Bulgarian Chemical Communications, Volume 48, Special Issue E (pp. 88 - 95) 2016

Parameter estimation of borehole thermal properties using artificial intelligence methods

R. K. Popov[1*], A. G. Georgiev[2,4], D. B. Dzhonova-Atanasova[3]

[1]*Faculty of Physics, Plovdiv University "Paisii Hilendarski", 24 Tzar Assen Str., 4000 Plovdiv, Bulgaria*
[2]*Faculty of Mechanical Engineering, Technical University - Sofia, Plovdiv Branch, 25 Tsanko Diustabanov Str., 4000 Plovdiv, Bulgaria*
[3] *Institute of Chemical Engineering, Bulgarian Academy of Sciences, Acad. G. Bonchev Str., Bl. 103, 1113, Sofia, Bulgaria*
[4]*European Polytechnic University, Dept. of "Green Energy", 23 Kiril and Metodiy str., 2300 Pernik, Bulgaria*

There are many estimation techniques, which are used in Thermal Response Test (TRT) data analysis. The commonly used models, Line Source Model, Cylindrical Source Model, numerical models do not take into account the nonlinear system effects like for example the phase change. The present work suggests the use of the input/output black box identification technique for TRT data analysis. A nonlinear autoregressive exogenous (ARX) model structure and stochastic search algorithms are used to estimate model parameters. Artificial intelligence techniques, Genetic Algorithm and Particle Swarm Optimization Algorithm are employed to avoid local maxima problems. The study is based on data sets obtained during real TRT tests without phase change effects. All analyses are performed in MATLAB environment. The purpose of this paper is to verify that the proposed algorithms are suitable for processing of TRT data with the aim of future identification of thermal parameters of boreholes with phase change effects. The given solution is also useful when common techniques fail in search for the global optimum if the search space is not differentiable or linear in the parameters.

Keywords: artificial intelligence methods, parameter estimation, thermal response test, thermal properties, MATLAB

INTRODUCTION

Thermal Response Test (TRT) is an internationally approved technique to identify geothermal underground parameters like effective ground thermal conductivity and borehole thermal resistance. It is considered to be the method which gives the highest accuracy of evaluation. Generally, these tests are performed with heat injection, using the same assumed power level as the one planned by the Borehole Heat Exchanger (BHE) system. For first time TRT was presented by Mogensen – his installation was designed as an immobile system [1]. Later, TRT was developed as a mobile measurement installation at the Oklahoma State University (USA) by Austin [2] and at the Technical University Lulea, Sweden [3]. Now, this type of measurement is used also in Germany [4], Canada, Norway, Netherlands, England, Turkey [5] and Chile [6]. Several TRTs are done in Bulgaria, too. Some activities in Bulgaria preceded the first official TRT [7].

Algorithm Overview

The genetic algorithm (GA), first proposed in [8], is a method for solving optimization problems which are not easy manageable by standard optimization methods, e.g. when the objective function is discontinuous, non-differentiable, stochastic, or highly nonlinear. It is used for both constrained and unconstrained optimization problems.

GA concept is taken from the principles of the evolution of the species by survival of the fittest. Like in a population of organisms [9] the solutions are created by reproduction of solutions and compete for survival in the next iteration. At each step the randomly selected individuals of the current population become parents and produce the children of the next generation. Over successive generations, the population improves to the optimal solution.

In the initial population *P(0)*, encoded randomly by strings [10], the more fit elements of each generation *(t)* are selected and processed by the basic genetic operators, crossover, and mutation, to create the next generation. On the basis of the evolution principles the best chromosome of a candidate solution is preserved. Thus the GA uses three main types of rules at each step towards the next generation of the current population: *Selection rules* select the individuals, called *parents*, which contribute to the population at the next generation; *Crossover rules* combine two parents to form children for the next generation; *Mutation rules* apply random changes to individual parents to form children.

* To whom all correspondence should be sent:
rum_pop@yahoo.com

The computational procedure is illustrated by the following GA pseudo code [11]:

```
Procedure GA
begin
   t=0
   initialize P(t)
   evaluate P(t)
   while not satisfy stopping rule do
      begin
        t= t+1
        select P(t) from P(t1)
        alter P(t)
        evaluate P(t)
      end
   end
```

In the present work, the GA is employed for model identification in Autoregressive integrated moving average (ARIMA) models because of the following very powerful characteristic: simultaneous searching of a population of points, not a single one, which results in finding the approximate optimum quickly and without falling into a local optimum. Moreover GA is free from the limitation of differentiability, unlike other mathematical techniques.

