Remote real-time data management for education of identification and modeling

Zbigniew Ogonowski, Silesian University of Technology, 44-100 Gliwice, Akademicka 16, Poland, Phone: +4832-2371084, zbigniew.ogonowski@polsl.pl

Abstract — *The paper presents new method of teaching of the process identification and modeling (IM) for control by using remote real-time data management. The system is installed in a single-family, small building and collects the weather and building thermal data. Bad habits of simulation-only technique in IM teaching and problems which arise along the real-world data processing are discussed. The new teaching method forces students to solve problems that are neglected in the simulation-approach.*

Index Terms — *identification and system parameter estimation, modeling, data processing, data transmission management,).*

INTRODUCTION

Teaching of the process identification and modeling (IM) for control serves as an example of multi-disciplinary training. The curriculum should contain topics of data supervision and processing, numerical methods and optimization techniques. Hence IM subject is usually projected for master studies to use bachelor preparation. Unfortunately, experiences in teaching as well as the reviewing of the respective textbooks prove that the muli-disciplinarity is very unbalanced.

Theory of identification and system parameter estimation has been deeply explored for over thirty years. Number of mathematical tools has been developed and methods of data managing and processing for identification issues [1,2,3,4]. There are still remains unsolved problems e.g. continuous model identification [5] and nonlinear model identification which seems to be the most challenging task [6]. However, the basis of the curriculum is linear model identification methods which is well established subject.

The very usual scheme of the subject is to teach the theory of identification methods with computer simulation as a laboratory support. That scheme is very useful: students have the opportunity to explore all features of the certain method as convergence, bias, accuracy etc. Knowing the simulated system all results can be precisely validated, all tests as whiteness of residuals, parameters variance, information criteria etc. can be verified easily [3]. The importance of this skill is obvious. Students have to be taught the competence for proper choice of an identification method applied to the certain problem.

However, if verify in practice the above know-how finds number of additional problem which never appears during simulations. The real-world plants are much more complex then the simulated systems. The practice brings important constraints and environmental difficulties which are left out in simulations. These issues are detailed in [7]. To illustrate the problems a simple thermal stand example is presented in [7]. This simple stand is now changed with the system aimed for IM teaching.

To cope with the real-world data and IM problems, the remote real-time data management system has been set up in the Institute of Automatic Control of STU. The system is installed in a single-family, standing alone, small building and is aimed to collect the weather and building thermal data. The paper presents telemetric issues as well as the environmental layer of the system. Students have an access to the system to make respective experiments. Experimentation is restricted due to a normal use of the house by occupants. Yet another restriction imposes weather which parameters cannot be changed at all. If, for instance, outer temperature becomes an excitation signal of the model then the choice of the signal can be difficult. Careful experiment preparation is then necessary. Obviously, students are allowed to use historical data, but to understand the processes of heat exchange in the context of IM problems, the own experiments are much more valuable.

Examples of sets of data are presented in the paper as well as the models that has been found by students.

SIMULATION-ONLY BAD HABITS

The most important bad habits of simulation-only technique in IM teaching are presented below. The common ground is that the only aim of teaching is training of the mathematics (rules, algorithms and formulas). This aim is, however, somewhat hidden because the task being solved by students is to find a model that behaves as the plant. Thus accuracy is the only measure of the success. Thus students use all possible mathematical tools to reduce properly defined measure of accuracy (i.e. model validation tests). Such a construction of the exercise provides the training of the mathematics and number of important issues are missed. They are as follows:

