



## SIMPLIFIED MULTIPLE LINEAR REGRESSION BASED MODEL FOR SOLAR COLLECTORS

### Author(s):

R. Kicsiny

### Affiliation:

Department of Mathematics, Institute for Mathematics and Informatics, Szent István University, Páter K. u. 1., Gödöllő, H-2103, Hungary

### Email address:

kicsiny.richard@gek.szie.hu

### Abstract

A recent accurate multiple linear regression (MLR) based collector model is simplified to gain a more easy-to-apply model with still good accuracy. The new model (SMLR model) is validated and compared with the former MLR based model (MLR model) and with a physically-based model used successfully in applications. Based on measurements, the SMLR model is nearly the same accurate as the physically-based one but more easy-to-apply than the physically-based and the MLR models. The computational demand is also lower than in case of any former model. Accordingly, the SMLR model is suggested for fast but relatively accurate collector modelling.

### Keywords

solar collectors, mathematical modelling, black-box model, multiple linear regression

### 1. Introduction

Mathematical modelling is the most widely used and theoretically established tool to investigate and develop solar thermal collectors as environmentally friendly technological heat producers. The two main categories of mathematical models for collectors are physically-based models, which represent exact physical laws (based on theory), and black-box models, which describe empirical correlations (based on experiences or measurements).

Among the most important physically-based models, the Hottel-Whillier-Bliss model [1] may be the earliest, which is frequently used to date. This model determines the collector temperature as a function of time and space. Buzás et al. [2] proposed a simpler model assuming that the collector temperature is homogeneous in space. This model is a linear ordinary differential equation (ODE) validated in [3] and is likely the simplest physically-based model used in the practice (see e.g. [4, 5]), which can still describe the transient collector processes with an appropriate accuracy.

The greatest advantage of black-box models is that it is not needed to know the physical laws of a collector precisely in order to create a model. Nevertheless, the model may be rather precise even if it is simple as in the case of [3]. The most frequent black-box model type is perhaps the artificial neural network (ANN) in

the field of collector modelling. Generally, ANNs are accurate tools but rather troublesome to apply because of the so-called training process. The convergence of the algorithm, which indicates the end of a training session, may be also time-consuming. According to Fischer et al. [6], a conveniently usable algorithm ensuring a reliable and fast determination of an appropriate ANN for a collector is still needed to be worked out.

Because of the above problems, a simple and general but still accurate black-box model, which can be applied easily and fast for a wide range of solar collectors, has been recently worked out in [3]. The model is based on the well-known methods of mathematical statistics, more precisely, the multiple linear regression (MLR). Based on the literature, MLR is a rare black-box modelling technique in the field of collectors despite of its simplicity. Considering the high precision (with an error of 4.6%), simple usability and low computational demand of the mentioned MLR-based model (MLR model in short) in [3], it is definitely worth trying to simplify further the MLR model to gain an even more easy-to-apply model with a still good accuracy. Such a simplification (likely the simplest possible MLR based model) is proposed in the present study.

### 2. Physically-based and MLR models

For the Reader's convenience, the physically-based collector model of Buzás et al. [2], which will be called physically-based model in short, and the MLR model [3] are recalled in this section. The scheme of the studied solar collector can be seen in Figure 1.

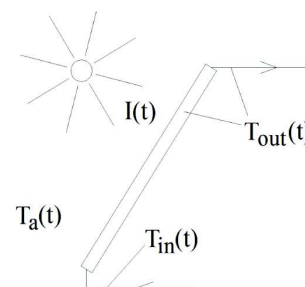


Figure 1. Scheme of the solar collector

#### *Physically-based model*

The physically-based model is the following ODE:

$$\frac{dT_{out}(t)}{dt} = \frac{A\eta_0}{\rho c V} I(t) + \frac{U_L A}{\rho c V} (T_a(t) - T_{out}(t)) + \frac{v}{V} (T_{in}(t) - T_{out}(t)). \quad (1)$$

### MLR model

The inputs of the MLR model are from appropriately chosen values of  $T_{in}$ ,  $I$ ,  $T_a$  and  $T_{out}$ . The output is from appropriately chosen values of  $T_{out}$ . The flow rate value  $v$  is a fixed positive constant or 0.

