Synergy International Conferences - Engineering, Agriculture and Green Industry Innovation
Gödöllő, Hungary, October 16-19. 2017.

MULTIPLE LINEAR REGRESSION BASED MODEL FOR THE TEMPERATURE OF THE UPPER, UNDISTURBED PART OF A SOLAR STORAGE

R. Kicsiny

Department of Mathematics, Institute of Environmental Systems, Faculty of Mechanical Engineering, Szent István University
Páter K. u. 1., Gödöllő, H-2100, Hungary
Tel.:+36 28 522-000/1413, E-mail: Kicsiny.Richard@gek.szie.hu

Abstract: Developing mathematical models for describing the temperature of solar storage tanks is of great importance for the practice, since the storage tank stores and provides the solar heat directly, in the form of hot water, for the consumer. In this study, a new, general and easy-to-apply multiple linear regression based model, called LR model, is proposed to predict the temperature at the upper, undisturbed part of solar storages. This model is likely one of the simplest linear black-box type models, which can still describe the transient changes of the upper temperature with a satisfactory precision. Comparing measured and simulated data on a real solar storage, the validation and the efficiency of the LR model is presented. The generality and the simple applicability of the model are also mentioned along with future research proposals.

Keywords: solar storage, upper temperature, mathematical modelling, black-box model, multiple linear regression

INTRODUCTION

Developing mathematical models for describing temperatures of solar storage tanks is of great importance for the practice, since these elements are unavoidable in any solar heating systems, where some heat should be stored in the form of hot fluid. There are two main categories of mathematical models for solar storages: physically-based models (or white-box models), which describe exact physical laws (on the basis of theory) and black-box models, which represent empirical correlations (based on experiences or measurements).

The difficulties of the exact physically-based modelling caused by the complex/complicated physical phenomena, like the effect of the inlet fluid flow on the thermal stratification (Meyer et al., 2000), can be often overcome by means of the empirical black-box models. The most frequent black-box model type may be the artificial neural network (ANN) in thermal engineering applications. Kalogirou et al. (1999) predict the useful heat extracted from a solar heating system and the temperature increase of the stored water with an ANN with a modelling error of 7-10%. This can be stated appropriate accuracy for such systems (Kalogirou, 2000). An ANN is worked out in (Géczy-Víg and Farkas, 2010) to model the temperatures of the layers in a storage tank for domestic hot water. In general, ANNs are rather arduous to apply because of the so-called training (or learning) process used for identification. For example, 30 943 data sets were used in (Géczy-Víg and Farkas, 2010) in a version of the trainings (identification) while only 2997 in the validation. Furthermore, many different back-propagation algorithms needed for training processes are available. The selection of the right one along with the proper number of layers and neurons, forming the ANN model itself, requires high experience and expertise of the user. The convergence of the training algorithm indicating the end of a training session may be also time-consuming.

Because of the above problems on arduousness, complexity and time consumption, the present work aims at establishing a simple and general but still precise black-box model, which can be used fast and easily for a wide range of storage tanks. The model is based on a well-known method of mathematical statistics, namely, the multiple linear regression (MLR). Based on studies in the literature, MLR is a missing black-box modelling technique in the field of storage tanks despite of its linear algebraic simplicity. MLR-based models have been worked out in recent works for other working components of solar heating systems, namely, for solar collectors (Kicsiny, 2016) and for pipes (Kicsiny, 2017). As proposing a new MLR-based model to predict the temperature of the upper part of a solar storage, called *LR model* in short, the present study can be considered as the continuation of these works.

The reason for the upper part is modelled in this work is that this part is less disturbed by the inlet flows to the storage (caused by the consumption load and the pump of the heating loop), and, accordingly, easier to model. Basically, a simple model, like the LR model, can be expected to describe relatively simple, undisturbed cases more efficiently.

LR MODEL

Fig. 1 shows the general scheme of the studied solar storage.