- **Operating point choice.** The model to be identified is linear and explains deviations from the operating point. The characteristics of the plant are usually unknown in advance and determination of the operating point can be difficult. After being established, the operating point should be kept in some way (at least during the identification experiment). It could happen that the identification experiment has to be curried out in the limited period (see the point 'Duration of the experiment and number of data'). Necessity of keeping the state of the plant close to the operating point remains in a conflict with necessity of plant excitation (see the point 'Excitation signal'). It can be seen that the real experiment has to be prepared in advance and *a priori* knowledge is of the great importance. The task is then realized iteratively and *a priori* knowledge is enlarged step by step. These problems are interconnected i.e. if one is solved then the other is easier to be tackled. These issues are neglected during simulation studies. It is difficult to create the above situations in the simulation world. Even if the plant is described with the first-principle non-linear relations, it is relatively easy to find operating point using standard numerical methods and software what does not happen in practice.
- **Initial conditions.** Data for identification should be obtained in the steady state to eliminate influence of initial conditions. Disturbances affecting the plant and the nature of input signal (e.g. two-state) make difficult fulfilling zero initial state in practice. Time which is needed to reach steady state can be viewed as yet another limitation of duration of the identification experiment. Including response on initial conditions in the identification data causes necessity of additional data pre-processing [7]. All this issues are neglected in simulation-only IM.
- **Duration of the experiment and number of data.** There are several limitations influencing the experiment duration. The first comes from technology e.g. batch processes. The second follows from impossibility of certain apparatus exclusion form the process to make identification experiment. The third reason is non-stationarity of the process. Limitation of the experiment duration causes limited number of data which cannot be eliminated with shorter sampling time due to deterioration of the model accuracy [8]. Yet again these issues are omitted in simulation-only IM teaching.
- **Excitation signal.** In practice generation of the excitation signals is very limited. It follows mainly from non-linearity of the actuators (at least amplitude and rate limits). To avoid that problem the excitation has to be properly chosen not to exceed the constraints. Good selection is using of PWD modulation in this case. However, to keep input deterministic the changes of input has to be applied only in sampling time hits. The most difficult is to determine excitation range a priori (amplitude and frequency i.e. signal dynamics). Some risk is necessary due to the plant has to be properly excited but state cannot exceed assumed bounds. It is also often happens that there is no possibility of input signal generation at all because the input is determined by the environment of the plant. This situation is the most difficult in IM problems and do not appear in simulations.
- **Necessity of iterations.** Practical identification is the iterative process which enlarges knowledge concerning the plant and the model. Starting from initial statement the first experiment is done. After data processing the first model is obtained and the second experiment can be curried being much better prepared.

The above presentation shows that number of practical identification aspects are neglected during simulation-only teaching. In this paper new method of IM teaching is proposed. The basic idea is to reverse the aim of teaching to change motivation of students. The formulation of IM task states as many as possible of the problems detailed above. If identification is understood as numerical data processing (usually relating to optimization technique) then within such formulation it becomes a secondary task. Because here, identification is understood much wiser i.e. as a collection of sub-tasks being developed by students and solved step by step. To do so students need to have an access to the real process. Usually it is difficult to provide it in the laboratory thus remote system is proposed. The system is installed in a single-family, standing alone, small building and is aimed to collect the weather and building thermal data. Students have an access to the system to make respective experiments. Restrictions are imposed due to normal use of the house by occupants and due to weather which parameters cannot be changed at all. The outer temperature serves as the example: if it becomes an excitation signal then the influence on input signal is not possible. Careful experiment preparation is then necessary with all the restrictions and problems stated above.

REMOTE REAL-TIME DATA MANAGEMENT SYSTEM

There are two parts of the remote system. The first part measures thermal states of the building including temperatures, medium flows and binary states of the heating system elements. The second part measures weather parameters as outer temperature, humidity, air pressure, wind speed and direction and precipitation. All data are stored in data-base. Communication is organized trough the Internet.

Thermal state of the building

The limitations imposed by normal use of the building cause special demands concerning measurement system. All connections between units have to be wire-less. Number of units and their specification have to cover all requirements according to temperature and other states of the heating system to be measured. The general structure of the measurement system is presented in Figure 1.

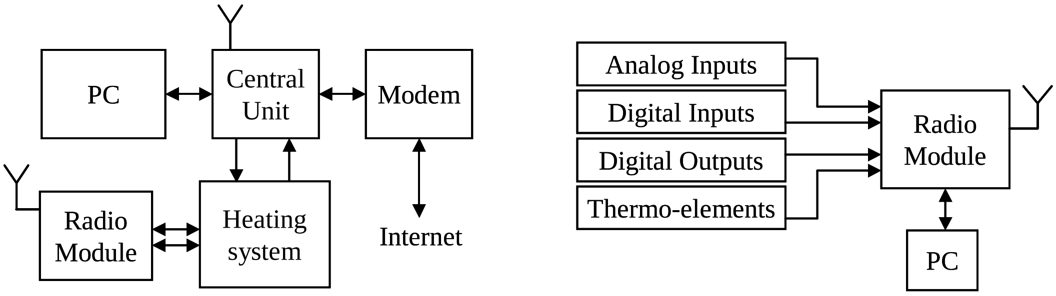

FIGURE 1
Structure of remote system (thermal state of the building).

