

Grey-box model for pipe temperature based on linear regression

Richárd Kicsiny

Department of Mathematics, Institute for Mathematics and Informatics, Szent István University, Páter
K. u. 1., 2100 Gödöllő, Hungary

E-mail address: Kicsiny.Richard@gek.szie.hu

Tel.: +3628522000/1413, fax: +3628410804

Abstract

Developing mathematical models describing pipe (or duct) temperature is of great importance, since pipes are unavoidable elements in most (hydraulic) heating systems, in which some heat transfer fluid flows/circulates between neighbouring working components (such systems are district, central or solar heating systems, etc.).

In the present study, the Newton's law of cooling is completed with a recent explicit equation determining the time delay of pipes. Based on measured data, the gained mathematical model, called physically-based model, describes the outlet (fluid) temperature of pipes with a convenient accuracy with respect to the practice.

A further model, called LR model, is worked out based on multiple linear regression. Based on measured data, the LR model can model the outlet temperature of pipes generally more precisely than the physically-based model if the flow rate is nonzero. In addition, the LR model has lower computational demand.

Since the physically-based model is still more precise under certain conditions, a third model, called grey-box model, is proposed as a combination of the physically-based and the LR model calculating every time according to the more advantageous one of them. Based on measured data, the grey-box model is the most precise model. In addition, this model has lower computational demand than the physically-based model.

Keywords: Pipe temperature; Mathematical modelling; Linear regression; Grey-box model

Nomenclature

t : time, s;

x : space coordinate along the pipe, m

Time-dependent functions

T : pipe temperature, °C;

T_a : ambient temperature of the pipe, °C;

T_{in} : inlet pipe temperature, °C;

T_{out} : outlet pipe temperature, °C;

$T_{out, meas}$: measured outlet pipe temperature, °C;

$T_{out, mod}$: modelled outlet pipe temperature, °C;

v : (pump) flow rate in the pipe, m³s⁻¹

Constant parameters

A : area of pipe cross section, m²;

c : specific heat capacity of the pipe fluid, Jkg⁻¹K⁻¹;

k : heat loss coefficient of the pipe, Wm⁻¹K⁻¹;

L : length of pipe, m;

V : pipe volume, m³;

Δt : time period between successive measurements on the pipe, s;

47 ρ : mass density of the pipe fluid, kgm^{-3}

48 **1. Introduction**

49 Developing mathematical models describing pipe (or duct) temperature is of great
50 importance, since pipes are unavoidable elements in most (hydraulic) heating systems, in
51 which some heat transfer fluid flows/circulates between neighbouring working components.
52 District heating systems, central heating systems and solar heating systems can be mentioned
53 here as particular examples.

54 In many works, the heat loss effect, and even more often, the time delaying effect, of pipes is
55 neglected. Nevertheless, one or both of the above pipe effects must be often considered, for
56 example, if the pipe is not so well insulated and/or the pipe is relatively long. This is the case
57 for district heating systems [1, 2] or e.g. in [3], where 5-10% more solar energy can be gained
58 at a real solar heating system with differential control if the pipes are taken into account
59 precisely enough.

60 As for most other working components, differential equations (DEs) are the most frequent
61 mathematical models to model pipes. See, for example, [4], when an inverse method is used
62 to estimate the inlet temperature on the basis of the governing DEs. In the simplest case, a
63 pipe is modelled with a single ordinary differential equation (ODE) assuming homogeneous
64 pipe (fluid) temperature (with respect to space) as the single state variable. The heat loss can
65 be considered in such models [5], so they can be accurate enough on an average in long terms,
66 but, if the transients are important, more sophisticated models should be used.

67 The most often model, describing both delay and heat loss, is the linear one dimensional
68 partial differential equation (PDE) corresponding to heat transfer and plug-flow (that is, the
69 mixing and temperature homogenization effects inside the pipe are neglected). In other words,
70 this model is the one dimensional linear transport equation [6] describing the pipe (fluid)
71 temperature supported with a member with respect to heat loss, optionally. This PDE is
72 applied to model temperature distribution in solar collectors in [7 and 8], in heat exchangers
73 with connecting pipes in case of neglected heat losses in [9], in chemical tubular reactor in
74 [10] and in a district heating system in [2]. In the latter difference, the PDE is used to
75 determine time delay directly as well. In [11], the PDE describing the temperature distribution
76 inside the pipe is applied to a water heating equipment of pilot-scale in case of perfectly
77 insulated pipes. By means of a simplifying procedure, the PDE is transformed into two
78 models for control purposes. One of them contains ODEs for perfectly mixed sections, while
79 the other one is a length integrated model, determining the average of the temperature along
80 the pipe. Similarly, PDEs describing temperature distribution are used for pipes and parallel-
81 plate channels in [12] and for a basic natural circulation loop in [13]. If the one dimensional
82 linear heat transfer equation corresponding to pipe temperature is applied to the moving “fluid
83 element” inside the pipe, essentially, the well-known Newton’s law of cooling [14] is gained.

84 Although, PDEs are usually more difficult to handle than ODEs and they cannot be solved
85 exactly, solutions with desired precision can be generally gained by means of discretization
86 methods. For example, in the TRNSYS software [15], which is widely used to simulate
87 transient thermal processes in different heating systems, a pipe is divided into segments, each
88 of which has homogeneous temperature and is modelled with an ODE. The pipes of district
89 heating systems are discretized in [1], after which, the temperature is calculated numerically.

90 Although, PDEs can be generally solved numerically with desired precision, there is a
91 problem of principle with respect to the (one dimensional) linear transport equation
92 corresponding to pipe (fluid) temperature. Namely, if the (pump) flow rate is intermittent, that
93 is, the flow rate is sometimes zero, the transport equation may have no classical solution or
94 may have not unique solutions. Neither case satisfies the natural physical expectation on
95 definiteness (see Remark 2.1 in [16] for more details). This problem is avoided with nonzero
96 flow rate (e.g. in [7 and 17]) or with numerical solution (e.g. in [10]).

The problem of discontinuity can be overcome if heating systems affected with the delaying effect of pipes are modelled with delay differential equations (DDEs) [18, 19]. Most references on thermal engineering problems deal with constant time delay. See, for example, [20] and [21] on DDEs with respect to heat conduction if the heat flux vector is endowed with time delay. In [22], the pipe outlet temperature is described in time by means of a delayed equation derived from the heat transfer PDE in case of constant flow rate. An explicit formula is given to express the time delay as a function of time in case of variable flow rate in [17], although, it can be used only if the flow rate is nonzero. The concept is improved in [16], where an explicit delay equation is proposed, which can be applied to intermittent flow rates as well.

So far, white-box models were discussed as they are founded on known physical phenomena. In case of black-box models, some experienced/measured correlations are represented empirically. In the present work, multiple linear regression (MLR) is used in constructing a black-box model to determine the outlet temperature of pipes as a function of the inlet and ambient temperatures. See [23], where solar collectors, as other working components of heating systems, are modelled by means of MLR.

The contributions are the following in details in the present study.

