Artificial Neural Networks and Genetic Algorithms for the Optimisation of Solar Thermal Systems

Soteris A. Kalogirou
Higher Technical Institute, P.O.Box 20423, Nicosia 2152, Cyprus, Tel.: +357-22-406466,
E-mail: skalogir@spidernet.com.cy

ABSTRACT

This paper presents a new method to optimise solar energy systems in order to maximise their economic benefits. The system is modelled with TRNSYS computer program. An artificial neural network is trained using a small number of annual TRNSYS simulation results, to learn the correlation of collector area and storage tank size on the auxiliary energy required by the system and thus on the net solar energy price. Subsequently a genetic algorithm is employed to estimate the optimum size of these two parameters, which maximise the net solar energy price, thus the design time is reduced substantially and the solution obtained is more accurate that the trial and error method used traditionally in these optimisations.

Keywords: Artificial neural networks, genetic algorithms, solar thermal systems

1. INTRODUCTION

In any solar energy installation, it is necessary to optimise the collector area in order to maximise the economic benefits of the system. In this work, economic benefits are considered as the lowest solar energy price, i.e., solar energy produced at the lower possible cost. The objective of this work is to present a method for the optimisation of solar energy systems based on two artificial intelligence techniques; namely, artificial neural networks and genetic algorithms. For this purpose, an artificial neural network is trained to learn the correlation of collector area and storage tank on the net solar energy price and a genetic algorithm is employed to estimate the optimum size of these two parameters by minimizing the net solar energy price of the installation. Thus, the design time could be reduced substantially and the optimum solution eventually reached is more correct than the solution reached by trial and error. The importance of this selection method is to make solar applications more viable and thus more attractive to potential users.

2. CHARACTERISTICS OF INDUSTRIAL PROCESS HEAT SYSTEM

As an example, an industrial process heat (IPH) system shown schematically in Fig. 1 is considered. The system consists of an array of collectors, a circulating pump and a storage tank. It includes also the necessary controls and thermal relief valve, which relieves energy when storage tank temperature is above a preset value. The system is once through, i.e., there is no hot water return to storage, thus used hot water is replaced by mains water. When the temperature of the stored water is above the required process temperature, this is mixed with mains water to obtain the required temperature. If no water of adequate temperature is available in the storage tank its temperature is topped-up with an auxiliary heater before use. For the modelling and simulation of the system, the well-known program TRNSYS is employed [1].

The system considered is one where 2000 kg/hr of hot water are used at a temperature of 85°C (load). The load pattern and other system characteristics are shown in Table 1. Flat plate collectors are employed for this system, which are by far the most used type of solar collector. Flat plate collectors are usually permanently fixed in position and require no tracking of the sun. The collectors should be oriented directly towards the equator, facing south in the northern hemisphere and north in the southern. The characteristics of the collectors considered are shown in Table 2.

Table 1 Characteristics of the basic system

Parameter	Value/Type
Load temperature	85°C
Load flow rate	2000 l/hr
Use pattern	5 days a week, 8.00-16.00 hours each day, load used for the first ¾ of each hour
Collector to storage distance	30m
Piping UA value	20 W/°C
Piping diameter	75mm
Relief valve set temperature	100°C

Table 2 Characteristics of the collector system

Parameter	Value
Type of collector	Flat-plate
Fixing of risers on absorber plate	Embedded
Absorber coating	Black mat paint
Glassing	Low-iron glass
Efficiency mode	n v_s (Ti - Ta)/I_T
Flow rate per unit area at test conditions	0.015 kg/s-m^2
(I) Intercept efficiency	0.792
(S) Negative of first-order coefficient of the efficiency	6.67 W/m$^{2\,°}$C
(b_o) Incidence angle modifier constant	0.1