Because of the boundedness of the flow rate,  $T_{in}(t-\tau_1)$  can play a role as an input in the MLR model if  $T_{out}(t)$  is the output, where the positive constant  $\tau_1$  is a time delay. Similar considerations hold for  $I$  and  $T_a$  as well because of the bounded propagation speed of their effects, so former  $I(t-\tau_2)$  and  $T_a(t-\tau_2)$  values can play roles as inputs in forming the output  $T_{out}(t)$ . (The time delays of  $I$  and  $T_a$  are assumed to be the same ( $\tau_2$ ) for the sake of simplicity.) Naturally, appropriate former value of  $T_{out}$  also affects the value of  $T_{out}(t)$  and participates as the initial value of the MLR model at time  $(t-\tau_2)$  in essence. Considering the collector as a black-box, distinct sub-models as parts of the MLR model were identified for significantly different operating conditions. For example, the collector behaves different if the pump is on ( $v>0$ ) or off ( $v=0$ ) permanently. Even, the effect of  $T_{in}$  was neglected in permanently switched off case, since there is no flow between the collector inlet and outlet.

Considering a typical day, when the temperature increase of  $T_{out}$  is significant, three different operating conditions were distinguished according to Figure 2.

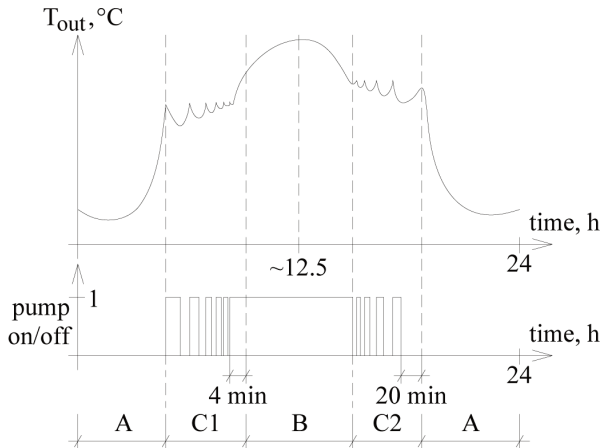


Figure 2. Outlet temperature and pump operation on a typical day

Case A corresponds to permanently switched off pump, Case B corresponds to permanently switched on pump and Case C corresponds to frequent switch-ons and -offs. It can be seen that there are two further significantly different operating cases within Case C:  $T_{out}$  basically increases before the solar noon and basically decreases after the solar noon, so Case C is divided into Cases C1 and C2. See [3] for more details.

The MLR model is composed of the following linear equations, which describe the corresponding sub-model of each operating case.

$$\text{Case A: } T_{out,mod}(t) = c_{1,A} I(t-\tau_2) + c_{2,A} T_a(t-\tau_2) + c_{3,A} T_{out}(t-\tau_2) + c_4 \quad (2a)$$

$$\text{Case B: } T_{out,mod}(t) = c_{in,B} T_{in}(t-\tau_1) + c_{I,B} I(t-\tau_2) + c_{a,B} T_a(t-\tau_2) + c_{out,B} T_{out}(t-\tau_2) + c_B \quad (2b)$$

$$\text{Case C1: } T_{out,mod}(t) = c_{in,C1} T_{in}(t-\tau_1) + c_{I,C1} I(t-\tau_2) + c_{a,C1} T_a(t-\tau_2) + c_{out,C1} T_{out}(t-\tau_2) + c_{C1} \quad (2c)$$

$$\text{Case C2: } T_{out,mod}(t) = c_{in,C2} T_{in}(t-\tau_1) + c_{I,C2} I(t-\tau_2) + c_{a,C2} T_a(t-\tau_2) + c_{out,C2} T_{out}(t-\tau_2) + c_{C2} \quad (2d)$$

$c_{1,A}$ ,  $c_{2,A}$ ,  $c_{3,A}$ ,  $c_4$ ,  $c_{in,B}$ ,  $c_{I,B}$ ,  $c_{a,B}$ ,  $c_{out,B}$ ,  $c_B$ ,  $c_{in,C1}$ ,  $c_{I,C1}$ ,  $c_{a,C1}$ ,  $c_{out,C1}$ ,  $c_{C1}$ ,  $c_{in,C2}$ ,  $c_{I,C2}$ ,  $c_{a,C2}$ ,  $c_{out,C2}$ ,  $c_{C2}$ , are constant parameters.

According to the definition of  $\tau_2$ , the measurements take place at times  $t=\tau_2, 2\tau_2, 3\tau_2, \dots$ . The modelled value of  $T_{out}$  (that is  $T_{out,mod}$ ) is determined at times  $t=\tau_2, 2\tau_2, 3\tau_2, \dots$  from the measured values of  $I(t=\tau_2)$ ,  $T_a(t=\tau_2)$ ,  $T_{out}(t=\tau_2)$  and  $T_{in}(t=\tau_1)$ , based on Equations. (2a)-(2d).