Concepts of the Particle Swarm Optimization Approach

As the neural network is a simplified model of a human brain and the genetic algorithm is inspired by the biological evolution, an optimization approach known as swamp intelligence is inspired by the collective behavior of the individuals in a social system when interacting with the environment and each other. Two of the popular swamp inspired methods of computational intelligence area are Ant Colony Optimization (ACO) and Particle Swarm Optimization (PSO).

The first is inspired by the behavior of ants and has many successful applications in discrete optimization problems [12].

The PSO concept was originally developed for graphical simulation of the choreography in a bird flock or a fish school and employed afterwards as an optimizer. It was proposed in [13] as a population based stochastic optimization technique, inspired by social behavior of bird flocking or fish schooling. PSO simulates the bird flocking according to the following scenario: a group of birds are randomly searching only one piece of food in an area. They do not know where the food is, but do know how far it is in each iteration. The best searching strategy is to follow the bird nearest to the food.

In solving optimization problems by using this scenario, each solution in the search space is “bird”, called “particle”. All the particles have fitness values, which are evaluated by the fitness function to be optimized. The particles have flying velocities and fly through the problem space by following the current optimum particles.

PSO shares many similarities with evolutionary computation techniques. Like in the GA, the system is initialized with a population of random solutions and searches for optima over successive generations. However, unlike the GA, the PSO has no evolution operators such as crossover and mutation. The potential solutions, called particles, fly through the problem space by following the current optimum particles.

At the first step in PSO the position and velocity of the particles are randomly initialized. In every iteration, each particle is updated by following two "best" values. The first one is the best position (fitness) it has achieved so far denoted by p_{best}. (The fitness value is also stored). The second one is the best position, obtained so far by any particle in the population, called global best value and denoted by g_{best}. Each particle communicates with its topological neighbors and knows the best position found so far by any of them. If a particle takes part of the population as its neighbors, the best value is a local best and is called l_{best}.

After finding the two best values, the particle updates its velocity and position by the following equations (1) and (2)

$$v[\,] = v[\,] + c1 * rand(\,)*(pbest[\,] - present[\,]) + + c2*rand(\,)*(gbest[\,] - present[\,]) \quad (1)$$

$$present[\,] = present[\,] + v[\,], \quad (2)$$

where $v[\,]$ is the particle velocity, $present[\,]$ is the current particle (solution). $pbest[\,]$ and $gbest[\,]$ are defined as stated before. $rand(\,)$ is a random number between (0,1). $c1$, $c2$ are learning factors. Usually $c1 = c2 = 2$.

The pseudo code of the procedure is as follows:

```
for each particle
  Initialize particle
end
do
    for each particle
        Calculate fitness value
        if the fitness value is better than the best
           fitness value (pbest) in history
        set current value as the new pbest
```

end
Choose the particle with the best fitness value of all the particles as the gbest
for each particle
Calculate particle velocity according equation (1)
Update particle position according equation (2)
end
while maximum iterations or minimum error criteria is not attained

On each dimension a maximum velocity V_{max} is specified by the user as a limit of the particles' velocities.

INSTALLATION FOR IMPLEMENTING OF THERMAL RESPONSE TEST

An original construction of a Thermal Response Test rig has been built recently at the Technical University of Sofia, branch Plovdiv. The equipment for implementing of TRT is situated on a mobile trailer consisting of two parts. The first one contains the working installation. The second part is formed as a living room for the investigators, who will implement the in-situ tests. The scheme of the installation is shown in Fig.1. It consists of the following parts: electrical boiler 1, calorimeter 2, pressure watch 3, expansion tank 4, thermo – manometer 5, filter 6, circulation pump 7, de-aeration pipe 8, quick couplings 9, valves 10 and electrical unit 11. A 41,10m deep hole was drilled on the territory of the Technical University – Sofia, branch Plovdiv. The borehole has a diameter of 0.18 m and has been backfilled with 11% bentonite and 2% cement solution. There are two temperature sensors Pt 100 for measuring the inlet and outlet borehole temperatures. There are five other temperature sensors placed at different depths inside the borehole.