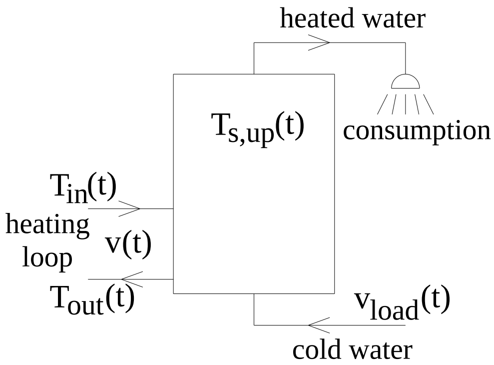

Fig. 1. Scheme of the solar storage

The storage tank can be heated up through a heating loop, within which the fluid enters the storage with temperature T_{in} and leaves it with T_{out}. A pump circulates the fluid in the heating loop with 0 or a prefixed constant flow rate value v according to on/off pump operation. Sometimes a consumer discharges some fluid from the storage with the flow rate v_{load}.

All time dependent variables T_{in}, $T_{s,up}$, v and v_{load} are measured periodically according to a time period Δt.

Because of the bounded propagation speed of physical effects and the bounded speed of measurements, the inputs of the LR model are $T_{in}(t-\tau)$, $v_{load}(t-\tau)$ and $T_{s,up}(t-\tau)$ with respect to the output (which is the modelled value of $T_{s,up}(t)$) at the current time *t*. Here τ is the time delay with respect to the effects of the inlets (more particularly T_{in} and v_{load}) to the interior of the storage tank. Clearly, the previously detected value of the temperature in the upper part of the storage (called simply *storage temperature* below), as some initial value of the model, has also essential effect on its current value $T_{s,up}(t)$. For simplicity, $T_{s,up}(t-\tau)$ is taken as this previous temperature to be considered in the model.

Considering the storage tank as a black-box, it can be admitted that distinct sub-models as parts of the LR model (as a black-box model) should be identified for significantly different operating conditions. More particularly, the storage tank behaves absolutely different if the pump is on (*v*>0) or off (*v*=0) permanently. Namely, under the same initial storage temperature, $T_{s,up}$ basically increases if the pump is on and decreases if the pump is off. Even, the effect of T_{in} can be neglected in case of permanently switched off pump, since there is no fluid flow into the storage from the heating loop. Considering a typical day, when the temperature increase of $T_{s,up}$ is significant (and the consumption load is not extremely high), four different operating cases should be distinguished according to Fig. 2.

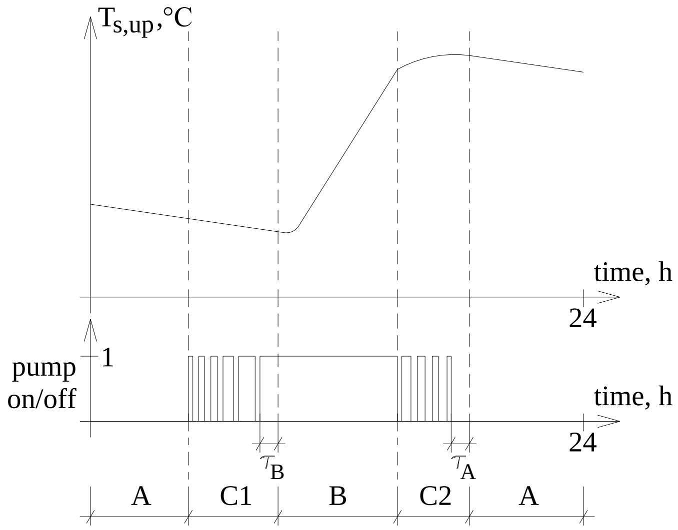

Fig. 2. Upper storage temperature and pump operation on a typical day

Cases A and *B* correspond to permanently switched off and switched on pump, respectively, while *Case C1* and *C2* correspond to frequent switch-offs and -ons. Since the storage temperature basically increases before the solar noon and decreases after it, it seems practical to divide *Case C* into two parts accordingly. The detailed specification of each *Case* can be found below (see also Fig. 2).