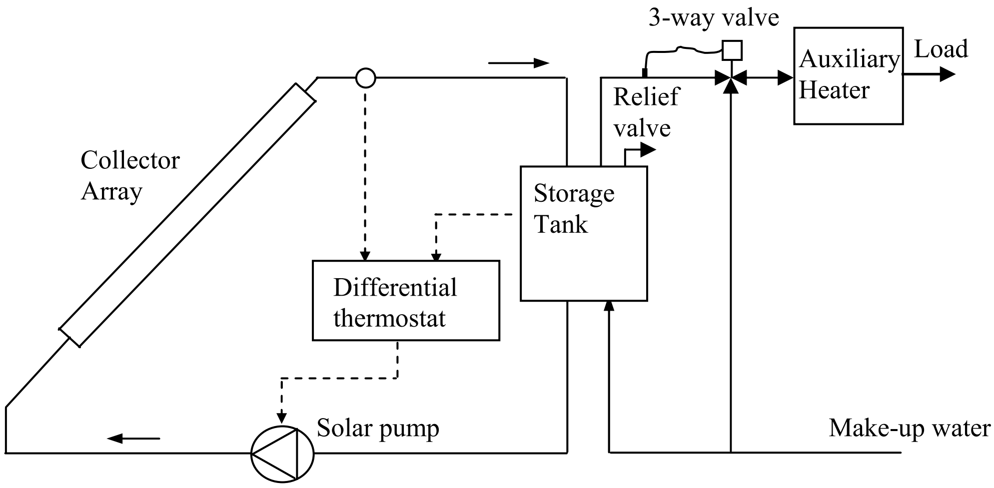

Fig. 1 Schematic diagram of the solar collector system

3. SYSTEM MODEL

For the modelling of the system the well-known program, TRNSYS is employed. TRNSYS is employing the standard collector performance equation in which the intercept (I) and slope (S) factors, shown in Eq. 1, are used to model the collector.

$$n = K_{\alpha\tau} I - S\frac{\Delta T}{G} \tag{1}$$

where G is the global solar radiation, $k_{\alpha\tau}$ is the incidence angle modifier and ΔT is equal to T_i-T_a, i.e., inlet temperature to the collector minus ambient temperature. The following model for the incidence angle modifier is employed:

$$k_{\alpha\tau} = 1 - b_o\left(\frac{1}{\cos(\theta)} - 1\right) \tag{2}$$

where b_o is a constant and θ is the angle of incidence. The useful energy extracted from the collectors is given by:

$$Q_u = F_R A\left[k_{\alpha\tau}(\tau\alpha)G - U_L(T_i - T_a)\right] \tag{3}$$

where F_R is the heat removal factor and τα is the tranmsittance-absorptance product.

The total useful energy for the whole year is obtained from:

$$Q_{u,a} = \sum_{d=1}^{365}\sum_{h=1}^{24} Q_u \quad (4)$$

and the auxiliary energy required, Q_{aux} is:

$$Q_{aux} = Q_{load} - Q_{u,a} - Q_{loss} \quad (5)$$

where Q_{load} is the energy required by the load and Q_{loss} is the energy lost from the storage tank and pipes.

As can be seen from the above equations the energy obtained from the solar collector field depends on the inlet temperature to the collector T_i, which depends on the load pattern and the losses from the storage tank and pipes.

The present investigation was conducted through the use of a Typical Meteorological Year (TMY) for Nicosia, Cyprus. This was generated from hourly measurements, of solar irradiance (global and diffuse on horizontal surface), ambient temperature, wind speed and direction, and humidity ratio, for a seven-year period, from 1986 to 1992 using the Filkenstein - Schafer statistical method [2]. The measurements were recorded by the Cyprus Meteorological Service at the Athalassa region, an area at the suburbs of the town of Nicosia. The TMY is considered as a representative year for the Cypriot environment. Using this approach the long-term integrated system performance can be evaluated and the dynamic system's behaviour can be obtained.

The economic analysis is performed in order to obtain the annual cost of the system and the net solar energy price (NSEP). The investment cost of the solar system is obtained from:

$$C_s = C_f + C_a A + C_v V \quad (6)$$

where C_f is the area independent cost, C_a is the area dependent cost, both applied to the solar collectors, and C_v is the cost of storage per m^3 of storage volume.

For the operation cost (C_o), maintenance and parasitic costs are considered. The former are estimated to be 2% of the initial investment. The latter account for the energy required (electricity) to drive the solar pump. The total annual cost is given by:

$$C = C_s \alpha + C_o \quad (7)$$

where α is the annuity given by:

$$\alpha = \frac{i(1+i)^N}{(1+i)^N - 1} \quad (8)$$

where i is the inflation rate and N is the number of years the system is operational, i.e., life of the system (20 years).