The central unit (CU) communicates with PC directly through the serial interface or indirectly via the modem and the Internet connection. Using input-output interface the CU can collect data from heating system. It can also buffer data and transfer them out (e.g. to the PC). The wire-less communicator allows for the connection of the CU with radio-module (RM). Basically the RM can perform similar function as the CU. However, the RM unit is much smaller with poorer interface thus it can be used as a distributed measurement unit rather than as a control one. Thus, the CU is addressed to be the final control unit, while the RM serves as measurement pods.

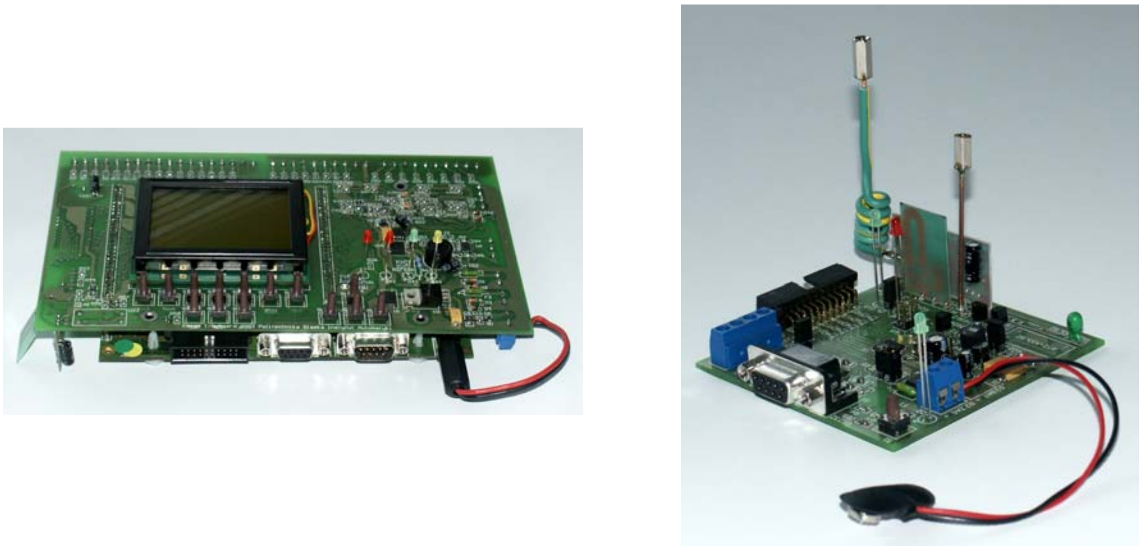

FIGURE 2
Central unit and radio module.

After careful selection it was decided to design the measurement system on the Atmel microprocessor platform. CU consists of two cards (Figure 2, left): user-designed upper card mostly devoted to local communication with the user and lower card (only part is visible in Figure 2) being the standard Atmel product EB40A with AT91R40008 microprocessor. The CU consists of A/D and D/A converters (both with 8 channels of 16 bit), two serial ports RS232, bus of Dallas digital thermo-elements, 16 digital inputs or outputs, wire-less 433MHz communicator (TeleControlli), 64kB EEPROM, 2MB Flash memory, LCD screen with keyboard and additional signaling.

RM unit is presented in Figure 2, right. Basic role of RM in the system is collection and transfer data obtained directly from the heating system. Data transfer is realized through the wire-less connection with CU. However, as it can be seen in Figure 1, the connection with PC (serial interface) is also possible thus RM can perform as the server. In this configuration RM-server collects data from other RMs remotely and transfer them to PC.