### 3. SMLR model

The MLR model is simplified in this section in the way of merging Cases A, B, C1 and C2. Thus there is only one operating case with only one mathematical relation (Equation (3) below) in the new model, which will be called SMLR model. The corresponding mathematical relation is the following:

$$T_{out,mod}(t) = c_{in} T_{in}(t-\tau_1) + c_I I(t-\tau_2) + c_a T_a(t-\tau_2) + c_{out} T_{out}(t-\tau_2) + c. \quad (1)$$

$c_{in}$ ,  $c_I$ ,  $c_a$ ,  $c_{out}$  and  $c$  are constant parameters to be identified.

Below, the SMLR model (Equation (3)) is identified and validated based on simulation and measured data then it is compared with the physically-based and MLR models in view of precision. The results and figures of the latter models used in the comparison are from. The identification and validation of the SMLR model are based on the same days as in case of the physically-based and MLR models in [3]. The used real flat plate collector field of 33.3 m<sup>2</sup> [7] at the Szent István University (SZIU) in Gödöllő, Hungary (SZIU collector in short) is also the same.  $T_{out}$ ,  $T_{in}$ ,  $I$ ,  $T_a$  and  $v$  are measured once in every minute at the SZIU collector. The measured value of  $T_{out}$  serves only for identification and comparison purposes, the measured value  $T_{out}(0)$  is fed into the models as initial condition. The technical details of the identification and validation of the SMLR model are very similar as in case of the MLR model in [3], so they are not fully specified below. The needed calculations have been done numerically in Matlab [8] used comprehensively to simulate solar engineering systems (see e.g. [9]).

#### Identification

Four measured days are selected for the identification in such a way that they cover a wide range of possible operating conditions of a selected season (summer). Since the operating conditions are well characterized with the operating states of the pump (switched on state or switched off state), two measured days (2<sup>nd</sup> July 2012, 24<sup>th</sup> June 2012) with relatively few pump switches (smooth operation) and two other days (28<sup>th</sup> June 2012, 8<sup>th</sup> June 2012) with relatively many switches (intermittent operation) are selected.

A standard MLR routine is applied based on the measured data to identify parameters  $c_{in}$ ,  $c_I$ ,  $c_a$ ,  $c_{out}$  and  $c$ . According to the minute-based measuring,  $\tau_2$  is set 1 min.  $V=0.027$  m<sup>3</sup> and  $v=0.98$  m<sup>3</sup>/h if the pump is on, so  $\tau_1 \approx 1.5$  min. The measured value of  $T_{in}$  should be used in the right hand side of Eq. (3) for identification (and for validation). Since  $\tau_1=1.5$  min is not suitable for the minute-based measuring,  $T_{in}(t-\tau_1)$  is substituted for  $(T_{in}(t-\tau_2) + T_{in}(t-2\tau_2))/2$  in the identification. The standard MLR routine (based on least squares method) is well-known, available and easy-to-apply in most statistical and spreadsheet programs (SPSS, Excel, etc.) with low computational demand, so it is not detailed here. It can be seen that the SMLR model with a single linear relation (Eq. (3)) has lower computational demand than the physically-based model with an ODE or the MLR model with four relations. The identified parameters of the SMLR model can be seen in Table 1.

Table 2 contains the average of error (time average of the difference between the modelled and measured outlet temperatures) and the average of absolute error (time average of the absolute difference between the modelled and measured outlet temperatures) values for two days (2<sup>nd</sup> July 2012, 28<sup>th</sup> June 2012) of the identification of all models. The average of absolute error values

are presented in proportion to the difference between the daily maximal and minimal measured outlet temperature values as well (in %). The mean of these % values with respect to all of the four days of the identification is also presented in Table 2 (6.6 % for the SMLR model).

Table 1. Parameter values of the SMLR model

$c_{in}$ , -	$c_f$ , m <sup>2</sup> K/W	$c_a$ , -	$c_{out}$ , -	$c$ , °C
-0.0017	0.0019	0.0471	0.9707	-0.2621

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

			Physically-based model	MLR model	SMLR model
Identification	2 <sup>nd</sup> July (smooth operation)	Average of error	-1.86 °C	-0.47 °C	1.43 °C
		Average of absolute error	4.33 °C; 7.0%	2.79 °C; 4.6%	3.88 °C; 6.3%
	28 <sup>th</sup> June (intermittent operation)	Average of error	-1.26 °C	-0.23 °C	-2.87 °C
		Average of absolute error	4.35 °C; 7.5%	3.01 °C; 5.2%	3.39 °C; 5.8%
	Mean % value for the whole identification (four days)	Average of absolute error	7.8%	4.7%	6.6%
Validation	3 <sup>rd</sup> August (smooth operation)	Average of error	-1.38 °C	-1.31 °C	-0.12 °C
		Average of absolute error	4.70 °C; 7.4%	2.85 °C; 4.5%	3.71 °C; 5.8%
	5 <sup>th</sup> August (intermittent operation)	Average of error	-2.57 °C	-1.58 °C	-1.00 °C
		Average of absolute error	4.66 °C; 8.0%	3.07 °C; 5.2%	3.95 °C; 6.7%
	Mean % value for the whole validation (3 <sup>rd</sup> July – 31 <sup>st</sup> August)	Average of absolute error	7.8%	4.6%	8.0%