Case A: The pump is switched off permanently. This *Case* consists of the time period from the beginning of the day to the first switch-on of the pump and all such periods, which start at a time when the pump has been (permanently) switched off for exactly τ_A time and stops either at the time of the next switch-on or at the end of the day.

Case B: The pump is switched on permanently. This *Case* consists of all such periods, which start at a time when the pump has been (permanently) switched on for exactly τ_B time and stops at the next switch-off.

Case C1: Time periods besides *Cases A* and *B* before the solar noon.

Case C2: Time periods besides *Cases A* and *B* after the solar noon.

Remark: τ_A is the time which should be passed after a switch-off to go on with *Case A*. More particularly, τ_A is the time which is generally enough for the pump to become permanently off (and not alternating), and, this behaviour is considered as the characteristic feature of *Case A*. Similarly, τ_B is the time which should be passed after a switch-on to go on with *Case B*. More particularly, τ_B is the time which is generally enough for the pump to become permanently on (and not alternating), and, this behaviour is considered as the characteristic feature of *Case B*.

For the best possible modelling precision, distinct sub-models based on MLR are worked out for each operating case. In the sub-models of *Cases B*, *C1* and *C2*, T_{in} is among the inputs, since there is some fluid flow into the storage tank from the heating loop in accordance with the permanently or intermittently switched on pump. In the sub-model of *Case A*, T_{in} is neglected according to the permanently switched off pump.

The LR model is composed of Eqs. (1A), (1B), (1C1) and (1C2), which are simple linear algebraic relations representing the corresponding sub-models of the separate operating cases.

Case A:
$$T_{s,up,\mathrm{mod}}(t)=c_{load,A}v_{load}(t-\tau)+c_{s,up,A}T_{s,up}(t-\tau) \tag{1A}$$

Case B:
$$T_{s,up,\mathrm{mod}}(t)=c_{in,B}T_{in}(t-\tau)+c_{load,B}v_{load}(t-\tau)+c_{s,up,B}T_{s,up}(t-\tau) \tag{1B}$$

Case C1:
$$T_{s,up,\mathrm{mod}}(t)=c_{in,C1}T_{in}(t-\tau)+c_{load,C1}v_{load}(t-\tau)+c_{s,up,C1}T_{s,up}(t-\tau) \tag{1C1}$$

Case C2:
$$T_{s,up,\mathrm{mod}}(t)=c_{in,C2}T_{in}(t-\tau)+c_{load,C2}v_{load}(t-\tau)+c_{s,up,C2}T_{s,up}(t-\tau) \tag{1C2}$$

$c_{load,A}$, $c_{s,up,A}$, $c_{in,B}$, $c_{load,B}$, $c_{s,up,B}$, $c_{in,C1}$, $c_{load,C1}$, $c_{s,up,C1}$, $c_{in,C2}$, $c_{load,C2}$, $c_{s,up,C2}$, are constant parameters to be identified.

IDENTIFICATION AND VALIDATION

In this section, the LR model (Eqs. (1A), (1B), (1C1) and (1C2)) is applied to a real storage tank for identification and validation. The calculations needed have been done in the Matlab software (Etter et al., 2004).

In the identification, $T_{s,up,meas}(t-\tau)$ is used as $T_{s,up}(t-\tau)$. During the validation of the already identified LR model, the previously modelled value $T_{s,up,\mathrm{mod}}(t-\tau)$ is used as $T_{s,up}(t-\tau)$ when modelling $T_{s,up}(t)$. Measured $T_{in}(t-\tau)$ and $v_{load}(t-\tau)$ values are available both during the identification and the validation. According to the specification of Δt, the measurements happen at times $t = 0, \Delta t, 2\Delta t, 3\Delta t, ...$ Practically, the modelled value of $T_{s,up}$ (that is $T_{s,up,\mathrm{mod}}$) is determined in the LR model also at times $t = 0, \Delta t, 2\Delta t, 3\Delta t, ...$ Furthermore, for simplicity, $\tau = \Delta t$ is assumed in the LR model. *Case A* holds and $T_{s,up}(t-\tau) = T_{s,up,meas}(0)$ is used as measured initial condition in Eq. (1A) at $t = \tau$ (at the beginning of the day).