The main criteria in RM design was low cost of the single module, high reliability, long-life battery supply (a few months), wide temperature range, small dimension, simplicity of the software, not less than 10 temperature points of measurement, input-output digital interface and A/C input converter, serial interface. All the requirements has been fulfilled. The base creates microprocessor Atmel ATMega8. The size of RM is 9x9cm. The general structure of this controller is presented in Figure 3. RM can measure up to 16 temperatures using two 1-wire bus of Dallas digital thermoelements DS18S20. The temperature range is –55 up to 150 C^0 with 0.1 C^0 accuracy. Other features of the unit are as follows: 5 channel digital input/output, 6 channel 10-bit A/D input, serial interface RS232, wire-less communicator RT4 and RRQ3 of TeleControlli, SPI interface (to software update), additional switches and LEDs to set structure, starting and resetting the unit. To make the software flexible the own list of instruction has been created. Number of additional problems had to be solved during the design and testing the RN unit as: disturbances, electrical protections (separations), control of radio-communication via software, measurement synchronization etc. The most important problem concerned power supply. The solution is a combination of ATMega8 facilities, hardware elements and respect software. Between the measurements the current consumption is reduced down to 50μA (basically the need of microprocessor in sleep mode). The only current consumption is while RM communicates remotely, however, this takes a part of one second. The tests confirmed that two R4 batteries can supply the RM in three month with 1 minute sampling interval. To improve radio-communication additional antennas has been applied. Tests proved that the signal can be transferred through two floors with reinforced concrete ceilings.

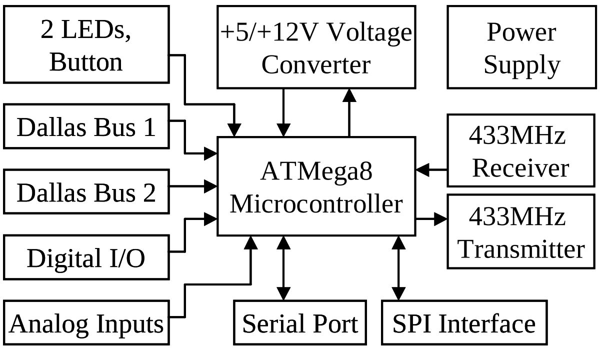

FIGURE 3
Structure of RM unit.

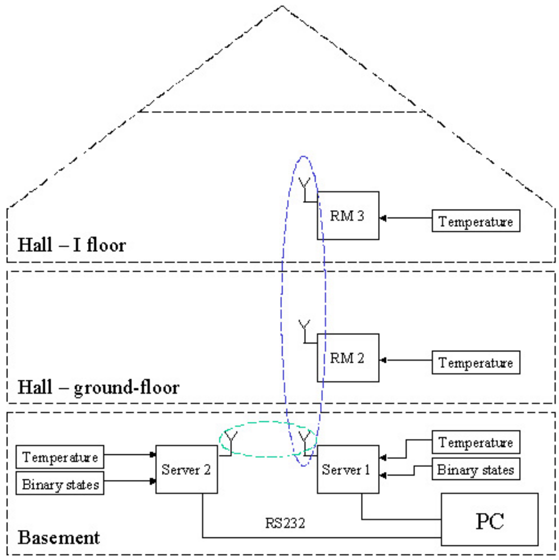

FIGURE 4
Configuration of the measuring system.

CU and RMs allows for different configuration of measurement system. Figure 4 presents an example of the configuration. Table 1 shows measured signals in this configuration.

Outer temperature	3	Boiler gear	1
Boiler temperature	2	Boiler state	1
Hot water temperature	4	Room relay state	1
Radiator temperature	4	Hot water relay state	1
Temperature in the kitchen	2	Pump gear	1
Temperature in the hall	1	Pump state	1
Temperature in the boiler room	1	Thermo-valve set-point	1
Other temperatures on 1st floor	3	Temperature of hot water tank	5
Other temperatures on the 2nd floor	3	Hot water flow	1

TABLE 1
Measured signals.

Weather parameter

Weather parameters are measured with Vaisala Weather Transmitter WXT510 installed on the building roof (Figure 5). This device is equipped with serial port (RS232 or RS485) and was connected to the CU by the wire-less connection via RM. Wind speed is measured in the range 0 ... 60 m/s (accuracy ±0.3 m/s or ±3% or ±5% depending on sub-range). Wind direction is measured all-round (0 ... 360° ±3°). Liquid precipitation is measured with respect to rainfall (resolution 0.01mm, accuracy 5%), rainfall duration counted each ten-second increment when droplet detected and rain intensity **(**one-minute running average in ten-second steps with the range 0 … 200 mm/h and the accuracy 0.1 mm/h**).** Barometric pressure is measured in the range 600 ... 1100 hPa and accuracy not less than ±1 hPa). Air temperature has the measurement range 52 ... +60 °C and accuracy ±0.3 °C. Relative humidity is of the range 0 ... 100 %RH and accuracy ±3 %RH within 0 ... 90 %RH and ±5 %RH within 90 ... 100 %RH.