The net energy collected from the collector (Q_{net}) is obtained by subtracting the annual amount of auxiliary energy from the annual amount of conventional fuel required to cover the load for a fuel only system, i.e.,

$$Q_{net} = Q_{load} - Q_{aux} \quad (9)$$

Finally, the solar energy price, in £/kWh, is obtained by:

$$NSEP = \frac{C}{Q_{net}} \quad (10)$$

4. SYSTEM SIMULATION

With the aid of TRNSYS and the TMY for Nicosia, Cyprus, a number of simulations were performed. The optimisation parameter used is the net solar energy price (NSEP), i.e., the optimum system is the one, which gives minimum NSEP. As the load pattern is the same for all the cases considered, the parameters on which the performance of the system depends are the collector area and the storage tank volume. By increasing the collector area more solar energy is collected but the solar system costs more, whereas by increasing the storage tank volume more energy can be stored but the losses from the storage tank to the environment are increased. Small storage tank volumes exhibit lower environmental losses but increase the losses of energy though the relief valve as they can easily reach the relief valve setting.

Various simulations were performed for a number of collector areas and a range of storage tank volumes. In this way, a database is created with the combination of the collector area and storage volume on auxiliary energy required by the system, for various cases. The idea is to create a small database of combinations of values and use these values to create a fitness function for the NSEP. The data collected from the present system from 7 runs of TRNSYS are shown in Table 3. As can be seen from Table 3 as the collector area and the storage tank are increased the auxiliary energy required is reduced because the solar system can collect and store more solar energy, but the system costs more. It is important to cover the low and upper extreme ends of the cases to be investigated plus some intermediate values. This is important for two reasons; first, the neural network learns all the range of possible values and thus it will not need to extrapolate and second the scaling functions for the input and output data need to be determined only once, based on the minimum and maximum values of the parameters in the dataset. For those extreme conditions, rules of thump can be used such as storage tank between 40 and 300 litres per square meter of collector area.

Table 3 Training dataset

Area (A) [m^2]	Storage tank volume (V) [m^3]	Q_{aux} [$kJx10^8$]
100	10	6.447
100	30	6.288
200	30	4.623
300	20	3.606
400	10	3.058
500	10	2.508
500	30	2.124

The traditional method for finding the optimum solution is to perform many runs of TRNSYS and decide by trial and error the characteristics of the system, which gives the optimum solution. This method may lead to solutions far away from the optimum as the method strongly depends on the intuition of the engineer and the peculiarities of the system and of the site weather conditions. Additionally depending on the computer system frequency and the complexity of the system, each run might need several minutes to hours to be performed. Thus, it is required for the present system not only to be able to find the optimum solution but also to reduce the time required for such task to be performed.

5. METHOD DESCRIPTION

A different approach to optimise the system based on artificial intelligence systems is suggested in this paper. The idea is to use genetic algorithms to find the optimum values of collector area and storage tank volume, which will minimise the net solar energy price of the system. For this purpose, an accurate correlation of collector area and storage tank volume on the auxiliary energy required to cover the load is required. In this work a neural network is used which when trained gives a complex polynomial equation correlating these parameters. The neural network and the genetic algorithm are briefly described in this section.

5.1 Group Method Data Handling (GMDH) Neural Network (NN)

There are various methods that can be used to model the data, i.e., correlate the collector area and storage tank volume with the auxiliary energy required. These could be based on simple regression analysis, multiple regression analysis, neural networks and many others. In this work, the neural network method is selected because of its accuracy. One type of neural networks, which is very suitable for the present application, is the group method of data handling (GMDH) neural network, which was used to model the data. GMDH works by building successive layers with links that are simple polynomial terms. These polynomial terms are created by using linear and non-linear regression. The initial layer is simply the input layer. The first layer created is made by computing regressions of the input variables and then choosing the best ones. The second layer is created by computing regressions of the values in the first layer along with the input variables. Again, only the best are chosen by the algorithm called survivors. This process continues until the network stops getting better (according to a prespecified selection criterion)

The resulting network can be represented as a complex polynomial description of the model. The complexity of the resulting polynomial depends on the variability of the training data. In some respects GMDH, it is very much like using regression analysis, but it is far more powerful than the latter. GMDH can build very complex models while avoiding overfitting problems. A by-product of GMDH is that it recognizes the best variables as it trains.

