off ratio of fan 2 of ASHP3	Total heating capacity of the system
Inlet water Temperature of ASHP2	Inlet volume of heat storage water tank

because of the heat loss of the tank and pipes. The ASHP turns on if the temperature is below 43°C. The temperature of the supply water tank is 45-50°C to meet the demand of the hotel.

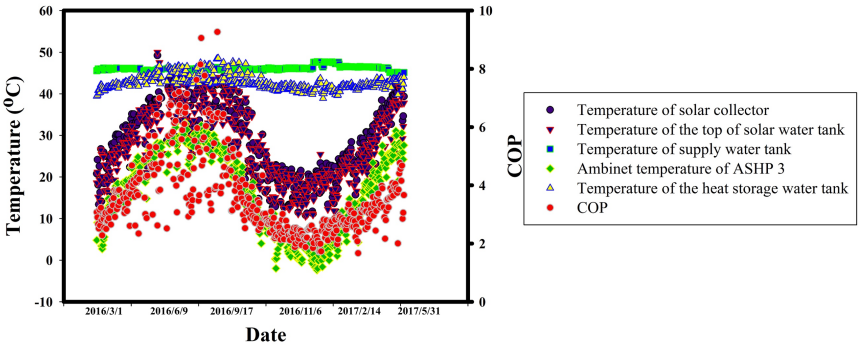


Fig. 2. The input and output variations.

The data set for the needs of this investigation contains 425 experimental samples derived from a real project in China, which comprises forty nine input variables, as listed in Table 1. These 49 inputs were monitored from 2016/3/1 to 2017/5/31. The output parameter is COP of the system, which defined as the total heating capacity divided by the total electric consumption. The ambient temperature, the temperature of the solar collector, the temperature of water tanks, and COP are shown in Fig. 2. The temperature of solar collector, temperature of the top of solar water tank, ambient temperature of ASHP3, and COP varies with date, while the temperatures of heat storage water tank and supply water tank are nearly constant.

2.2 Modelling

Input variable selection. Input variable selection refers to identifying the optimal or a set of optimum variables with the highest influence on the desired output. Owing to the complexity of our system, the high number of influential parameters on the system, and the complex inter-relations among them, proper selection of the most significant input variables for predicting the COP of the system is desperately needed to develop a model along that has a high precision, robustness, and reliability, but less complexity and computational time. It has been proven that more or irrelevant input variables create many drawbacks in the model. Some of the drawbacks are: (1) difficulty in explaining the model; (2) inaccuracies caused by irrelevant parameters; (3) complexity in the developed model due to a high number of required inputs; and (4) the requirement of the time-consuming task of collecting more data. These factors may consequently cause the generalization capacity of the model to deteriorate.

In this research work, a multi-objective optimization method based on non-dominated sorting genetic algorithm, NSGA-II, was developed and utilized to determine the optimum number of input parameters for the desired output. More details about genetic algorithms and their procedures are given elsewhere [6]. The mean square error (MSE), which is the calculated error between the predicted values by the developed predictive models and the experimental values considering different sets of input variables, and the number of input variables are considered as the cost functions to be minimized using the evolutionary algorithm. The MSE is described as:

$$MSE = \frac{1}{2}\sum_{i\in N}(t_i - p_i)^2 \tag{1}$$

where N represents the training data set, t_i and p_i are the target and predicted vales for the *ith* training data. In the approach, the final non-dominated solutions are considered as the final responses.

Multilayer perceptron-artificial neural network model. Artificial neural networks (ANNs) represent parallel information processing methods, which are capable of explaining nonlinear and complex logic operations and determining the pattern within data via input-output training pairs from a data set. The adaptive nature of ANNs and their high ability in the learning of system's behavior from the information

that introduced to them during training phase make them popular in different research areas to tackle with different large-scale ill-defined problems [7]. The feed-forward neural network is the most common used network architecture. In general, three layers are applied to construct the MLP-ANN model, with every layer performing a specific action. This approach includes simple processing elements called neurons, which are connected to one another in a parallel structure by weights.

The number of neurons in the input and output layers is defined based on the problem of interest, while the optimum numbers and sizes of hidden layers are determined by adding neurons in a systematic form during training process. These numbers are flexible and directly depend on the complexity of problem, the number of training data samples, and the noise in the representative data [8].