FIGURE 5
Vaisala Weather Transmitter WXT510.

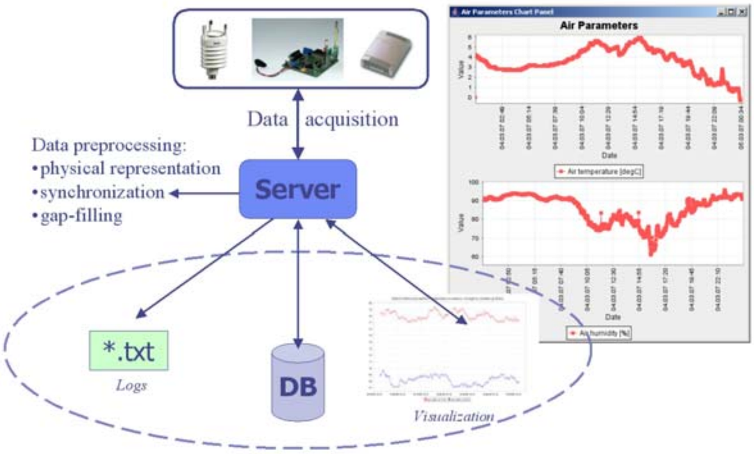

FIGURE 6
Measurement Server.

For indoor humidity measurements Vaisala humidity transmitter (HMW60) was coupled with RM analog input. Humidity transmitter is mounted on the hall wall. The unit is complete with 4-20mA output, temperature compensated and 10-35VDC input power. The HUMICAP 180 RH sensor, Pt 1000 RTD allows for humidity measurement in the whole range with accuracy ±2%.

Data-base and communication

The Measurement Server (Figure 6) written in Java language collects data from wireless measurement system and weather transmitter every minute. Collected data are recalculated to the physical units and stored in the data base. The Server provides also synchronization service together with gapfilling based on suitable approximation algorithm. The online data visualization as well as the historical charts can be also delivered.

Communication with the Measurement Server is organized trough the Internet. TCP/IP communication allows for client connection (in the form of Java Applets) and monitoring of all measured signals. Adequate security mechanisms are applied (user logging, IP filtration etc.). The communication protocol is universal which means that students can access Measurement Server with their own software written e.g. in C++ or C Scharp. Java technology allows also for remote logging with mobiles (GPRS) or palmtops etc. Generally speaking the Internet interface allows students configuring, experimenting and data management in almost unlimited manner.

EXAMPLE OF IM TASK

The task being solved by students is formulated in ‘a basic’ terms i.e. to identify certain element(s) of the heating system. Conceptually the heating system is easy to understand, thus students can formulate sub-task easily. These sub-task correspond to IM practical issues presented in the first section. Below an example of the task and obtained data are presented.

Figure 7 shows example of thermal model (structure) of the building. The goal of the identification was to obtain model which can be useful for control algorithms synthesis. The indoor method of control was used so the model had to explain chosen inner states. This additional degree of freedom simplified identification procedure. Not all measured data were then used. So the Figure 7 follows from some simplifying assumptions.

The general model consists of elementary models (e.g. boiler model, radiators model etc.) which explain the heat transfer in respective parts of the building. However, student task refers to a part only unless the more complex task for a group of students is formulated. In this example the task concern identification of only K_h transfer function of the hall. The measured signal T_h is not only influenced with the hall response on radiator temperature T_r changes but it is also disturbed by other signals as outer and kitchen temperatures and by the fire place. Students realize that if one can extract from data the periods with only one of these disturbances varying then the representation of the model can be as shown in Figure 8, and iterative solution can be applied as follows:

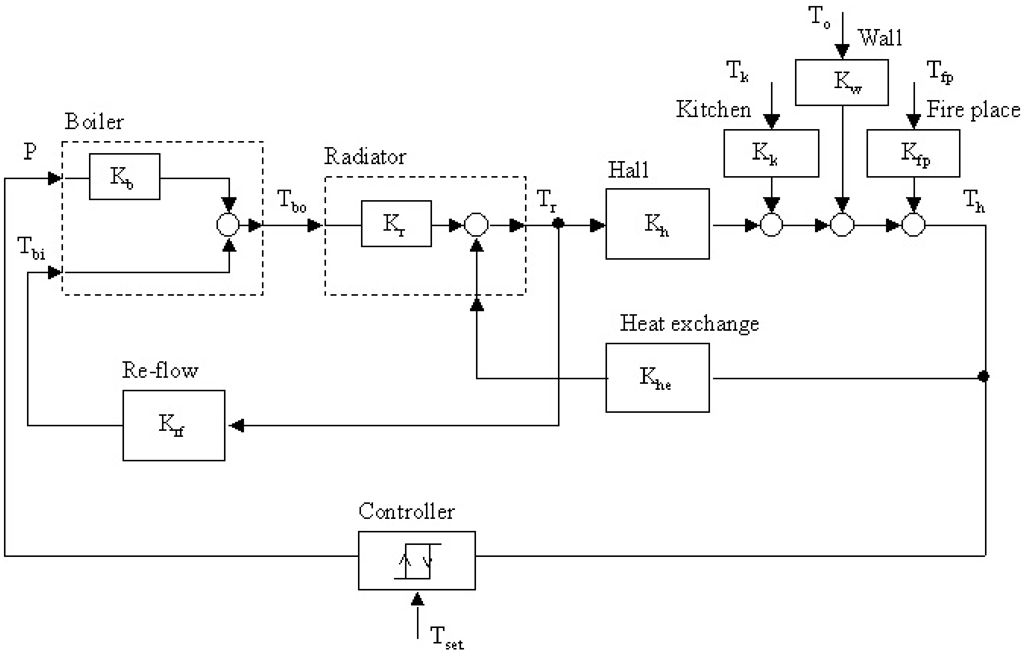

FIGURE 7
Example of thermal model.

I iteration

1. part
   - identification of K_{p1} model using S_I and S_o
   - simulation of K_{p1}, response for $S_I - S'_{11}$ is obtained
   - calculation of $N'_{11}=S_o-S'_{11}$
2. part
   - identification of K_{n1} model using N_I and N'_{11}
   - simulation of K_{n1} response for $N_I - N'_{12}$ is obtained
   - calculation of $S'_{12}=S_o-N'_{12}$

II iteration

1. part
   - identification of K_{p2} model using S_I and S'_{12}
   - simulation of K_{p2} response for $S_I - S'_{21}$ is obtained
   - calculation of $N'_{21}=S_o-S'_{21}$
2. part
   - identification of K_{n2} model using N_I and N'_{21}
   - simulation of K_{n2} response for $N_I - N'_{22}$ is obtained
   - calculation of $S'_{22}=S_o-N'_{22}$

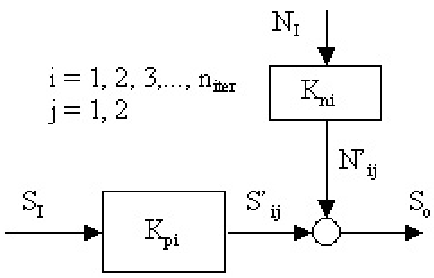

FIGURE 8
Model structure for iterative method.

Necessary data collected during IM experiment are presented in Figure 9. The figure is respectively scaled to show all signals. The upper one refers to T_h, the middle (two-state) represents state of the boiler, lower lines corresponds to the respective temperatures.

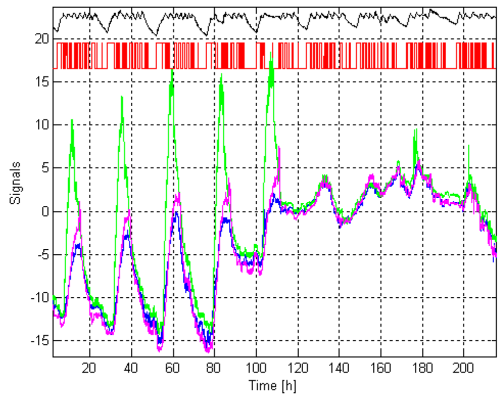

FIGURE 9
Example of data-set.

Figure 10 presents results of the identification: left is the first iteration, right is the third. The improvement has ben obtained by proper data selection as indicates above iterative procedure.