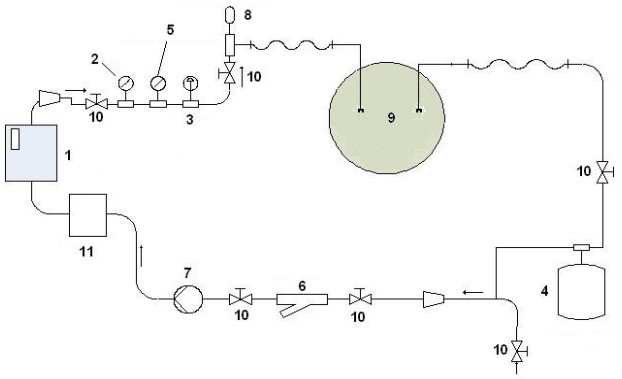

Fig.1. Installation setup

THERMAL RESPONSE TEST - EXPERIMENT AND EVALUATION

The tests were implemented in January, March and April 2009. The following parameters were measured: the inlet and outlet fluid temperatures of the borehole, the ambient air temperature and five temperatures in the borehole body. The electrical power (about 1500 W) was controlled and maintained constant during the whole test. The flow rate of water was 4.06 l/min.

There are different methods to evaluate the experimental data and to calculate the unknown thermal conductivity λ and borehole thermal resistance. The Line Source Model (LSM) is the widely used and simplest method [3]. The delivered heat is considered as coming from an infinite line source (the borehole). The following equation represents the heating process:

$$T_{f,m} = \frac{Q}{4\pi\lambda H}\ln(t) + \left[\frac{Q}{H}\left(\frac{1}{4\pi\lambda}\left(\ln\left(\frac{4a}{r_b^2}\right) - \gamma\right) + R_b\right) + T_s\right] \quad (3)$$

for $t \geq \frac{5r_b^2}{a}$

where $T_{f,m} = (T_{f,i} + T_{f,o})/2$ - mean fluid inlet/ outlet temperature*, K;*

Q – delivered heat power, W;
λ – thermal conductivity, W/mK;
H – borehole depth, m;
t – time from start, s;
a – thermal diffusivity, m^2/s;
r_b – radius of the borehole, m;
$\gamma = 0.5772$ – Euler's constant;
R_b – borehole thermal resistance, mK/W;
T_s – initial soil temperature, K.

Test Evaluation

The average undisturbed ground temperature T_s is a key parameter in Eq. (3) and should be measured prior to the test start before switching on the electrical heaters, when the borehole is at thermal equilibrium with its surroundings. T_s was determined by pumping the heat carrier fluid out of the borehole pipes and measuring its outlet temperature over a time of 10s. T_s was then calculated as the average of the measurement data. In the presented experiments, T_s was found to be about 16.3°C. As soon as T_s is measured the electric heaters are switched on and a constant heat starts to be injected in the BHE.

The flow rate was fixed at a constant value during the tests. The electrical heater power and the electrical power of the circulating pump were maintained constant automatically. For the

experiment purpose the installation was filled with water and a pressure of 2.2 bars was established. Most of the main characteristics of the performed Bulgarian TRTs in 2009 are shown in Table 1.

Table 1. Main characteristics of the Bulgarian TRTs carried out in 2009

Properties	January	March	April
Date	11-21	16-24	13-17
Duration, days	10	8	4.3
Flow rate, l/ min	4.06	3.83	4.6
Electrical heater power, W	1500	2000	1500
Circulating pump power, W	100	100	100
Water pressure, bar	2.2	2.2	2.2
Undisturbed soil temperature, °C	16.3	16.3	16.35
Measuring interval step, s	60	60	60

The control of the test rig is the most challenging part of the system. All data are automatically controlled by a specially designed system for the laboratory trailer needs, installed on the control board. The system is fully automatic and writes down all measured data in text files. If appropriate software is available, the data collecting process could be visualized in real time or after finishing the experiment. In the experiments the aim of a constant heat flow was realized by a constant frequency control of the circulation pump and boiler.