1. In the determination of the outlet (fluid) temperature, the Newton's law of cooling is used to model the heat loss to the ambience of the pipe completed with the explicit delay equation of [16] to determine the time delay. This white-box model is called physically-based model henceforth.
2. Furthermore, an MLR based model, called LR model, is worked out to determine the outlet temperature of pipes as a simple linear function of the inlet and ambient temperatures. It is presented based on measured data that the LR model is generally more precise than the physically-based one if the flow rate is nonzero. In addition, the LR model has lower computational demand.
3. Since the physically-based model is still more precise under certain conditions, a third model, called grey-box model, is proposed as a combination of the physically-based and the LR model calculating every time according to the more advantageous one of them. Based on measured data, the grey-box model is more precise than any of the other two models. In addition, the grey-box model has lower computational demand than the physically-based model.

The Matlab software [24] has been applied in this work to carry out the needed calculations. The organization of the paper is the following: Section 2 gives common features on the pipe operation, the measurements and the modelling corresponding to all of the studied models. In Section 3 and 4, the physically-based and the LR model are constructed and their identification and validation are given based on measured data. The grey-box model is proposed and validated in Section 5. The detailed comparison of the models is given in Section 6. Conclusions and recommendations for future research works are given in Section 7.

2. Common features on pipe operation, measurements and modelling

2.1. Pipe operation

Fig. 1 shows the pipe to be modelled.

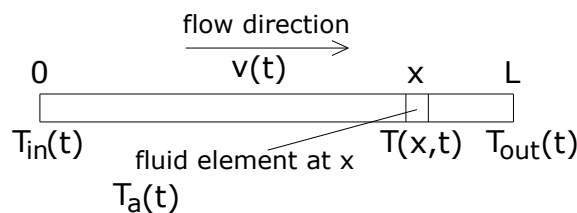


Fig. 1. Temperature along the pipe at time t

Obviously, it takes a certain time for the “fluid element” leaving the pipe inlet at time τ to reach the outlet at time t . That is, the delay d equals to $t - \tau$. If the (pump) flow rate in the pipe v is not constant but a function of time $v(t)$, then $\tau(t)$ also depends on time t , as well as the delay $d(t)$, see Eq. (1).

$$d(t) = t - \tau \quad (1)$$

In particular, $\tau(t) < t$ can be determined according to Eq. (3) if the initial contents of the pipe have already been completely discharged, that is, if Ineq. (2) holds.

$$\int_0^t v(\hat{t}) d\hat{t} \geq V, \quad (2)$$

$$\tau(t) = \max \left\{ \tilde{t} \left| \int_{\tilde{t}}^t v(\hat{t}) d\hat{t} = V \right. \right\}, \quad (3)$$

If Ineq. (2) does not hold for a time t , then $\tau(t)$ is meaningless and not defined for that t . (See [16] for more details on the above concept.)

As regards the practice, $\tau(t)$ in Eq. (3) can be easily estimated with numerical integration if $v(t)$ is measured.

In the present study, two cases of operation are distinguished on each day as follows.

Case On: $v(t) > 0$ and Ineq. (2) holds. In other words, this case includes those points of time, when the fluid is flowing and the initial contents of the pipe have already been completely discharged.

Case Off: $v(t) > 0$ and Ineq. (2) does not hold or $v(t) = 0$. In other words, this case includes those points of time, when the fluid is not flowing or it is flowing but the inlet has no effect on the outlet, since the initial contents of the pipe have not been completely discharged yet. This case is further divided into sub-cases, which are the maximal connected time intervals included completely in *Case Off*, that is, all sufficiently small distinct time intervals before or after such a sub-case has no common point with *Case Off*.

It should be noted that the mixing and temperature homogenization effects inside the pipe are neglected in the present study.

2.2. Measurements and modelling

The measured data of a real solar heating system (called *SZIU system*) installed at the Szent István University in Gödöllő (Hungary) is used in the present work (see e.g. [25, 26] for the description of the *SZIU system*). Among other system components, the pipe from the heat exchanger to the solar collector (it means only one direction) is monitored, that is, T_a , T_{in} , v and T_{out} (for comparison) are measured once a minute, so $\Delta t = 1$ min. The pipe is in the open air, its constant parameters [5] can be seen in Table 1. It should be mentioned that the pipe length is 80 m, because of that it takes about 7 min for the pump to totally discharge it, so the delaying (and heat loss) effect of the pipe is considerable. In most cases, the pipe effects are even more significant, since the pump is generally not switched on permanently but switched on/off with variable frequency. It should be also mentioned, that *Case On* and the switched on state of the pump coincide after discharging the initial pipe contents (cf. the description of *Case On* in Section 2.1) and that the pump works in differential (on/off) mode, that is, the flow rate is 0 or $0.000272 \text{ m}^3\text{s}^{-1}$ ($= 0.98 \text{ m}^3\text{h}^{-1}$) according to measurements in case of permanent pumping.

The measured data of four days (2 July, 24 June, 28 June and 8 June, 2012) are used in the identification. In 2 July and 24 June, the pump switches on/off relatively rarely (smooth

operation), in 28 June and 8 June, the pump switches relatively often (intermittent operation). The measured data of 56 days are available from 3 July to 31 August, 2012 and used in the validation.

The modelled (calculated) outlet temperature is determined for each measured time, so the modelled and measured values of the outlet temperature ($T_{out,mod}$ and $T_{out,meas}$) can be directly compared. The following indices of comparison, with respect to the currently investigated time period, are used in the present work: the average of error is the time average of $(T_{out,mod} - T_{out,meas})$, the average of absolute error is the time average of the absolute value $|T_{out,mod} - T_{out,meas}|$. The latter is considered also in proportion to the (positive) difference between the maximal and minimal value of $T_{out,meas}$, expressed in %.

3. Physically-based model

The physically-based model is the well-trying Newton's law of cooling [14] according to Eq. (4).

$$\frac{dT}{dt} = \frac{k(T_a - T)}{c\rho A} \quad (4)$$

k is the parameter to be identified in this model. In the two different operating cases, Eq. (4) is used as follows.

Case On: T in Eq. (4) is the temperature of the moving fluid element leaving the inlet at time $\tau(t)$ and reaching the outlet at time t , so the initial temperature is $T_{in}(\tau(t))$.

Case Off: T in Eq. (4) is the temperature of the stagnating fluid element at the pipe outlet.

3.1. Identification

In the identification process, Eq. (4) is solved (numerically) then the modelled outlet temperature for each measured time (gained from the solution of Eq. (4)) and the measured outlet temperature can be compared. Of course, the measured values of T_a , T_{in} and v are used in the solution. More details are given below for the two operating cases.

Case On: In practice, $\tau(t)$ is the last measured time, for which Ineq. (2) holds and each measured time belonging to this operating case is considered as time t .

Case Off: Eq. (4) is solved for each sub-case in *Case Off* separately (see Section 2.1 for the definition of the sub-cases). The initial temperature is $T_{out,meas}$ at the beginning (at the first measured time) of the current sub-case, otherwise, it is $T_{out,mod}(t - \Delta t)$. Each measured time belonging to this operating case is considered as time t .

If Eq. (4) is solved for a whole day (not only for *Case On* or *Case Off* separately), then the initial temperature in the sub-cases of *Case Off* is always $T_{out,mod}(t - \Delta t)$, except the beginning of the day, when the initial temperature is, naturally, the measured outlet temperature.

k is to be identified separately for *Cases On* and *Off* in such a way that the mean % value of the four average of absolute error values is minimal for the periods of *Case On* in the four identified days and for the periods of *Case Off* in the four identified days, respectively. The accordingly identified values of k (0.50 and 1.30 Wm⁻¹K⁻¹) can be seen in Table 1 for each operating case. Henceforth, the identified model is used to model the outlet temperature.