The real storage tank, which is to be modelled, is the solar storage of a measured solar heating system (Farkas et al., 2000) installed at the Szent István University (SZIU) in Gödöllő, Hungary. This storage will be called below *SZIU storage* in short. The SZIU storage contains preheated domestic water for a kindergarten at the campus of the university. The heat is transferred from a solar collector field into the SZIU storage by means of a heating loop equipped with a pump working in on/off operation. The measurements are carried out once a minute, that is, Δt =1 min. The volume of the SZIU storage is 2 m^3. Based on observations, τ_A and τ_B can be set 10 min. These and other important parameter values can be found in Table 1.

Table 1. Parameter values of the LR model and the SZIU storage

Δt, s	60
τ, s	60
τ_A, s	600
τ_B, s	600
$c_{in,B}$, -	0.0049
$c_{in,C1}$, -	-0.000096
$c_{in,C2}$, -	0.0011
$c_{load,A}$, Ksm^{-3}	37.7445
$c_{load,B}$, Ksm^{-3}	31.0834
$c_{load,C1}$, Ksm^{-3}	13.3407
$c_{load,C2}$, Ksm^{-3}	29.4138
$c_{s,up,A}$, -	0.9999
$c_{s,up,B}$, -	0.9952
$c_{s,up,C1}$, -	1.0000
$c_{s,up,C2}$, -	0.9989

The following indices (corresponding to the currently investigated day) are used in this paper for the evaluation. The average of error is the time average of $(T_{s,up,\mathrm{mod}} - T_{s,up,meas})$, the average of absolute error is the time average of the absolute value $\left|T_{s,up,\mathrm{mod}} - T_{s,up,meas}\right|$. The average of absolute error is determined also in % dividing it by the (positive) difference between the maximal and minimal value of $T_{s,up,meas}$.

Identification

The measured data of four days have been selected for the identification in such a way that they cover a wide range of possible operating conditions of a selected season (summer). Two days (8th June, 2012; 28th June, 2012) are with relatively high consumption load (more than 1000 litres) and two ones (24th June, 2012; 2nd July, 2012) are with relatively low consumption load (less than 200 litres). Based on many computer experiments (not detailed here), such four days proved to be enough for the identified model to have a rather good accuracy. For the sake of practice, these four days have been selected from the first third of the summer. In this way, the already identified model can be conveniently used in the remained two summer months. (To apply the model for the whole year, the identification could be carried out easily for each season separately for maximal yearly precision).

Four independent standard MLR routines have been applied based on the measured data of each separate operating case (*Cases A*, *B*, *C1* and *C2*) of the LR model to identify parameters $c_{load,A}$, $c_{s,up,A}$, $c_{in,B}$, $c_{load,B}$, $c_{s,up,B}$, $c_{in,C1}$, $c_{load,C1}$, $c_{s,up,C1}$, $c_{in,C2}$, $c_{load,C2}$, $c_{s,up,C2}$ in Eqs. (1A), (1B), (1C1) and (1C2) in the LR model. The standard MLR routine (based on least squares method) is well-known and available in most statistical and spreadsheet programs (SPSS, Excel, etc.), so it is not detailed here. The identified parameters of the LR model can be seen in Table 1.

Table 2 contains the average of error and the average of absolute error values for a selected day (2nd July, 2012) of the identification (with the already identified LR model). The average of absolute error value is presented in proportion to the positive difference between the daily maximal and minimal measured storage temperature values as well (in %). The mean of these % values relating to all of the four days of the identification can be also seen in Table 2 (2.6 %).