The application of equation (3) to the real problem is connected with a systematic error, which diminishes with time and increases with borehole radius. Eq. (3) can be rewritten for evaluation purposes in the form of:

$$T_{f,m}(t) = k_1 \ln(t) + k_2 \qquad (4)$$

Further on the test evaluation is based on Eq. (4), which is fitted to the experimental data, and the curve constants k_1 and k_2 are determined. Usually a Least Squares Method (LSM) is used to calculate k_1 and k_2. The thermal ground conductivity λ and the borehole resistance R_b can be adequately calculated, by comparing Eqs. (3) and (4) when the curve parameters k_1 and k_2 are known. In this article instead of LMS method, the GA and PSO search approaches are employed to fit the data and to compare the agreement.

Test Results

The mean fluid temperature in the borehole (average temperature of the inlet and outlet fluid temperature of the borehole) and the ambient temperature are shown for the test done in January in Fig.2. The figure shows that the experimental fluid temperature rises slow and with small deviations.

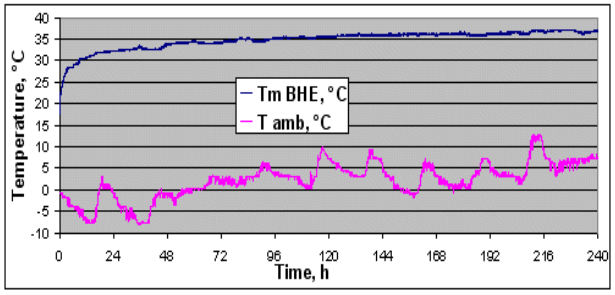

Fig.2. Profile of the mean fluid temperature in the borehole and the ambient temperature in January 2009

The number of the data points is very large (about 14400). In order to decrease the processing time of the GA and PSO stochastic search algorithms the number of the points was reduced by a factor of 30, using averaged values. The averaged trends of the same temperatures are showed in Fig.3.

Test Data Evaluation, Using Standard LSM Regression Method

The Line Source Model gives more exact temperature estimations for longer terms of time. It takes some hours to the real BHE to behave as an ideal line source. Therefore, usually the data correspondent to the first 7 to 30 hours of experiment is not taken into account in the analysis.

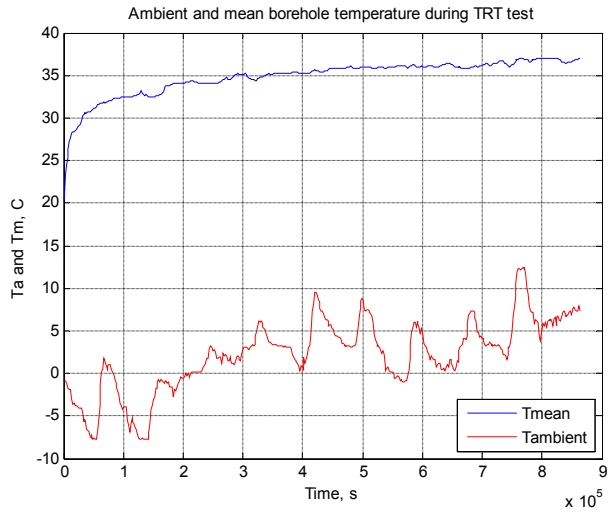

Fig.3. The reduced number of values of fluid mean temperature in the borehole and the ambient temperature in January 2009

In the present work, this period is accepted to be 20 hours, as it depends on the estimated data. Fig.4 shows the logarithmic time dependence of the temperature and the slope of the associated regression line. As stated previously, the thermal conductivity λ is related to the slope of the resulting line by Eq. (4). The resulting value calculated

during the test for λ is *0.83* W/mK and for R_b – *0.532* mK/W.

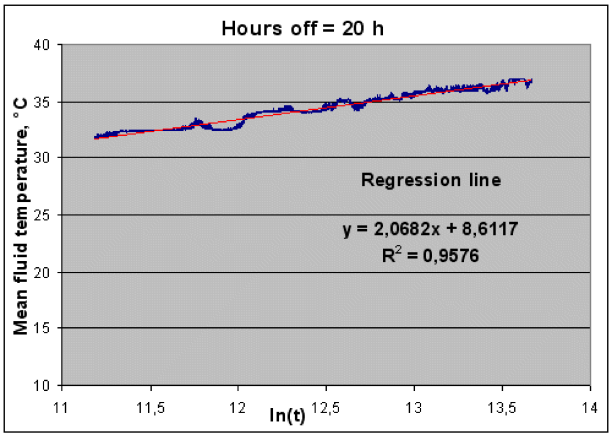

Fig.4. Logarithmic time plot of the mean temperature for the entire test length in January 2009 (excluding the first 20 hours)

Test data evaluation, using artificial intelligence methods

First we have to choose the input factors, which are to be varied. In GA algorithm they are:

- Crossover_p – Crossover probability parameter;
- Mutation_r – Mutation rate.