Comparing all measured and modelled values of the outlet temperature in the whole identification period, the R^2 (square of correlation coefficient) value is 0.9547 and 0.9802 in *Case On* and *Case Off*, respectively (see Table 2). The average of absolute error is 5.2% and 3.3% in *Case On* and *Case Off*, respectively, regarding the whole identification (see Table 3).

For two single days of the identification (2 July and 28 June), the average of error and average of absolute error values with respect to *Case On* and *Case Off* can be found in Table 3. The identified Eq. (4) has been solved also for whole days. The corresponding values for 2 July, 28 June and for the whole identification can be found in Table 3 (3.9%, 4.1% and 4.2% with respect to the average of absolute error). Fig. 2 shows the comparison graphically with respect to time for the mentioned two days. For the sake of more convenient comparison, the switching state of the pump (on/off) can be also seen in the figure.

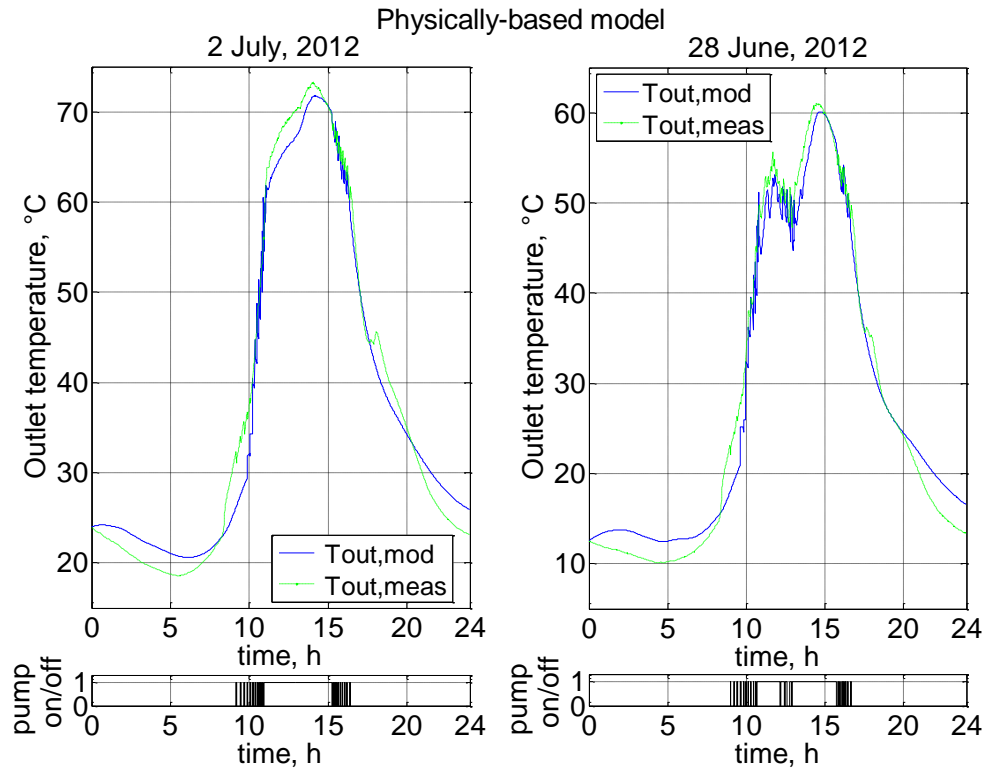


Fig. 2. Measured and modelled outlet temperature in case of the physically-based model on two days of the identification

3.2. Validation

In the validation, the modelling process is the same as in the identification in Section 3.1 (including e.g. the selection of the initial conditions in *Case On*, *Case Off* and for whole days). The above identified values of k are used here as well.

Comparing all measured and modelled values of the outlet temperature in the whole validation period, the average of absolute error is 7.6% and 2.9% in *Case On* and *Case Off*, respectively. For whole days (cf. Section 3.1), this value is 4.3% (see also Table 3).

For two single days of the validation, 3 August (smooth operation) and 5 August (intermittent operation), the average of absolute error values with respect to whole days are 4.0% and 4.7%, respectively (see Table 3 for more details). Fig. 3 shows the comparison graphically, along with the switching state of the pump, with respect to time for the mentioned days.

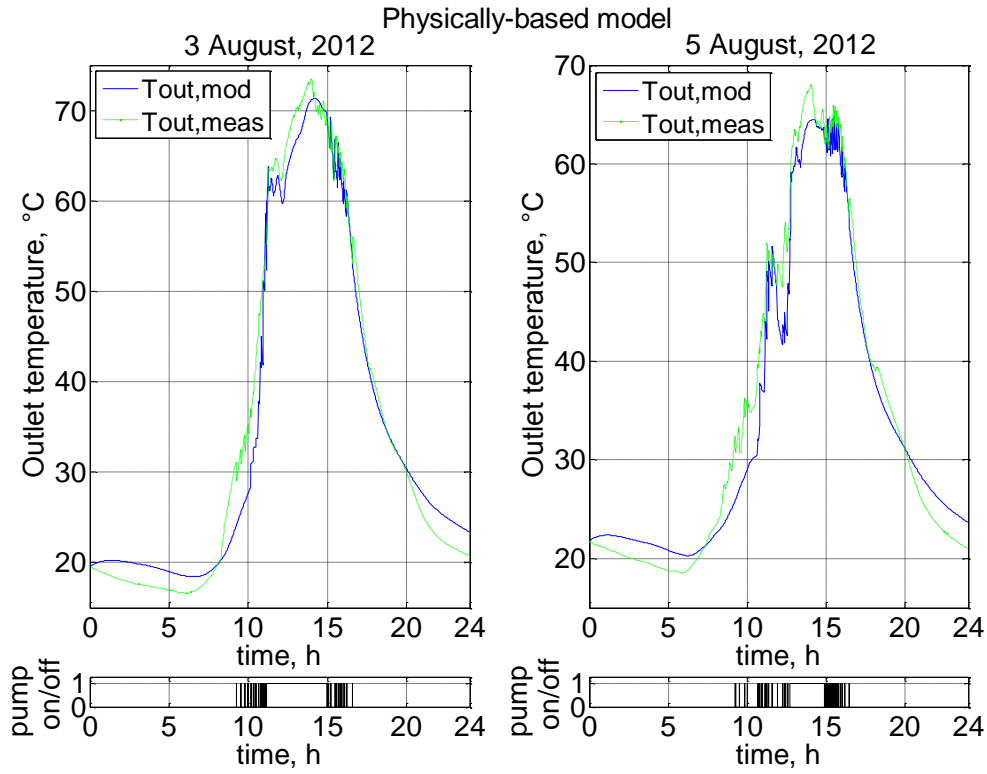


Fig. 3. Measured and modelled outlet temperature in case of the physically-based model on two days of the validation

4. LR model

The LR model contains different relations for the two operating cases. Eqs. (5a) and (5b) correspond to *Case On* and *Case Off*, respectively.

$$\text{Case On: } T_{out,mod}(t) = c_{in}T_{in}(\tau(t)) + c_d d(t) + c_{a,d}T_a(\tau(t)) \quad (5a)$$

$$\text{Case Off: } T_{out,mod}(t) = c_{a,\Delta}T_a(t - \Delta t) + c_{out}T_{out}(t - \Delta t) \quad (5b)$$

In this model, the parameters to be identified are c_{in} , c_d , $c_{a,d}$, $c_{a,\Delta}$ and c_{out} .