Table 2. Average of error and average of absolute error values with the LR model

Identification	2nd July	Average of error	0.08 °C
		Average of absolute error	0.11 °C; 1.5%
	Mean % value for the whole identification (four days)	Average of absolute error	2.6%
Validation	3rd August	Average of error	0.15 °C
		Average of absolute error	0.15 °C; 2.6%
	Mean % value for the whole validation (3rd July – 31st August)	Average of absolute error	7.7%

Fig. 3 compares the modelled and measured storage temperatures of the LR model for a selected day (2nd July, 2012) of the identification. The operating state of the pump is also shown in the figure.

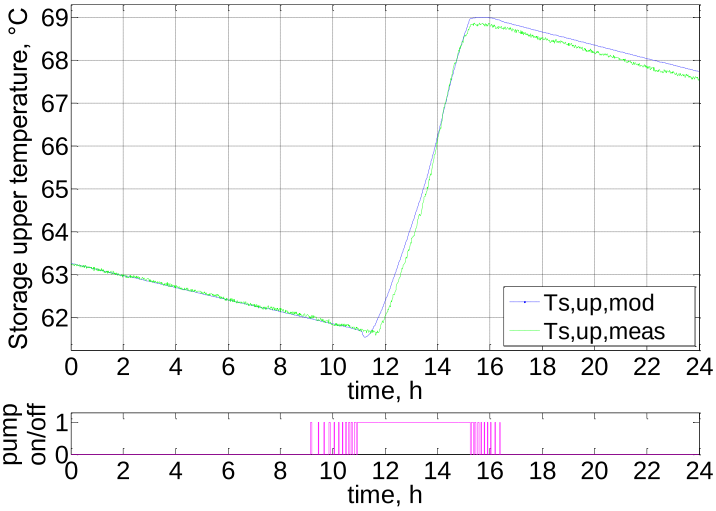

Fig. 3. Modelled $T_{s,up,\mathrm{mod}}$ and measured $T_{s,up,meas}$ upper storage temperatures on 2nd July 2012

Validation

In the validation, the LR model is applied with the corresponding measured inputs of the remaining two summer months. More precisely, one input of the model is changed in comparison with the inputs of the identification, namely, the modelled value $T_{s,up,\mathrm{mod}}(t-\tau)$ is used as $T_{s,up}(t-\tau)$ in the LR model (1A), (1B), (1C1) and (1C2) (not $T_{s,up,meas}(t-\tau)$). The modelled days are from 3rd July, 2012 to 31st August, 2012, which means 56 days for the validation according to minor technical interruptions in the operation.

The modelled and measured storage temperatures are compared and evaluated. Table 2 contains the average of error and the average of absolute error values for a selected day (3rd August, 2012) of the validation. The average of absolute error value is presented in proportion to the positive difference between the daily maximal and minimal measured storage temperature values as well (in %). The mean of these % values relating to the whole modelled time period 3rd July – 31st August is also presented in Table 2 (7.7%).

Fig. 4 compares the modelled and measured storage temperatures of the LR model for a selected day 3rd August, 2012) of the validation. The operating state of the pump is also shown in the figure.

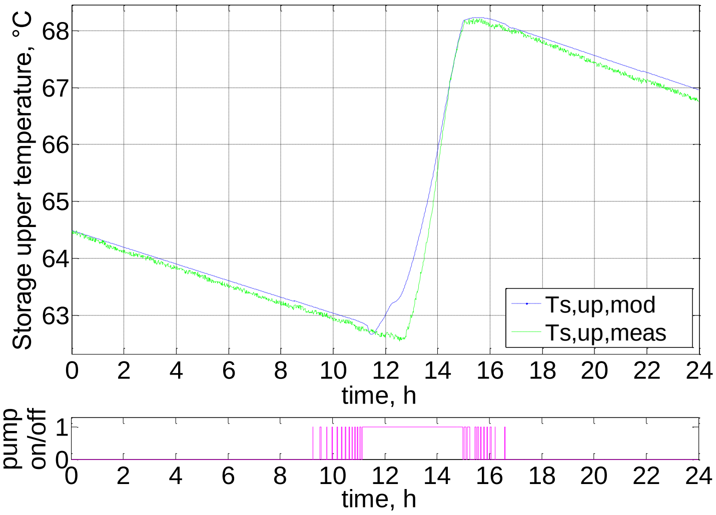

Fig. 4. Modelled $T_{s,up,\mathrm{mod}}$ and measured $T_{s,up,meas}$ upper storage temperatures on 3rd August 2012

CONCLUSION

Based on the validation, it can be stated that the temperature in the upper (undisturbed) part of a solar storage can be modelled rather precisely with the LR model (with an error of 7.7%) in view of the general solar engineering aims (studying the thermal processes and developing solar storage tanks).