Output parameter is AvSAE variable (error) end value. It is the average value of the sum of absolute errors between data and model outputs at the end of the estimation process. If 50 generations are performed in GA search, then last column in variable AvSAE (50) is used. In the next step central composite experiment design is performed. Only one central point is used in case of obtaining uniform value distribution in the parameter space. The Matlab function "ccdesign" have been used for this purpose. After that factor variance bounds are chosen:

- Crossover_p = 0 … 1;
- Mutation_r = 0.05 … 0.5

Then matrix table is multiplied by factor n=10 in case of running not only 9 but 90 tests, because output value AvSAE ( ) is statistically dependent on the run. Now parameters are bounded to parameter space, using Matlab function "coded2real". Next step is to randomize runs' order in the test matrix. The positions of the input variables in the parameter space are shown in Fig.5.

The varied factors in the PSO algorithm are chosen to be:

- correction_factor;
- inertia.

The bonds chosen in this task are 1.6 to 2.4 for the Correction factor and 0.4 to 0.8 for the Inertia parameter. The central composite design for the PSO – test and the factors' variation in a Monte Carlo simulation are presented respectively in Fig.6.

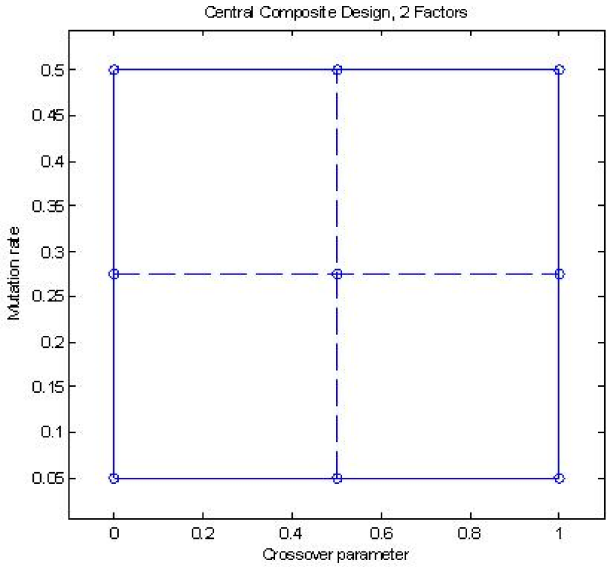

Fig.5. Positions of the input variables in the parameter space in GA tests

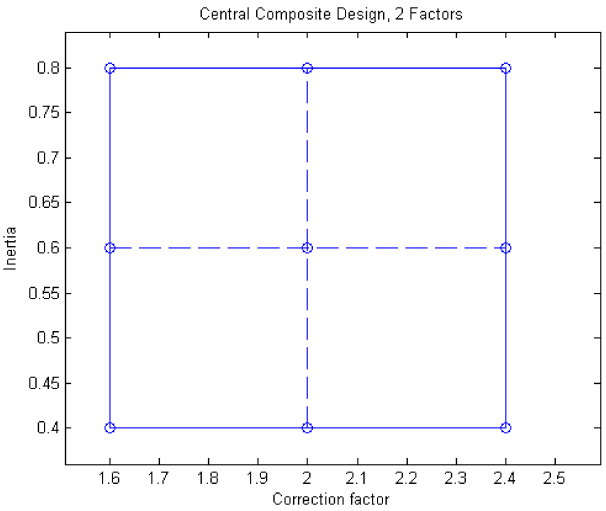

Fig.6. Positions of the input variables in the parameter space in PSO tests

Genetic algorithm-based search function developed in Matlab

The Genetic Algorithm-based search procedure is coded as a Matlab – function with the next syntax:

[A B AvSAE] = ga_proc(err, popsize, crossover_p, mutation_r, n_generations),

where:

A is a vector, containing evaluated 'a' parameters;
B is a vector, containing evaluated 'b' parameters;
AvSAE is a vector of average sum of absolute error of the model for all passed generations;
err is an error tolerance;
popsize is the size of population;
crossover_p is the crossover probability parameter;
mutation_r is the mutation rate;
n_generations is the number of generations.