4.1. Identification

Case On: For a measured time t belonging to *Case On*, $\tau(t)$ is determined in the same way as in Section 3.1 then $d(t)$ can be gained from Eq. (1). If these values are known, along with the measured $T_{in}(\tau(t))$, $T_a(\tau(t))$ and $T_{out}(t)$ values, then c_{in} , c_d , $c_{a,d}$ can be easily and fast identified by means of a standard MLR routine, which can be found in any spreadsheet handling or statistical programs (in Excel, SPSS, etc.).

Case Off: For each measured time t belonging to *Case Off* (except the first one at the beginning of the current day), the measured values of $T_a(t - \Delta t)$, $T_{out}(t - \Delta t)$ and $T_{out}(t)$ are available, based on which $c_{a,\Delta}$ and c_{out} can be identified by means of a standard MLR routine. For the sake of simplicity, $T_{out,meas}$ of the first measured time is used also as the measured value of $T_{out}(t - \Delta t)$ corresponding to the first measured time t .

The accordingly identified values of c_{in} , c_d , $c_{a,d}$, $c_{a,\Delta}$ and c_{out} can be seen in Table 1 for each operating case. Henceforth, the identified model is used to model the outlet temperature. Comparing all measured and modelled values of the outlet temperature in the whole identification period, the R^2 value is 0.9615 and 0.9685 in *Case On* and *Case Off*, respectively

(see also Table 2). The average of absolute error is 4.1% and 5.3% in *Case On* and *Case Off*, respectively, regarding the whole identification (see also Table 3). For two single days of the identification (2 July and 28 June), the average of error and average of absolute error values with respect to *Case On* and *Case Off* can be found in Table 3. Regarding whole days, the corresponding values for 2 July, 28 June and for the whole identification can be found in Table 3 (6.3%, 6.1% and 5.4% with respect to the average of absolute error). Fig. 4 shows the comparison graphically, along with the switching state of the pump, with respect to time for the mentioned days.

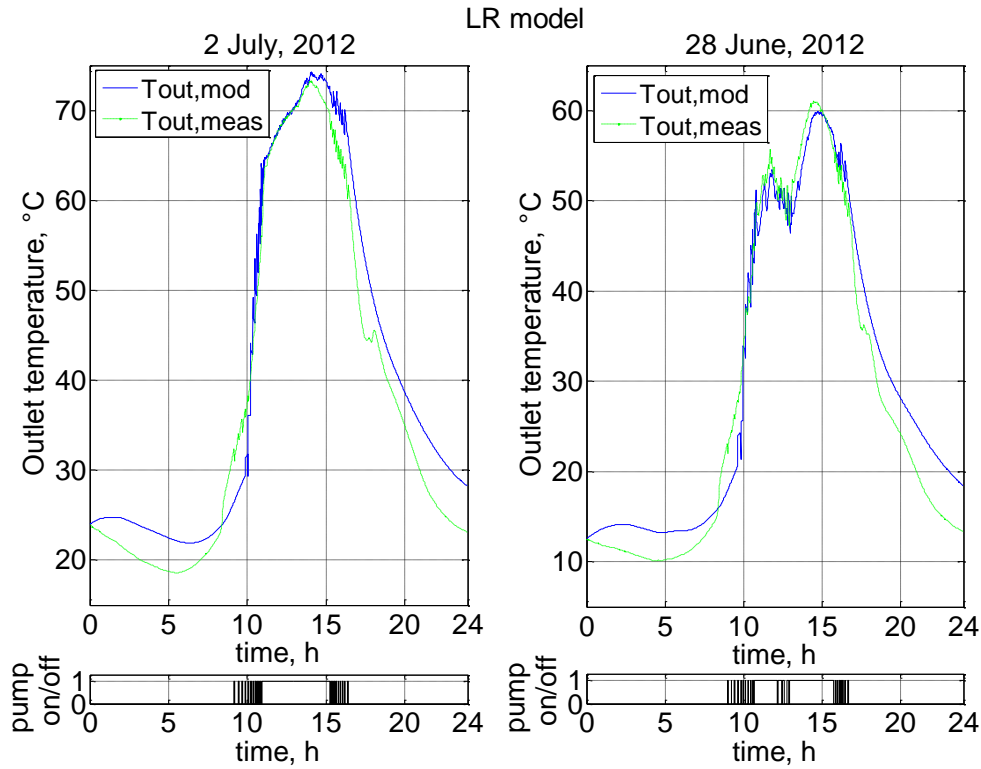


Fig. 4. Measured and modelled outlet temperature in case of the LR model on two days of the identification

4.2. Validation

The already identified LR model is used in the validation.

Case On: $\tau(t)$ and $d(t)$ is determined in the same way as in Section 4.1, furthermore, $T_{in}(\tau(t))$ and $T_a(\tau(t))$ are measured. From these data, $T_{out,mod}(t)$ can be easily and fast calculated by means of the simple linear algebraic relation Eq. (5a).

Case Off: For each measured time t belonging to *Case Off* (except the first one at the beginning of the current day), the measured value of $T_a(t - \Delta t)$ is available and $T_{out,mod}(t - \Delta t)$ (determined one step earlier) is used as $T_{out}(t - \Delta t)$ in Eq. (5b). Then $T_{out,mod}(t)$ can be simply calculated from Eq. (5b). At the first measured time, $T_{out,meas}$ is considered as $T_{out,mod}(t)$ (initial condition).

Comparing all measured and modelled values of the outlet temperature in the whole validation period, the average of absolute error is 6.4% and 4.5% in *Case On* and *Case Off*, respectively. For whole days, this value is 5.1% (see also Table 3).

For two single days of the validation, 3 August and 5 August, the average of absolute error values with respect to whole days are 6.6% and 7.1%, respectively (see Table 3 for more details). Fig. 5 shows the comparison graphically, along with the switching state of the pump, with respect to time for the mentioned days.