The GMDH network is implemented with polynomial terms in the links and a genetic component to decide how many layers are built. The result of training at the output layer can be represented as a polynomial function of the inputs. The layer building GMDH procedure continues as long as the evaluation criteria continue to diminish. GMDH algorithm then checks if this is so and continues or stops training. There may also be other conditions, which determine when training is stopped.

5.2 Genetic Algorithm

A genetic algorithm is an optimum search technique based on the concepts of natural selection and survival of the fittest. It works with a fixed-size population of possible solutions of a problem, called individuals, which are evolving in time. A genetic algorithm utilizes three principal genetic operators: selection, crossover, and mutation. Genetic algorithms (GA) are suitable for finding the optimum solution in problems were a fitness function is present. Genetic algorithms use a "fitness" measure to determine which of the individuals in the population survive and reproduce. Thus, survival of the fittest causes good solutions to progress. A genetic algorithm works by selective breeding of a population of "individuals", each of which could be a potential solution to the problem. In this case, the genetic algorithm is seeking to breed an individual that in this case minimizes the net solar energy price of the IPH system.

During each step (called a generation) in the reproduction process, the individuals in the current generation are evaluated by a so-called fitness function value, which is a measure of how well the individual solves the problem. Then each individual is reproduced in proportion to its fitness: the higher the fitness, the higher its chance to participate in mating (crossover) and to produce an offspring. A small number of newborn offsprings undergo the action of the mutation operator. After many generations, only those individuals who have the best genetics (from the point of view of the fitness function) survive. The best individual provides an optimum or near optimum solution to the problem.

The larger the breeding pool size, the greater the potential of it producing a better individual. However, the networks produced by every individual must be applied to the test set on every reproductive cycle, so larger breeding pools take longer time. After testing all of the individuals in the pool, a new "generation" of individuals is produced for testing.

During the setting up of the GA the user has to specify the adjustable chromosomes, i.e., the parameters that would be modified during evolution to obtain the minimum value of the fitness function. In this work, these are the collector area and the storage tank volume. Additionally the user has to specify the ranges of these values. It is important that the ranges specified to be the same as the extreme cases used when setting up the neural network. In the present work, these are equal to 100-500m^2 for the collector area and 10-30m^3 for the storage tank volume.

6. OPTIMUM SOLAR SYSTEM

The training dataset (Table 3) were learned by the NN with very good accuracy (R^2-value equal to 0.9986). A plot of the actual (modelled) and network predicted data is shown in Fig. 3. It should be noted that a multiple linear regression method could only produce correlation with R^2=0.9653 which is not acceptable for the kind of predictions required in this type of problems. A similar figure for data, which are completely unknown to the network and used for validation of the ability of the network to produce accurate results is shown in Fig. 4. It should be noted that in this case also the network provided good predictions with R^2=0.9906.

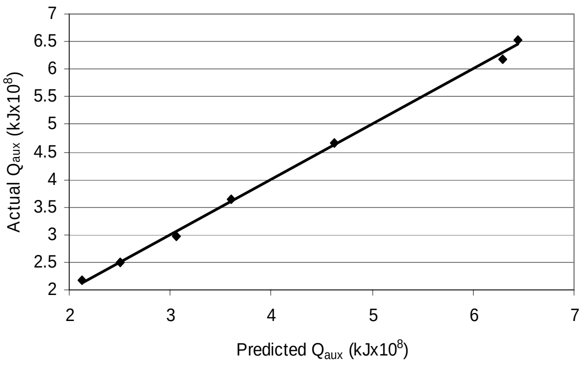

Fig. 3 Comparison of actual (modelled) and NN predicted data for the training dataset