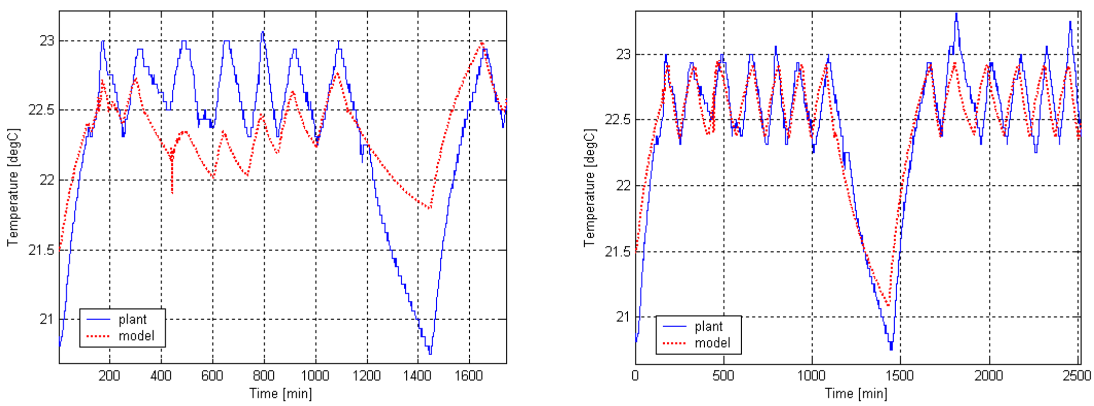

FIGURE 10
Result of the iterative identification.

The second part of the remote system allows formulation of number of IM task concerning weather influence which create almost unlimited possibility of IM teaching. These includes non-stationary models, non-linear problems, discrete-continuous transformations etc. Example of outer temperature, humidity, wind speed and direction is shown in Figure 11. Respective models of weather conditions influence on thermal comfort and fuel consumption can be identified using these data.

Example of the task for student is to use artificial neural network (ANN) models to find the parameter that influences the indoor temperature the most. The goal was achieved by comparing the network parameters (weight factors between input neurons and hidden layers) and network response analysis (response for change of only one input signal, while other signals being kept constant and equal to its mean value). Each learning process for different set of data (the same signals from different time periods) was performed until the goal-error is reached. The exemplary network response with real data comparison is presented in Figure 12. However, the first step of data analysis was simple hidden layer coefficients comparison. Such comparison delivered information about relative gain influence of measured signals on the indoor temperature. Conclusions from this step allowed for a choice of signals which influence ANN outputs the most. In this example students found important outer temperature, wind speed and direction. Interesting was also the conclusion that the air pressure influences significantly the fuel consumption which is usually hidden while efficiency of the heating system is evaluated. This follows from the nature of the combustion process and is highly connected with the type of heat source used in the system. For example this influence will diminish a lot if the closed combustion chamber is applied.. Second step of the analysis concerned network response which is depicted in Figure 12. It is clear that indoor temperature in biased the most with the outer temperature and the wind speed.

The above IM task serves as an example of iterative identification. It is aimed to eliminate necessity of identification of models which explain reaction of the heating system which are to small for the further analysis. Thus, next step can be reformulated i.e. eventual second experiment (second set of data) should be projected in another way.