Figure 3 compares the modelled and measured outlet temperatures in case of the physically-based and SMLR models

for a day of the identification. The operating state of the pump is also shown in the figure.

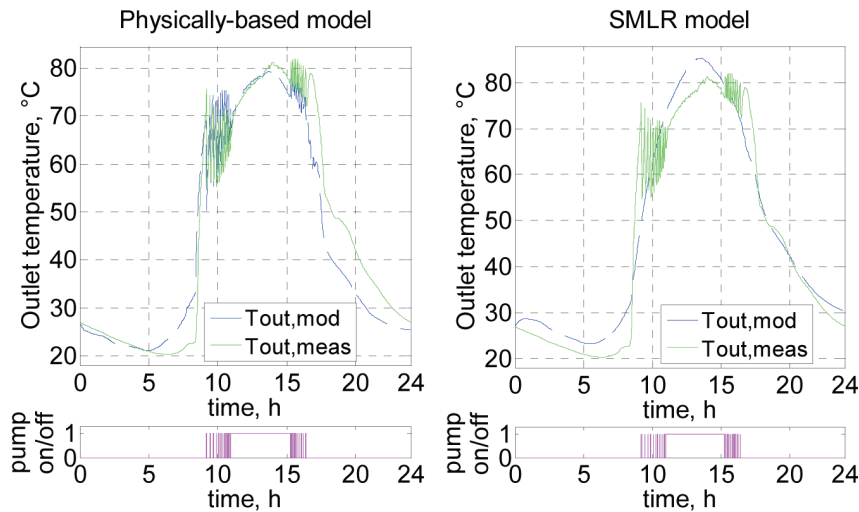


Figure 3. Modelled  $T_{out,mod}$  and measured  $T_{out,meas}$  collector temperatures on 2<sup>nd</sup> July 2012 in case of the physically-based and SMLR models

### Validation

In the validation, all identified models are applied with the corresponding measured inputs of the remaining two summer months. More precisely, one input is changed in comparison with the identification, namely, the modelled value  $T_{out,mod}(t-\tau_2)$  is used as  $T_{out}(t-\tau_2)$  in the models (2a)-(2d) and (3) (not  $T_{out,mod}(t-\tau_2)$ ), since the outlet temperature is to be modelled in the validation of course and not to be measured. The modelled days are from 3<sup>rd</sup>

July to 31<sup>st</sup> August 2012, which are 56 days according to minor technical interruptions.

Table 2 shows the resulted error values (the same as in the Identification section) of each model for two days and for the whole validation. Figure 4 shows the modelled and measured outlet temperatures in case of the physically-based and SMLR models for a day of the validation. The operating state of the pump is also shown in the figure

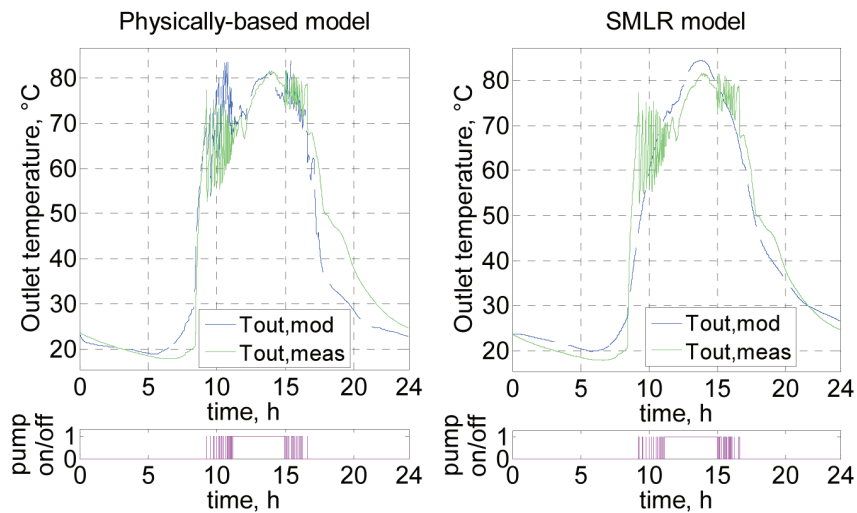


Figure 4. Modelled  $T_{out,mod}$  and measured  $T_{out,meas}$  collector temperatures on 3<sup>rd</sup> August 2012 in case of the physically-based and SMLR models