Back-Propagation (BP), which is a supervised training algorithm, is widely applied for training MLP-ANNs to update the weights and biases [9]. Suppose an input vector $X=[\chi_1, \chi_2, \chi_3, \ldots, \chi_n]$ and weight vector $W=[w_1, w_2, w_3, \ldots, w_n]$, the signals are fed to the input layer and the input weighted sum $\sum_{i=1}^{n} w_i x_i$ is calculated. This produced sum is transferred to a predefined transfer function, as shown in Equation (2).

$$y(x_1, x_2, x_3, \ldots, x_n) = f(\sum_{i=1}^{n} w_i x_i + b) \tag{2}$$

where *f (*)* indicates the transfer function. At the next step, following the rule minimizing the error between the experimental COP values and the values predicted by the approach, an objective function such as MSE is defined. The calculated error is considered as a criterion for tuning of the network's weights and biases.

To assess and compare the performance of the aforementioned technique, two well-known statistical error metrics like the coefficient of determination (R^2) and root mean square error (RMSE) were employed. These parameters are depicted in Equations (3) and (4):

$$R^2 = 1 - (\sum_{i=1}^{n}(t_i - p_i)^2 / \sum_{i=1}^{n}(t_i - \bar{t}_i)^2) \tag{3}$$

$$RMSE = \left(\frac{1}{n}\sum_{i=1}^{n}(t_i - p_i)^2\right)^{0.5} \tag{4}$$

where n, t_i, p_i, and $\bar{t}_i$ are the total number of data, the *ith* target data value, the *ith* data value predicted by the model, and the average of the target data values, respectively.

3 Results and discussion

To estimate the COP of the hybrid system, the novel variable selection algorithm was applied and run considering the forty nine recorded variables for selecting the most influential input variables on the output parameter. To highlight the merits and effectiveness of the proposed novel method, an additional predictive model was constructed and developed based on forty nine input variables and the results were compared to appraise the integrity and the potential of the suggested technique. Next, considering the determined input variables, the MLP-ANN predictive method was trained and

tested. Ultimately, to achieve the main goal of this research through the current study, a comparative study between the MLP-ANN and the other benchmark models was carried out to ascertain the generality, applicability, and comprehensiveness of the developed MLP-ANN in estimating precise values of the output parameter.

3.1 Implementing the variable selection technique

Fig. 3 represents the Pareto frontier solution for the problem of interest associated with the MSE and number of input variables objective functions for the multi-objective optimization.

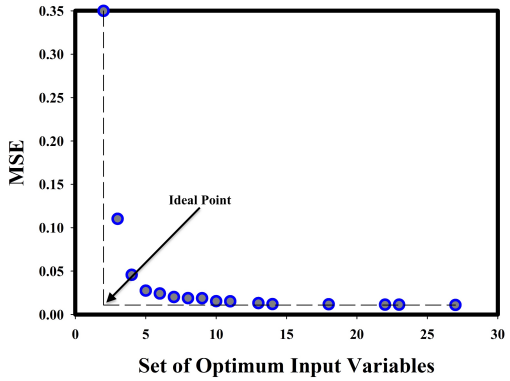


Fig. 3. Pareto frontier: best trade off values for the objective functions.

Obviously observed that at the start the MSE value decreases significantly as the number of input variable increases. It should be mentioned that in multi-objective optimization and the generated Pareto solution, each point has the possibility to be considered as an optimized point. In this context, the best solution for each problem is determined based upon the criteria of the decision maker and might be considered a different point as for the optimum solution according to the desires. Hence, the 0.0201 value for the MSE occurs when seven input variables, i.e., (1) temperature of solar collector, (2) temperature of the electric heater, (3) on-off ratio of fan 1 of ASHP1, (4) on-off ratio of compressor of ASHP2, (5) on-off ratio of fan 2 of ASHP3, (6) on-off ratio of pump between heat pump and heat storage water tank of ASHP3, (7) total heating capacity of the system, were considered for the desired output, while the value decreases slightly as the number of input variable increases. Therefore, this point is considered as the optimal situation by the purpose of constructing and developing a high performance and reliable predictive model.