Because of the simple linear algebraic relations of the LR model, the computational demand is low, which may make it ideal for model-based controls. Even, this model is likely one of the simplest black-box models still describing the transient changes of the upper storage temperature with a satisfactory precision. The advantages of the simple usability and low computational demand can be seen especially in comparison with other black-box models, e.g. the often used ANN models, which have essentially the same precision.

Future research may deal with modelling the lower part of a solar storage, which is more disturbed by the inlet flows to the storage (caused by the consumption load and the pump of the heating loop), and, accordingly, more difficult to model. Also, the already existing MLR-based models of separate working components (worked out in (Kicsiny, 2016; 2017) and the present study) may be connected in the future to form an easy-to-use and, hopefully, precise MLR-based model for complete solar heating systems.

NOMENCLATURE

t: time, s;

T_{in}: inlet temperature of the solar storage from the heating loop, °C;

T_{out}: outlet temperature of the solar storage to the heating loop, °C;

$T_{s,up}$: temperature of the upper part of the solar storage, °C;

$T_{s,up,meas}$: measured temperature of the upper part of the solar storage, °C;

$T_{s,up,\mathrm{mod}}$: modelled temperature of the upper part of the solar storage, °C;

v: volumetric flow rate of the heating loop, m^3/s;

v_{load}: volumetric flow rate of the heating loop, m^3/s;

Δt: time period between successive measurements, s;

τ: time delay with respect to the effects of the inlets to (the upper part of) the storage tank, s;

τ_A: time lag before Case A in the LR model, s;

τ_B: time lag before Case B in the LR model, s

ACKNOWLEDGEMENTS

The author thanks the Department of Physics and Process Control at the Faculty of Mechanical Engineering (SZIU) for the possibility of measuring on the SZIU storage and his colleagues at the Department of Mathematics for their contribution.

This paper was supported by the János Bolyai Research Scholarship of the Hungarian Academy of Sciences.

REFERENCES

Etter, D.M., Kuncicky, D., Moore, H. (2004), Introduction to MATLAB 7, Springer

Farkas, I., Buzás, J., Lágymányosi, A., Kalmár, I., Kaboldy, E., Nagy, L. (2000), A combined solar hot water system for the use of swimming pool and kindergarten operation, Energy and the environment, Vol. I. /ed. by B. Frankovic/, Croatian Solar Energy Association, Opatija, 2000., pp. 81-88.

Géczy-Víg, P., Farkas, I. (2010), Neural network modelling of thermal stratification in a solar DHW storage, Solar Energy, Vol. 84, pp. 801-806.

Kalogirou, S.A., Panteliou, S., Dentsoras, A. (1999), Modelling of solar domestic water heating systems using artificial neural networks, Solar Energy, Vol. 65, No. 6, pp. 335-342.

Kalogirou, S.A. (2000), Applications of artificial neural-networks for energy systems, Applied Energy, Vol. 67, pp. 17-35.

Kicsiny, R. (2016), Improved multiple linear regression based models for solar collectors, Renewable Energy, Vol. 91, pp. 224–232.

Kicsiny, R. (2017), Grey-box model for pipe temperature based on linear regression, International Journal of Heat and Mass Transfer, Vol. 107, pp. 13–20.

Meyer, J.P., Raubenheimer, P.J.A., Kruger, E. (2000), The influence of return loop flow rate on stratification in a vertical hot water storage tank connected to a heat pump water heater, Heat Transfer Engineering, Vol. 21, No. 2, pp. 67-73.