#### 4. Conclusion

Based on the validation, the SMLR model (with an error of 8.0%) is considerably less accurate than the MLR model (with an error of 4.6%) but nearly the same accurate as the physically-based model (with an error of 7.8%), which has been used successfully in the practice. The SMLR model has the lowest computational demand and is much simpler to use than the other models. The SMLR model is easy to identify for any collector with the same inputs and output as above (evacuated tube collectors, parabolic trough collectors, etc.), so the model is general. Thus the SMLR model can be suggested for fast but still relatively precise collector modelling.

It should be mentioned that the collector temperature is measured on the outside surface of the outlet pipe of the SZIU collector and not directly in the collector fluid. Thus the outlet temperature cannot be expected to be modelled perfectly because of the inaccuracy caused by this measuring and that the collector field is charged with significant disturbances (shadowing effect of clouds), which are hard to predict. Also, the rather small volume of the collector field involves high and fast changes in the collector temperature under the influence of the disturbances. These difficulties reinforce that the accuracy of the SMLR model can be called well.

#### Nomenclature

$t$	– time, s;	
$T_{out}$	– homogeneous temperature and also outlet (fluid) temperature of the collector,	°C;
$I$	– global solar irradiance on the collector surface,	W/m <sup>2</sup> ;
$T_a$	– ambient temperature of the collector,	°C;
$T_{in}$	– inlet (fluid) temperature of the collector,	°C;
$A$	– collector surface area,	m <sup>2</sup> ;
$\eta_0$	– optical efficiency of the collector,	- ;
$\rho$	– collector fluid density,	kg/m <sup>3</sup> ;
$c$	– specific heat capacity of the collector fluid,	J/(kgK);
$V$	– collector volume,	m <sup>3</sup> ;
$U_L$	– overall heat loss coefficient of the collector,	W/(m <sup>2</sup> K);
$v$	– volumetric flow rate in the collector (assumed to be constant),	m <sup>3</sup> /s;
$\tau_1$	– time of flowing from the collector inlet to the outlet in case of permanently switched on pump,	s;
$\tau_2$	– sampling time of the measurements (time between successive measurements of the measured variables),	s

#### Acknowledgements

The author thanks Prof. István Farkas and the Department of Physics and Process Control (SZIU) for the possibility of measuring on the SZIU collector and his colleagues, especially Dr. László Székely, in the Department of Mathematics in the Faculty of Mechanical Engineering (SZIU) for their contribution.

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

#### References

- [1.] Duffie J. A., Beckman W. A.: 2013. Solar engineering of thermal processes, 4th ed., John Wiley and Sons, New York.
- [2.] Buzás J., Farkas I., Biró A., Németh R.: 1998. Modelling and simulation of a solar thermal system, Mathematics and Computers in Simulation, Vol. 48, pp. 33-46.
- [3.] Kicsiny R.: 2014. Multiple linear regression based model for solar collectors, Solar Energy, Vol. 110, pp. 496-506. <http://dx.doi.org/10.1016/j.solener.2014.10.003>
- [4.] Kumar R., Rosen M. A.: 2010. Thermal performance of integrated collector storage solar water heater with corrugated absorber surface, Applied Thermal Engineering, Vol. 30, pp. 1764-1768. <http://dx.doi.org/10.1016/j.applthermaleng.2010.04.007>
- [5.] Buzás J., Kicsiny R.: 2014. Transfer functions of solar collectors for dynamical analysis and control design, Renewable Energy, Vol. 68, pp. 146-155. <http://dx.doi.org/10.1016/j.renene.2014.01.037>
- [6.] Fischer S., Frey P., Drück H.: 2012. A comparison between state-of-the-art and neural network modelling of solar collectors, Solar Energy, Vol. 86, pp. 3268-3277. <http://dx.doi.org/10.1016/j.solener.2012.09.002>
- [7.] 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.
- [8.] Etter D. M., Kuncicky D., Moore H.: 2004. Introduction to MATLAB 7, Prentice Hall, 2004, Springer.
- [9.] Buonomano A., Calise F., Palombo A.: 2013. Solar heating and cooling systems by CPVT and ET solar collectors: a novel transient simulation model, Applied Energy, Vol. 103, pp. 588-606. <http://dx.doi.org/10.1016/j.apenergy.2012.10.023>