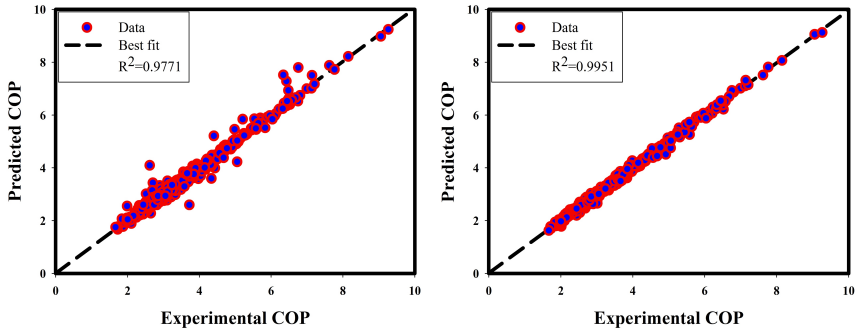
3.2 MLP-ANN output results

The best network topology for forecasting the COP was determined by designing different architectures using the BP algorithm and changing the number of hidden

layers, the number of neurons in the layers, and the transfer functions in the hidden and output layers. Table 2 summarizes the functions that used to construct the optimum MLP-ANN scheme.

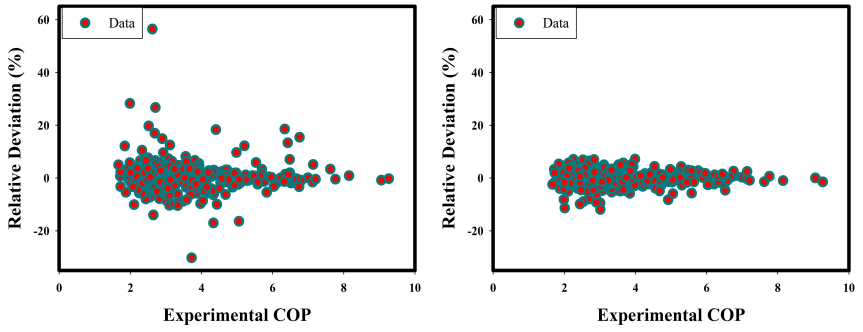
Table 2. User defined functions implemented in the MLP-ANN predictive model by considering seven input variables.

Number of layers	Number of neurons			Training function	Transfer function	
	Input	Hidden	Output		Hidden	Output
3	7	7	1	TrainBR	Tansig	purelin



(a) (b)

Fig. 4. Estimated COP vs. experimental by considering different input variables: (a) 49; (b) 7.



(a) (b)

Fig. 5. Relative deviation vs. experimental by considering different input variables: (a) 49; (b) 7.

The regression plot presented in Fig. 4 illustrates the correlation between experimental data samples and estimated by the developed MLP-ANN considering different input vectors. The calculated R^2 values using the MLP-ANN for before and after implementing the variable selection method are 0.9771 and 0.9951 respectively, which

ascertain the agreement between actual and estimated values. As is evident, the values of R^2 are quite close to unity, while the computed value after using the variable selection technique and determining the most effective input variables is higher, which shows a favorable and adequate performance of the MLP-ANN model when is combined with the novel algorithm. To support this claim, Fig. 5 presents relative error deviation percentage plot of estimations of both developed models as a function of actual COP values. This figure indicates that the model with seven input variables exhibits lower relative errors compared to another model with forty nine input variables as the relative errors of most data points predicted by the seven input variables model fall in the region bounded by two lines with relative deviations of -11.9% and +7.14%, while another developed model is capable of predicting the COP within -30.2% and +56.4% error band based on the given data. This fact indicates that the model is able to reproduce the whole set of experimental data with the lowest possible error. The computed statistical error indicators are presented graphically in Fig. 6. Once again, it is clear from Fig. 6 that the MLP-ANN model combined with the novel variable selectin technique has the best performance on estimating the COP. It can be observed from Fig. 6 that the selected input variables for developing the MLP-ANN model resulted in a higher R^2 and lower RMSE values.