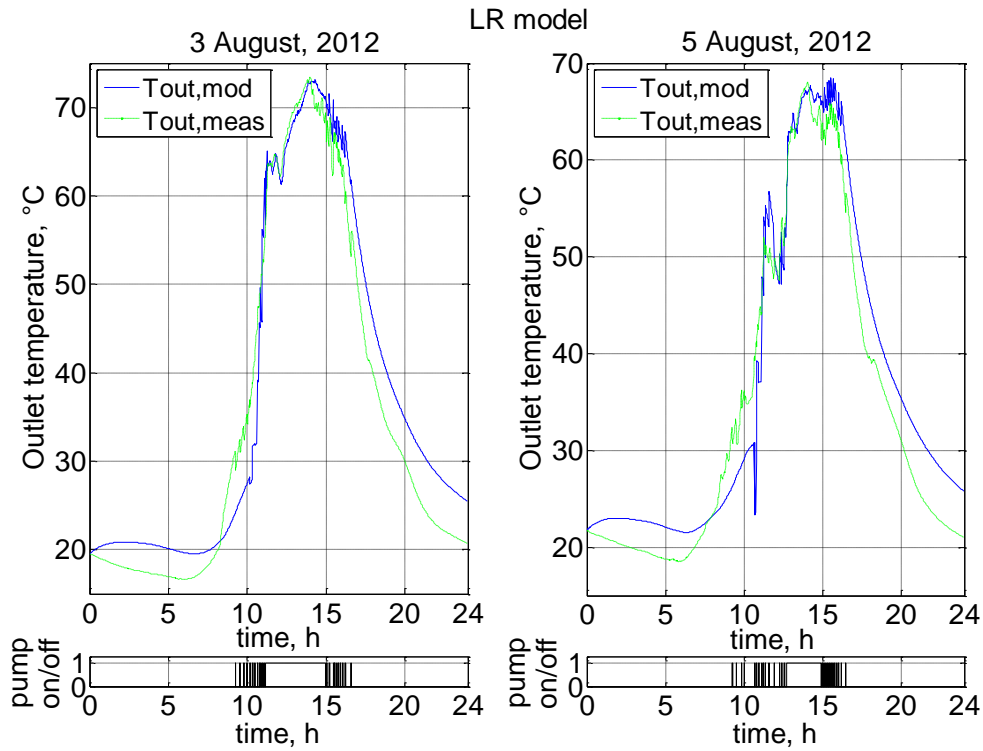


Fig. 5. Measured and modelled outlet temperature in case of the LR model on two days of the validation

5. Grey-box model

It can be seen from Table 3 that the physically-based model is more precise in *Case Off*, while the LR model is more precise in *Case On*. Namely, the mean % value of the average of absolute error for the whole validation in *Case Off* is smaller (2.9%) with the physically-based model than with the LR model (4.5%), while in *Case On*, this value is smaller with the LR model (6.4%) than with the physically-based model (7.6%). This relation between the models is reinforced by the R^2 values as this value is bigger with the physically-based model (0.9802) than with the LR model (0.9685) in *Case Off* and bigger with the LR model (0.9615) than with the physically-based model (0.9547) in *Case On*.

These results inspire to introduce a grey-box model as a composition of a white-box model (the physically-based model) and a black-box model (the LR model), which utilizes the advantages of both ones. The identified grey-box model consists of the above identified models, more precisely, it calculates $T_{out,mod}$ according to the identified physically-based model in *Case Off* and according to the identified LR model in *Case On*. Further details are the following.

Case On: $\tau(t)$ and $d(t)$ is determined in the same way as in Section 4.1, furthermore, $T_{in}(\tau(t))$ and $T_a(\tau(t))$ are measured. From these data, $T_{out,mod}(t)$ can be easily and fast calculated by means of the simple linear algebraic relation Eq. (5a).

Case Off: Eq. (4) is solved for each measured time t in *Case Off*. The initial temperature is $T_{out,meas}$ at the first measured time (at the beginning of the day), otherwise, the initial time is $(t - \Delta t)$ and the initial temperature is $T_{out,mod}(t - \Delta t)$.

The grey-box model is applied only for whole days (it would not be interesting/new for the separate operating cases). For two single days of the identification, 2 July and 28 June, and for the whole identification period, the average of absolute error values are 3.9%, 4.1% and 3.8%, respectively. Fig. 6 shows the comparison graphically, along with the switching state of the pump, with respect to time for the mentioned two days.

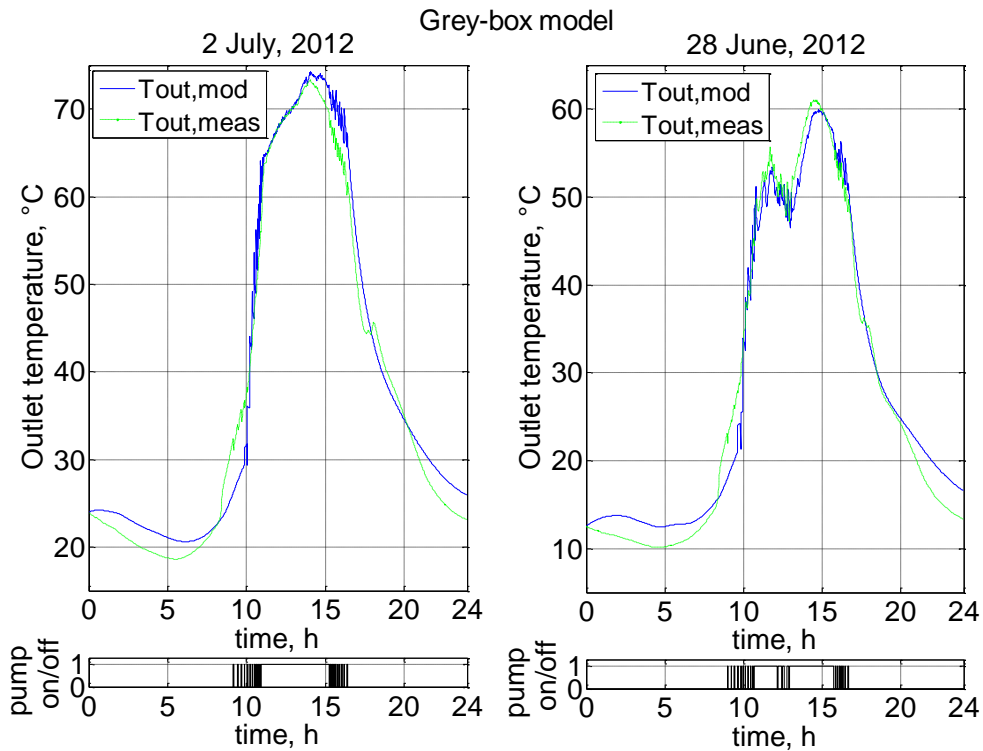


Fig. 6. Measured and modelled outlet temperature in case of the grey-box model on two days of the identification

For two single days of the validation, 3 August and 5 August, and for the whole validation period, the average of absolute error values are 4.2%, 4.5% and 3.6%, respectively. Fig. 7 shows the comparison graphically, along with the switching state of the pump, with respect to time for the mentioned two days.