Particle Swarm Optimization algorithm function developed in MATLAB

The Particle Swarm Optimization -based search algorithm is written as a MATLAB – function with the following syntax:

[A B AvSAE] = pso_proc (swarm_size, correction_factor, inertia, iterations, err);, where:

A is a vector, containing evaluated 'a' parameters;
B is a vector, containing evaluated 'b' parameters;
AvSAE is a vector of average sum of absolute error of the model for all passed generations;
err is an error tolerance;
swarm_size is the size of the swarm;
correction_factor is the particle velocity correction factor;
inertia is the inertia factor
iterations is the number of iterations. number of generations.

Data evaluation using statistical search approaches: GA and PSO

The 90 data fitting experiments were performed using GA and the same number of tests was carried out for the PSO approach. The results for the first 10 tests in each case are listed in the Table 2 (GA) and Table 3 (PSO).

Table 2. Monte Carlo simulation test results using GA

Run Number	Cross-over_p	Mutation _r	AvSAE_ fin_value	Rb	Lambda
1	1	0.050	0.91366	0.19607	0.88152
2	0.2	0.050	1.03060	0.10000	0.50000
3	1	0.275	0.22162	0.47185	0.87500
4	1	0.275	0.43971	0.10000	0.67388
5	0.6	0.275	0.94431	0.68193	2.44960
6	0.6	0.500	0.82120	0.60871	1.46410
7	0.2	0.050	1.22590	0.78088	5.48230
8	1	0.275	0.28684	0.10000	0.74713
9	0.2	0.500	0.40257	0.27394	0.68405
10	0.2	0.050	1.04250	0.10000	0.50000

Table 3. Monte Carlo simulation test results using PSO

Run Number	Correction factor	Inertia	AvSAE_ fin_value	Rb	Lambda
1	2	0.4	0.21792	0.46962	0.90854
2	2.4	0.4	0.21815	0.65663	1.80560
3	1.6	0.4	0.63708	0.60304	1.49490
4	2	0.6	0.21785	0.46567	0.86166
5	2	0.8	0.21785	0.46085	0.85269
6	1.6	0.6	0.21785	0.45883	0.84906
7	2.4	0.6	0.21790	0.46255	0.80732
8	2.4	0.6	0.21790	0.59181	1.25950
9	2.4	0.4	0.21786	0.45730	0.84441
10	1.6	0.6	0.21785	0.45608	0.84335

The GA tests duration was measured to be 24:52,256 min and the duration of the PSO tests was 21:05,385 min.

3D combined scatter plot of AvSAe versus GA parameters (factors) during the tests is showed in Fig.7 and Fig.8. Best values for the crossover parameter and the mutation rate are 1 and 0,5. For the correction factor and the inertia (in PSO) the combination of the values 2.4 and 0.6 reports lower error rate. The scatter plots show that all the factors are significant.

In the next step all test results, reporting AvSAe values higher than 0.3, were rejected. The calculated average values of the remaining estimated parameters are:

- in GA - tests:

$R_b = 0.4181513$ *mK/W*; $\lambda = 0.874063$ *W/mK;*

- in PSO - tests:

$R_b = 0.49098$ *mK/W*; $\lambda = 0.92624$ *W/ mK.*

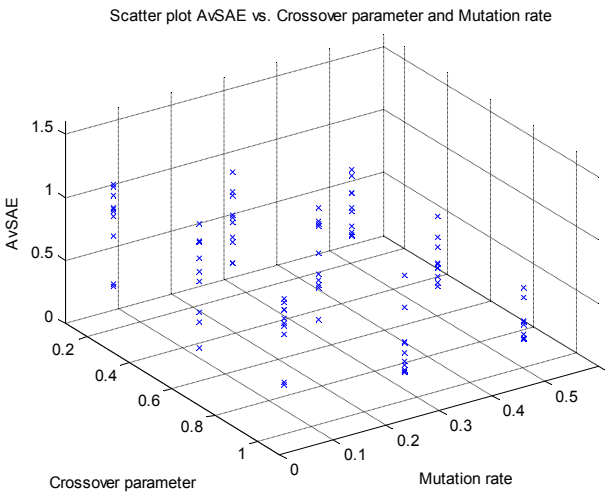

Fig.7. Positions scatter plot of AvSAE vs. crossover parameter and mutation rate in GA tests