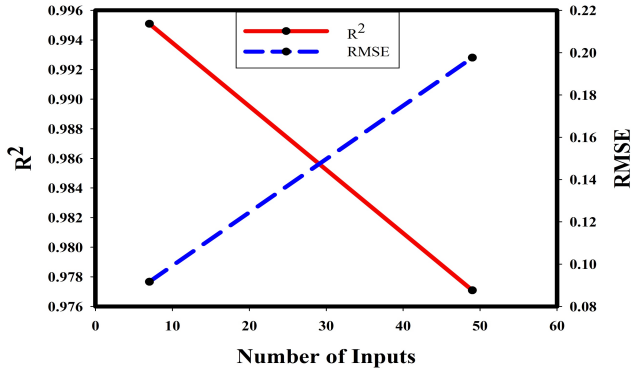


Fig. 6. The variation of statistical criteria for different models by considering different input variables.

3.3 A comparison with the other benchmark models

The well-known statistical error indicators, R^2 and RMSE, are calculated and summarized in Table 3 when the GA-LSSVM, Hybrid-ANFIS, and GMDH techniques are implemented to evaluate the predictive ability of the proposed MLP-ANN model from different aspects. Based on the calculated results reported in Table 3, the MLP-ANN provides the most accurate results and falls in the category of an excellent fit by presenting the highest precision and integrity for estimation of the COP based on the given data, followed by the GA-LSSVM and Hybrid-ANFIS approaches. However, the GMDH technique was not among the best models and it provided lower performance.

Table 3. Accuracy comparison results by considering seven input variables.

Parameter	Statistical criteria	Network			
		MLP-ANN	GA-LSSVM	Hybrid-ANFIS	GMDH
COP	R^2	0.9951	0.9929	0.9864	0.9357
	RMSE	0.0917	0.1114	0.1533	0.3335

4 Conclusion

The present contribution computed the performance of a solar-assisted heat pump device, which is based on integration of solar water heaters and air source heat pumps to produce domestic hot water. To determine a subset of the most suitable input variables, a novel method based on NSGA-II optimization method was proposed and implemented on the data set. As the first study of its type, the new technique paves the way for selecting the most predominant input variables among the factors that influence prediction of the desired output. It was indicated that one of the merits of this unique algorithm was decreasing the computational time and costs. Compared to the model including all of forty nine input variables, another positive outcome of this approach was increasing the coefficient of determination from 0.9771 to 0.9951 and decreasing the root mean square error from 0.1977 to 0.0917 when it combined with a MLP-ANN. A comparison of the MLP-ANN technique with the GA-LSSVM, Hybrid-ANFIS, and GMDH based on the same data set and input variables was carried out to assess the estimation robustness and precision. According to the results, the developed MLP-ANN method outperformed the others. Finally, it is safe to assert that the proposed hybrid model is suitable and practical for other complex systems due to its advantages.

Acknowledgments

This study was supported by the Innovative Research Groups of the National Natural Science Foundation of China (Grant No. 51521005).

References

1. Tsinghua University Building Energy Saving Research Center: Annual Report on China Building Energy Efficiency. China Architecture and Building Press, Beijing (2011) (in Chinese).
2. Panaras, G., Mathioulakis, E., Belessiotis, V.: Investigation of the performance of a combined solar thermal heat pump hot water system. Solar Energy 93, 169-182 (2013).
3. Poppi, S., Bales, C., Haller, M.Y., Heinz, A.: Influence of boundary conditions and component size on electricity demand in solar thermal and heat pump combisystems. Applied Energy 162, 1062-1073 (2016).

4. Poppi, S., Bales, C., Heinz, A., Hengel, F., Chèze, D., Mojic, I., Cialani, C.: Analysis of system improvements in solar thermal and air source heat pump combisystems. Applied Energy 173, 606-623 (2016).
5. Svozil, D., Kvasnicka, V., Pospichal, J.: Introduction to multi-layer feed-forward neural networks. Chemometrics and Intelligent Laboratory Systems 39, 43-62 (1997).
6. Deb, k., Pratap, A., Agarwal, S., Meyarivan, T.: A fast and elitist multiobjective genetic algorithm: NSGA-II. IEEE Transactions on Evolutionary Computation 6, 182-197 (2002).
7. Haykin, S.: Neural Networks, A comprehensive foundation. Prentice Hall Inc, New Jersey (1999).
8. Du, K.-L., Swamy, M.N.S.: Neural networks in a soft computing framework. Springer, London (2006).
9. Goh, A.T.C.: Back-propagation neural networks for modeling complex systems. Artificial Intelligence in Engineering 9, 143-151 (1995).