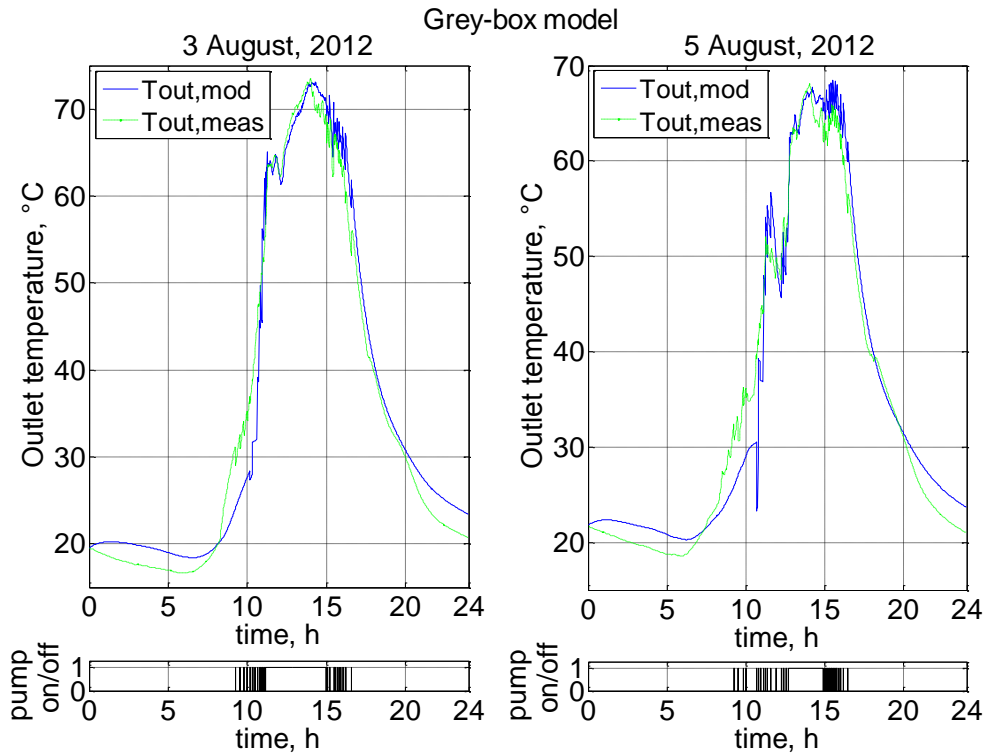


Fig. 7. Measured and modelled outlet temperature in case of the grey-box model on two days of the validation

6. Comparison

The parameter values of the three identified models are summarized in Table 1.

Table 1. Parameter values of the models

	Physically-based model		LR model		Grey-box model	
	<i>Case On</i>	<i>Case Off</i>	<i>Case On</i>	<i>Case Off</i>	<i>Case On</i>	<i>Case Off</i>
A, m^2	0.0014	0.0014	-	-	-	0.0014
$c, \text{Jkg}^{-1}\text{K}^{-1}$	3623	3623	-	-	-	3623
$c_{a,d}, -$ (identified)	-	-	0.5702	-	0.5702	-
$c_{a,\Delta}, -$ (identified)	-	-	-	0.0110	-	-
$c_d, ^\circ\text{Cs}^{-1}$ (identified)	-	-	-0.0031	-	-0.0031	-
$c_{in}, -$ (identified)	-	-	0.6997	-	0.6997	-
$c_{out}, -$ (identified)	-	-	-	0.9896	-	-
$k, \text{Wm}^{-1}\text{K}^{-1}$ (identified)	0.50	1.30	-	-	-	1.30
V, m^3	0.111	-	0.111	-	0.111	-
$\Delta t, \text{s}$	60	60	60	60	60	60
ρ, kgm^{-3}	1034	1034	-	-	-	1034

Table 2 contains the R^2 values between the measured outlet temperature and the outlet temperature calculated by the identified physically-based and the identified LR model for the whole time period of the identification.

Table 2. R^2 values of the physically-based and the LR model for the whole identification

	Physically-based model	LR model
<i>Case On</i>	0.9547	0.9615
<i>Case Off</i>	0.9802	0.9685

Table 3 summarizes the average of error and average of absolute error values produced by the three identified models for two particular days both from the identification and from the validation and for the whole identification and validation periods.

Table 3. Average of error and average of absolute error values with the models

				Physically-based model	LR model	Grey-box model
Identification	2 July, 2012 (smooth operation)	<i>Case On</i>	Average of error	0.17 °C	1.21 °C	-
			Average of absolute error	1.42 °C, 3.8%	1.34 °C, 3.6%	-
		<i>Case Off</i>	Average of error	0.68 °C	2.54 °C	-
			Average of absolute error	1.81 °C, 3.6%	3.06 °C, 6.0%	-
		Whole day	Average of error	-0.12 °C	2.70 °C	1.11 °C
			Average of absolute error	2.11 °C,	3.43 °C,	2.15 °C,

Validation				3.9%	6.3%	3.9%
	28 June, 2012 (intermittent operation)	<i>Case On</i>	Average of error	-0.23 °C	-1.37 °C	-
			Average of absolute error	1.77 °C, 5.3%	1.88 °C, 5.6%	-
		<i>Case Off</i>	Average of error	1.23 °C	2.59 °C	-
			Average of absolute error	1.82 °C, 4.0%	3.00 °C, 6.6%	-
		Whole day	Average of error	0.21 °C	1.77 °C	0.67 °C
			Average of absolute error	2.09 °C, 4.1%	3.11 °C, 6.1%	2.07 °C, 4.1%
	Mean % value for the whole identification (four days)	<i>Case On</i>	Average of absolute error	5.2%	4.1%	-
		<i>Case Off</i>	Average of absolute error	3.3%	5.3%	-
		Whole days	Average of absolute error	4.2%	5.4%	3.8%
	3 August, 2012 (smooth operation)	<i>Case On</i>	Average of error	0.28 °C	0.26 °C	-
			Average of absolute error	1.85 °C, 4.9%	1.60 °C, 4.3%	-
		<i>Case Off</i>	Average of error	0.74 °C	2.33 °C	-
			Average of absolute error	1.74 °C, 3.3%	3.11 °C, 6.0%	-
		Whole day	Average of error	-0.22 °C	2.19 °C	0.78 °C
			Average of absolute error	2.28 °C, 4.0%	3.76 °C, 6.6%	2.37 °C, 4.2%
	5 August, 2012 (intermittent operation)	<i>Case On</i>	Average of error	-0.24 °C	0.82 °C	-
			Average of absolute error	1.80 °C, 6.3%	1.80 °C, 6.2%	-
		<i>Case Off</i>	Average of error	0.43 °C	2.02 °C	-
			Average of absolute error	1.62 °C, 3.4%	2.74 °C, 5.8%	-
		Whole day	Average of error	-0.58 °C	2.13 °C	0.63 °C
			Average of absolute error	2.33 °C, 4.7%	3.50 °C, 7.1%	2.24 °C, 4.5%
	Mean % value for the whole validation (3 July – 31 August)	<i>Case On</i>	Average of absolute error	7.6%	<u>6.4%</u>	-
		<i>Case Off</i>	Average of absolute error	<u>2.9%</u>	4.5%	-
		Whole days	Average of absolute error	4.3%	5.1%	<u>3.6%</u>

Figs. 8 and 9 compare the physically-based and grey-box model graphically, along with the switching state of the pump, with respect to time for two days of the validation.