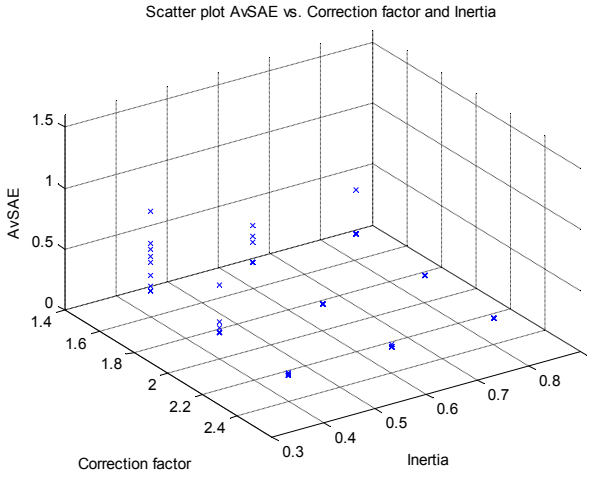

Fig.8. Scatter plot of AvSAE versus correction factor and inertia in PSO tests

The values of AvSAE (error) for both algorithms PSO and GA for the tests with the same number are compared in the Fig.9.

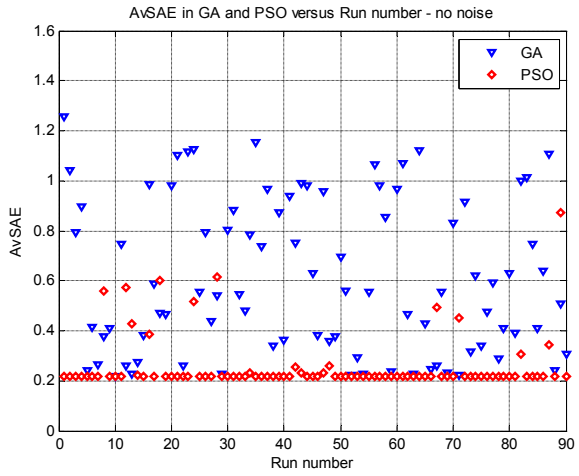

Fig.9. Value of AvSAE for both algorithms PSO (diamonds) and GA (triangles)

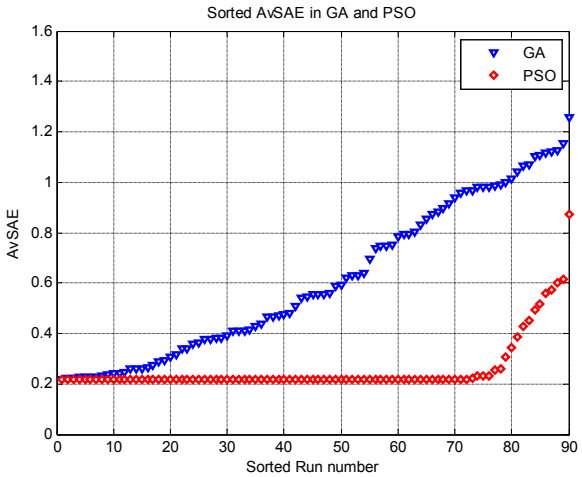

Fig.10. Sorted by values of AvSAE for both algorithms PSO (diamonds) and GA (triangles)

The next Fig.10 shows sorted values of AvSAE for both algorithms PSO (in red) and GA (in blue). It can be mentioned, that PSO procedure reports more accurate estimation of the system parameters.

CONCLUSIONS

The estimated values of the thermal properties, using different algorithms are presented in Table 4.

Table 4. Estimated values of the thermal properties, using different algorithms

Algorithm	R_b **, mK/W**	λ **, W/ mK**
LSM linear regression	0.5320000	0.830000
GA	0.4181513	0.874063
PSO	0.4909800	0.926240

The results demonstrate a good accordance in estimates and possibility to apply GA and PSO – based approaches in TRT data evaluation procedure. In the reported case no phase change process occurs in the BHE and the main advantages of these two artificial intelligence techniques are not demonstrated in their full range. So an additional experimental investigation should be performed to obtain data from the BHE, working with CO_2, especially in a freezing/melting process. That will give the possibility to identify thermal parameters of Phase Change Materials (PCM) and Slurries, used in thermal storages.