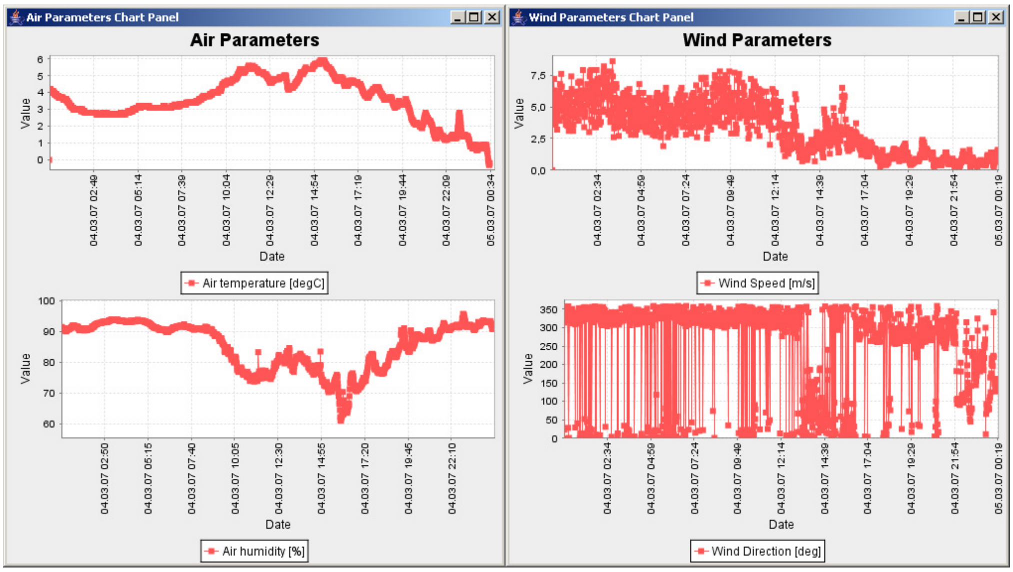

FIGURE 11
Weather parameters data.

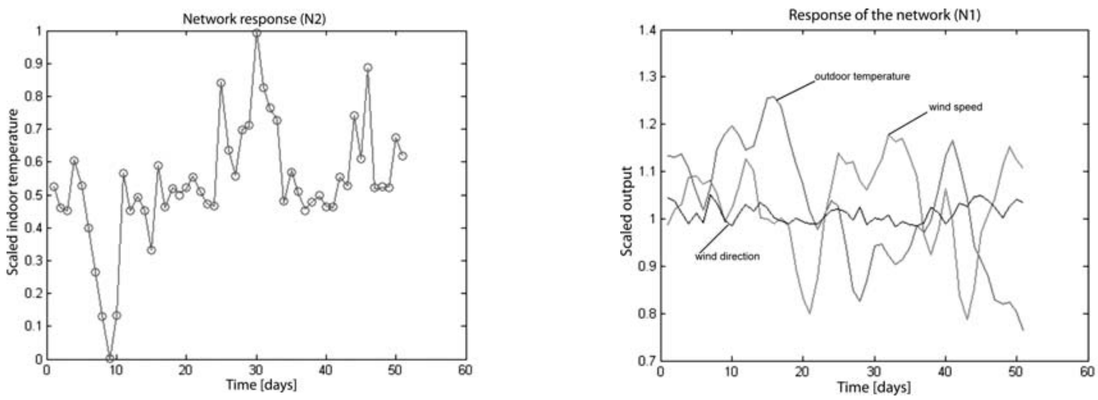

FIGURE 12
Artificial neural network responses.

CONCLUSION

It follows from the above presentation that remote real-time data management for education of identification and modeling allows for interactive teaching aimed for reversing of the bad habits following from simulation-only training. One should formulate the IM task on the basis of real process and concrete needs. This can lead student to develop the necessity of the proper formulation of the IM experiment. During the experiment number of practical problems need to be solved. This is the crucial issue of the proposed methodology of teaching. Solving of these problems (see the first section) students develop mathematical tools in a natural way and understand them much more precisely than in the case of simulation-only training.

The remote real-time data management for education of identification and modeling allows for almost unlimited access to all its elements. Students can then create their own experimental schemes to use different identification methods.

REFERENCES

[1] Söderström, T., Stoica P., "System identification", Prentice Hall International, London, 1989.
[2] Johansson, R. "System modeling and identification", Prentice Hall, Englewood Cliffs, New Jersey, 1993.
[3] Ljung, L. "System identification - Theory for the user", Prentice Hall, New Jersey,1999.
[4] Pintelon, R., Schoukens, J., "System identification. A frequency domain approach", IEEE Press, New York, 2001.
[5] Garnier, H., Wang, L., "Identification of continuous-time models from sampled data", Springer-Verlag, London, 2008.
[6] Khalil, H.K., "Nonliear Systems", Macmillan Publishing Co., Canada, 1992.
[7] Ogonowski, Z., "Real-world experimenting for education of dynamical system modeling and identification", *International Conference on Engineering Education*, Gliwice, Poland, 2005.
[8] Novak, E., "Deterministic and Stochastic Error Bounds in Numerical Analysis" *Lecture Notes in Math.* 1349 (1988), Springer-Verlag, Berlin.