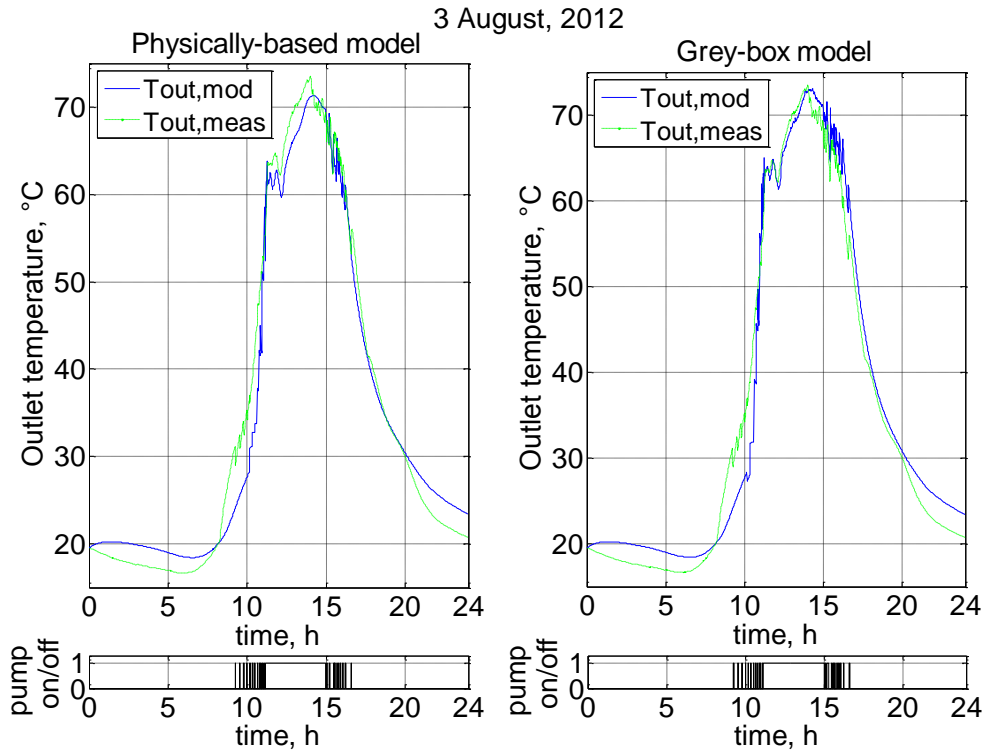


Fig. 8. Measured and modelled outlet temperature in case of the physically-based and the grey-box model on 3rd August, 2012 (validation)

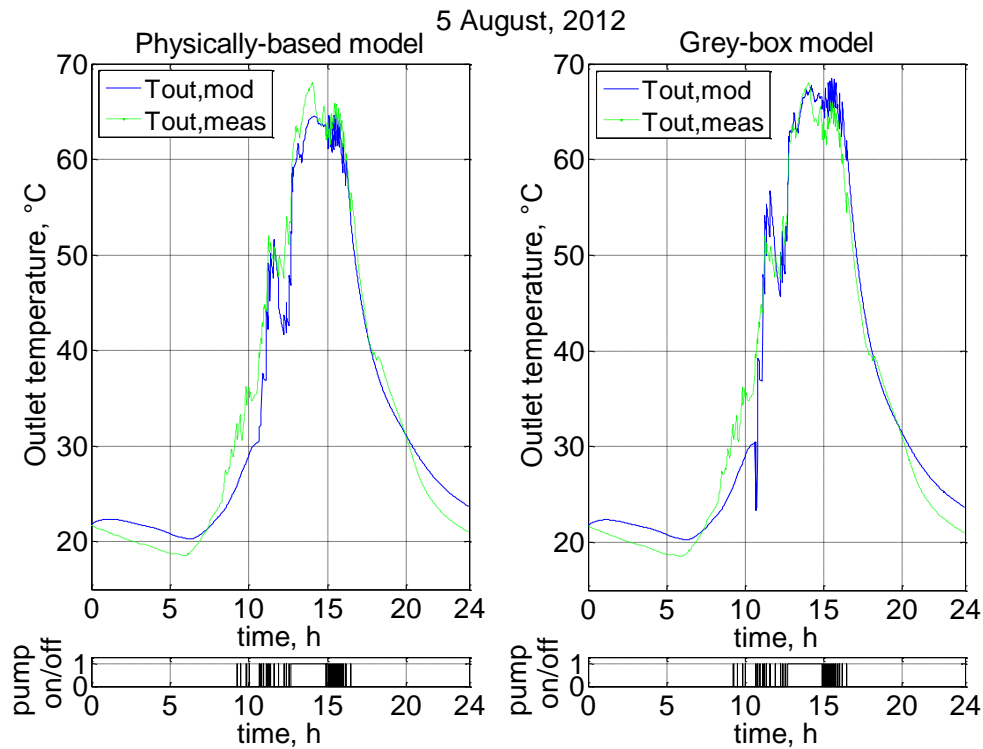


Fig. 9. Measured and modelled outlet temperature in case of the physically-based and the grey-box model on 5th August, 2012 (validation)

The last three rows of Table 3 corresponding to the mean % value of the average of absolute error values with respect to the whole validation period (from 3 July to 31 August, 2012) show the essence of the results. It can be seen that the physically-based model is more precise (the modelling error, that is, the average of absolute error is smaller) in *Case Off*, than the LR model. The related values are 2.9% and 4.5%, respectively. The relation is opposite in *Case On* with 7.6% and 6.4%, respectively. It can be hoped that the grey-box model, which applies

the physically-based model in *Case Off* and the LR model in *Case On* is better, that is, more precise for whole days than any of the other two models. This is really the case as it can be seen in the last row of Table 3 that the grey-box model has the smallest modelling error, namely 3.6%, among the three identified models, with respect to whole days.

Remark

1. In *Case Off*, the physically-based model calculates the outlet temperature for all sub-cases according to ODE (4). Since the measured data are collected once a minute, the ODE is solved with constant (measured) input T_a for every minute. In this (constant) case, the ODE can be solved analytically. The explicit solution is an exponential-type function. The n -th order Taylor polynomial used generally to estimate the exponential function (at a time t) is presented in Eq. (6).

$$1 + t + \frac{t^2}{2} + \frac{t^3}{3!} + \dots + \frac{t^n}{n!} \approx e^t \quad (6)$$

Based on Eq. (6), a computer needs to carry out basically $1+2+\dots+n-1=(n-1)n/2$ multiplications of t , $n-1$ scalar multiplications (e.g. t^3 has to be multiplied by $1/(3!)$) and $n-2$ additions every minute to calculate the outlet temperature. The LR model uses Eq. (5b) in *Case Off*, which means only 2 scalar multiplications and 1 addition each minute.

In *Case On*, the physically-based model (Eq. (4)) calculates the outlet temperature for every minute between the current value of time $\tau(t)$ and t , so the computational demand for any time t in *Case On* is similar as in any sub-case of *Case Off*. In *Case On*, the LR model (Eq. (5a)) needs to carry out only one calculation, that is, 3 scalar multiplications and 2 additions, for the current time t (and not more calculations for every minute from time $\tau(t)$ to t).

The above evaluation shows that the computational demand is considerably lower with the LR model than with the physically-based model in *Case Off* and even lower with the LR model in *Case On*. Since the grey-box model is the combination of these two models, its computational demand is still lower than that of the physically-based model. Clearly, this means considerable advantage with the LR model and the grey-box model, if, for example, an online process control needs to carry out the calculations in real time.

2. The LR and the grey box model can be used very generally for any design of tubes in any hydraulic circuit (not only in solar heating systems), since they need only basic and very common pipe features: the inlet temperature, pipe flow rate and the volume (length and diameter) of the pipe. These parameters are always known or can be measured easily. Then the identification process of Section 4.1 can be conveniently made. It is natural that any model (e.g. any well-trying physically-based model) needs the known or measured values of some physical or geometrical parameters (pipe diameter, pipe length, pipe flow rate). It is also natural that some other parameters, which cannot be measured directly (e.g. heat transfer coefficient), have to be identified, before the practical application of the model, in the same way as in the identification process of Section 4.1. These issues do not reduce the generality and the practical usability of the LR and the grey box model.