Based on the results of the parameter estimation experiments additionally the next conclusions are to be drawn:

1. PSO algorithm reports better performance than GA. The calculation time for the two procedures is similar because the time for model evaluation is determining, equal in both cases.
2. ARX Model structure (linear or nonlinear) is well suited to parameter estimation procedure, using usual or Evolutionary algorithms;
3. Population size of 30 citizens is well enough for low order linear ARX model parameter estimation procedure, using GA and PSO – based approaches;
4. When using GA and PSO – based approaches a number of 30 iterations gives in the linear case a good estimation accuracy AvSAE = 0.2 in our experiments;
5. When using GA and PSO – based approaches accuracy and efficiency become statistical parameters, which depends on the initial population generation and mutation variance;
6. For low order linear ARX model parameter estimation procedure the GA – based approach

gives a little bit higher performance than PSO – based approach;
7. For low order linear ARX model parameter estimation procedure the PSO – based approach gives much more higher accuracy than GA – based approach;
8. Significance analysis in all cases has shown that all varied parameters are significant;
9. The best values for the Crossover parameter and the Mutation rate in GA algorithm are: 1.0 and 0.5;
10. The best values for the Correction factor and the Inertia rate in PSO algorithm are: 2.4 and 0.6.

ACKNOWLEDGEMENTS

This work has been supported financially by the research project 152ПД0045-24/ 26.03.2015 and the research project 152ПД0046-24/ 27.03.2015 of the Technical University of Sofia, which is gratefully acknowledged by the author.

REFERENCES

1 Mogensen P. (1983). Fluid to Duct Wall Heat Transfer in Duct System Heat Storages. Proc. of Int. Conf. on Subsurface Heat Storage in Theory and Practice, Stockholm, Sweden, June 6-8, 1983, p. 652-657.
2 Austin, W. A. (1998), Development of an In-Situ System for Measuring Ground Thermal Properties. Master's thesis. Oklahoma State University. Stillwater, Oklahoma.
3 Eklöf, C., and S. Gehlin, S.: TED – A Mobile Equipment for Thermal Response Test, Master's Thesis 1996:198E. Luleå University of Technology, Sweden, (1996).
4 Sanner, B.; Reuss, M.; Mands, E. and Müller, J. (2000). Thermal Response Test - Experiences in Germany. Proc. of Terrastock 2000, 8th Int. Conf. on Thermal Energy Storage, Stuttgart, August 28-September 1, 2000, Vol. 1, p. 177-182.
5 Paksoy, H., Gurbuz, Z., Turgut, B., Dikici, D. and Evliya, H.: Aquifer thermal storage (ATES) for airconditioningof a supermarket in Turkey. Proc. of World Renewable Energy Congress-VII 2002, Cologne, Germany, (2002), 10_n66.pdf.
6 Roth, P., Georgiev, A., Busso, A., Barraza, E. First In-situ Determination of Ground and Borehole Thermal Properties in Latin America, Renewable Energy, 29, (2004), 1947-1963.
7 Georgiev, A., Pekov, O., Angelov, A., Popov, R., Urchueguía, J. F., and Witte, H.: First steps of ground accumulation in Bulgaria. Proc. of the World Renewable Energy Congress-IX, Italy, Florence, (2006).
8 Holland J.H., Adaptation in natural and artificial system, Ann Arbor, The University of Michigan Press, 1975.
9 Kargupta, H., Smith, R. E., System identification with evolving polynomial networks, Proceeding of the 4th International Conference on Genetic Algorithm, University of California, San Diego, USA, 1991, pp. 370-376.
10 Kristinsson, K., Dumont, G., System identification and control using Genetic Algorithms, IEEE Trans. on Systems, Man and Cybernetics, 1992 22 (5), pp. 1033–46.
11 Kent Allen, James G. Williams: Encyclopedia of Computer Science and Technology: Volume 21 - Supplement 6: ADA and Distributed Systems to Visual Languages, CRC Press, 1989.
12 Dorigo, M. (1992). Optimization, Learning and Natural Algorithms, PhD thesis, Politecnico di Milano, Italy.
13 Kennedy, J.; Eberhart, R. (1995). "Particle Swarm Optimization". Proc. of IEEE Int. Conf. on Neural Networks IV. pp. 1942–1948.