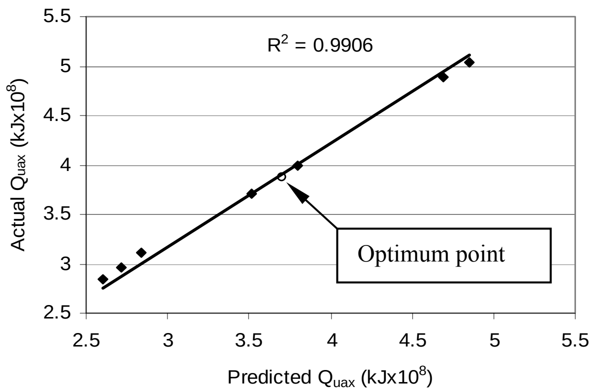

Fig. 4 Comparison of actual (modelled) and NN predicted data for the validation dataset

The final equation obtained from GMDH is:

$$Y=- 0.29565 - 0.93105 N_1 + 0.32545 N_1^2 - 0.07628 N_2^3 \quad (11)$$

All the data required by the GMDH need to be scaled from –1 to 1. Therefore, parameters N_1 (collector area), N_2 (storage volume) as well as Q_{aux}, obtained form Y (Eq. 11) needs to be scaled in the same interval. This is done with:

$$y_i = \frac{2(x_i - x_{\min})}{x_{\max} - x_{\min}} - 1 \quad (12)$$

Figure 5 shows the fitness function against the number of generations during the running of the genetic algorithm program. As can be seen the best fitness is found at the 45th generation, which is very fast. It should be noted however that the shape of the graph depends on the initial conditions. The particular case presented is for initial collector area equal to 100m^2 and initial storage tank volume equal to 30m^3. As can be seen after about 12 generations values near the optimum ones have been obtained.

The optimum solution reached is; collector area equal to 227.7m^2 and storage tank size equal to 13.2m^3. The net solar energy price for this solution is 0.038471 £/kWh, whereas the net solar energy price of a more practical solution with A=230m^2 and V=13m^3 are 0.038473 £/kWh. This performance of the GMDH network to predict Q_{aux} for this case is very good as it is shown by the point marked on Fig. 4, which compares the auxiliary energy predicted with the neural network with that obtained from TRNSYS. This case is checked on purpose to evaluate the prediction accuracy for the optimum case.

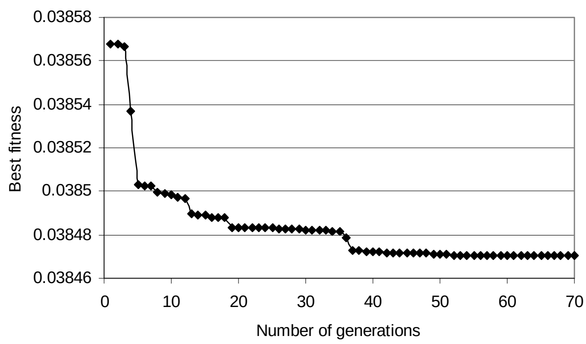

Fig. 5 Evolution of the best fitness

It should be noted that both the training of the neural network and the genetic algorithm program required just a few seconds to be performed whereas each run of TRNSYS requires 2.5 minutes, all on a Pentium 400 MHz machine. Thus, by reducing the number of runs of TRNSYS the time required to find the optimum solution is greatly reduced and the solution reached is more correct than by using the traditional trial and error method which most of the times relies on the intuition of the user to find a good solution.

7. CONCLUSIONS

A new optimisation method is presented in this paper. Initially the system is modelled with TRNSYS and GMDH neural network, which was trained at high accuracy, as the R^2-value obtained is 0.9986. Subsequently a genetic algorithm was used to select the combination of system components, i.e., solar collector area and storage tank volume, which minimises the net solar energy price of the system. It is believed that the present method would decrease the time required by design engineers to find the optimum solution and in many cases to obtain a selection, which could not be easily spotted by traditional trial and error method, which in most of the cases depends on the intuition of the engineer.

REFERENCES

1. TRNSYS program Manual. Solar Energy Laboratory, University of Wisconsin, Madison, USA, 1996.
2. Kalogirou S., 2003. Generation of a Typical Meteorological Year (TMY-2) for Nicosia, Cyprus, *Renewable Energy*, Vol. 28, No. 15, pp. 2317-2334.