7. Conclusion

In the present study, the well-trying Newton's law of cooling has been used to model the temperature change inside a pipe. Since a pipe involves delay for a fluid (element) flowing from the inlet to the outlet, this model has been completed with the mathematical description of the delay (based on former results in the literature). This completed mathematical construction has been used as the basic physically-based (white-box) model to determine the outlet (fluid) temperature of pipes.

Any day has been divided into two operating cases, *Case On* and *Case Off*, according to the current (pump) flow rate and discharging state of the pipe. Furthermore, a new mathematical model, called LR model, has been worked out in this study based on MLR. Both the

physically-based and the LR model have been identified and validated based on measured data. The LR model proved to be more precise in *Case On* but less precise in *Case Off*. These results inspired to propose a further model, a grey-box model, as a composition of the physically-based and the LR model utilizing the advantages of both ones. The identified grey-box model calculates $T_{out,mod}$ according to the identified physically-based model in *Case Off* and the identified LR model in *Case On*. The last row of Table 3 verifies that the grey-box model is in fact the more precise one, since its modelling error is the smallest, namely 3.6%, against 4.3% in case of the physically-based model and 5.1% in case of the LR model. It is much simpler and faster to use the linear algebraic equations (Eq. (5a) and (5b)) than the ODE (Eq. (4)) to calculate the outlet temperature, so the LR model has the lowest, while the physically-based model has the highest computational demand. Clearly, the grey-box model is between them, that is, its computational demand is still lower than that of the physically-based model. Regarding the modelling error, all three models are satisfactorily precise for general engineering/modelling purposes, nevertheless, the grey box-model is outstanding (with an error of 3.6%), so it can be highly recommended for the practice and for further researches. In the present study, mathematical models for pipes have been worked out based on MLR. Since MLR based models are likely the simplest possible models (because of the simple linear relations), which can still reach a very good accuracy (around or below 5%), it is worth continuing MLR based researches. For example, further MLR based models may be worked out for other working components of heating systems or whole heating systems in future research works. In particular, MLR based models for solar collectors were proposed in [23]. These models could be completed with the above LR or grey-box models for pipes and with further MLR based models for solar storages (and for heat exchangers) to construct an MLR based model for whole solar heating systems.

Acknowledgement

The author thanks the Editor for the encouraging help in the submission process, the staff of the Department of Mathematics (Faculty of Mechanical Engineering, SZIU) for the contribution and Prof. István Farkas and the Department of Physics and Process Control for the measured data.

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

References

- [1] V.D. Stevanovic, B. Zivkovic, S. Prica, B. Maslovaric, V. Karamarkovic, V. Trkulja, Prediction of thermal transients in district heating systems, *Energy Conversion and Management* 50 (2009) 2167–2173.
- [2] P. Jie, Z. Tian, S. Yuan, N. Zhu, Modeling the dynamic characteristics of a district heating network, *Energy* 39 (2012) 126–134.
- [3] R. Kicsiny, I. Farkas, Improved differential control for solar heating systems, *Solar Energy* 86 (2012) 3489–3498.
- [4] C.-K. Chen, L.-W. Wu, Y.-T. Yang, Temperature and heat flux in turbulent circular pipe flow, *Journal of Heat Transfer* 128 (1) (2005) 44–52.
- [5] R. Kicsiny, J. Nagy, Cs. Szalóki, Extended ordinary differential equation models for solar heating systems with pipes, *Applied Energy* 129 (2014) 166–176.
- [6] L.C. Evans, *Partial differential equations*, American Mathematical Society, 2010.
- [7] A. Orbach, C. Rorres, R. Fischl, Optimal control of a solar collector loop using a distributed-lumped model, *Automatica* 17 (3) (1981) 535–539.
- [8] S.R. Dhariwal, U.S. Mirdha, Analytical expressions for the response of flat-plate collector to various transient conditions, *Energy Conversion and Management* 46 (2005) 1809–1836.
- [9] P.C. Chiu, E.H.K. Fung, Flow-forced dynamics of a triple heat exchanger – Part I: Formulation of mathematical model, *International Journal of Heat and Fluid Flow* 11 (1) (1990) 10–16.
- [10] Y. Touré, J. Biston, G. Gilles, Modeling of a distributed parameter process with a variable boundary: Application to its control, *Chemical Engineering Science* 49 (1) (1994) 61–73.

- [11] L.R. Tóth, L. Nagy, F. Szeifert, Nonlinear inversion-based control of a distributed parameter heating system, *Applied Thermal Engineering* 43 (2012) 174–179.
- [12] R.M. Cotta, M.N. Özişik, Laminar forced convection inside ducts with periodic variation of inlet temperature, *International Journal of Heat and Mass Transfer* 29 (10) (1986) 1495-1501.
- [13] D.N. Basu, S. Bhattacharyya, P.K. Das, Dynamic response of a single-phase rectangular natural circulation loop to different excitations of input power, *International Journal of Heat and Mass Transfer* 65 (2013) 131-142.
- [14] P.C. Gibbons, *Physics*, second ed., Barron's Educational Series, Inc., 2008.
- [15] S.A. Klein et al., TRNSYS 16 – A transient system simulation program, Solar Energy Laboratory, University of Wisconsin-Madison, 2005.
- [16] R. Kicsiny, New delay differential equation models for heating systems with pipes, *International Journal of Heat and Mass Transfer* 79 (2014) 807–815.
- [17] S. Alotaibi, M. Sen, B. Goodwine, K.T. Yang, Flow-based control of temperature in long ducts, *International Journal of Heat and Mass Transfer* 47 (2004) 4995-5009.
- [18] J.K. Hale, S.M. Verduyn Lunel, *Introduction to functional differential equations*, Springer, 1993.
- [19] Y. Kuang, *Delay differential equations: with applications in population dynamics*, Academic Press, 1993.
- [20] S. Su, W. Dai, P.M. Jordan, R.E. Mickens, Comparison of the solutions of a phase-lagging heat transport equation and damped wave equation, *International Journal of Heat and Mass Transfer* 48 (2005) 2233–2241.
- [21] J. Escolano, F. Rodríguez, M.A. Castro, F. Vives, J.A. Martín, Exact and analytic-numerical solutions of bidimensional lagging models of heat conduction, *Mathematical and Computer Modelling* 54 (2011) 1841–1845.
- [22] G. Díaz, M. Sen, K.T. Yang, Effect of delay in thermal systems with long ducts, *International Journal of Thermal Sciences* 43 (2004) 249–254.
- [23] R. Kicsiny, Improved multiple linear regression based models for solar collectors, *Renewable Energy* 91 (2016) 224–232.
- [24] D.M. Etter, D. Kuncicky, H. Moore, *Introduction to MATLAB 7*, Springer, 2004.
- [25] I. Farkas, J. Buzás, A. Lágymányosi, I. Kalmár, E. Kaboldy, L. Nagy, A combined solar hot water system for the use of swimming pool and kindergarten operation, in: B. Frankovic (Ed.) *Energy and the environment*, Vol. I., Croatian Solar Energy Association, Opatija, 2000, pp. 81-88.
- [26] R. Kicsiny, Solution for a class of closed-loop leader-follower games with convexity conditions on the payoffs, *Annals of Operations Research* (2016), DOI 10.1007/s10479-016-2327-9 91