Lumped-parameters Control-oriented Gray-box Modelling of Direct Liquid Cooling Systems

Michele Lionello[1], Mirco Rampazzo[1], Alessandro Beghi[1], Damiano Varagnolo[2], Mattias Vesterlund[3]

***Abstract*—Liquid cooling systems have better heat dissipation capabilities than air based ones, and are expected to become a standard choice in future data centers, due to the ever increasing power density and heat rejection needs of the compute infrastructure. A convenient side-effect of implementing liquid cooling is that it facilitates the efficient recovery of the heat waste. However designing and managing these heat recovery infrastructures benefit from having control-oriented models that can accurately describe how different operating conditions of the to-be-cooled heat sources will affect the thermal status of the coolant. The aim of this manuscript is to derive control-oriented models of direct cooling systems, i.e., systems where the compute infrastructure is immersed in a vessel filled with dielectric fluid. More specifically we derive, starting from physical interpretations, a general lumped-parameters gray box dynamical model that has - as inputs - the electrical consumption of the heat sources and the working point of the heat recovery system, and has - as outputs - the temperature distribution of the coolant in the most relevant points of the system. Beyond proposing this modelling methodology we also validate the generalization capabilities of the obtainable models. In specific, we test the achievable statistical performances in a field case, plus compare with the ones of classical black box system identification strategies. We thus report that in the considered field case our gray box model reached a fit index of 91.08% when simulating test sets, while the best black box model we have been able to identify reached (on the same test sets) fit indexes of only 72.56%.**

***Index Terms*—liquid cooling, gray box modelling, lumped parameters, system identification, heat recovery**

I. Introduction

Data centers facilities, the supporting backbone of the telecom industries infrastructure, are composed of rooms filled with enterprise servers and equipment dedicated to the storing, managing, and distribution of information [1]. The exponential growth of demand of data centers services has increased the associated energy usage to the point that operating data centers in an efficient way is nowadays crucial.

In particular, the cooling infrastructure within a data center may account for up to 40% of the total electricity usage, [2]. Increasing the cooling efficiency represents thus an opportunity to reduce data centers energy costs and environmental impact.

We then notice that most of the currently existing data centers use air-based cooling systems, in the sense that the heat produced by the servers is removed by chilled air that is drawn through the servers. In typical setups servers are arranged in racks so that cooling air first enters the room into a so-called *cold aisle* placed in front of the racks, then passes through the servers, and exits on their back into a so-called *hot aisle*. Here, hot air rises and moves to a Computer Room Air Conditioning (CRAC) unit, where air is cooled and recirculated.

However, air is fundamentally an inefficient cooling method due to its low density and low heat removal capacity [3]. With the emergence of high-performance microprocessors in servers, this limitation will be more and more tangible [4]. Since liquids have generally far superior thermophysical properties than air, the expected technological solution is employing either *direct* or *indirect liquid cooling* strategies.

To be more precise, indirect liquid cooling solutions implement a Coolant Distribution Unit (CDU), i.e., a closed and controlled liquid coolant circuit with two properties: *1)* the coolant is chilled by an external cooling source, and *2)* the circuit is attached to the to-be-cooled electronic devices, so that their cooling is performed mainly by conduction [5]. The main drawback of this approach is that it has typically low versatility: it requires indeed installing sealed enclosures, rack and server levels piping systems, and everything must be tailored to the specific servers and facility layouts [6].

In Direct Liquid Cooling (DLC) solutions instead servers are directly immersed in a dielectric fluid (e.g., mineral oil) that ensures electrical insulation. The main advantages of DLC solutions are the adaptability of the cooling solution, since no sealed enclosures and piping are required at the server level, the reduced density compared to air-based cooling systems, and the opportunity to use the coolant as an energy storage.

All the various liquid cooling approaches share the possibility of implementing improved heat recovery strategies. To be precise, air cooled data centers may exploit heat pumps to recover the heat from the exhaust air into district heating supply flows (as being currently done in Stockholm, Sweden [7]). However, since liquid coolant exhaust has a far bigger exergetic content than air coolant exhaust, recovering heat from a liquid cooling system implies smaller needs for heat pumping – i.e., a more energy efficient process.

To further improve the energetic efficiency of the heat recovery process, the benefits from the advantages of using liquid cooling may be coupled with the unique opportunities offered by implementing advanced control systems, that are known to have the potential of enhancing both systems efficiency and performance by mostly acting on software components.

Nowadays, it is a common practice in the design of control systems to make extensive use of Computer Aided Control

[1]Department of Information Engineering, University of Padova, Padova, Italy. Email: {lionello, rampazzom, beghi}@dei.unipd.it

[2]Department of Engineering Cybernetics, Norwegian University of Science and Technology, Trondheim, Norway. Email: damiano.varagnolo@ntnu.no

[3]Swedish Institute of Computer Science North, Research Institutes of Sweden, Luleå, Sweden. Email: mattias.vesterlund@ri.se

Systems Design (CACSD) software tools [8]. These tools allow to simulate the relevant system dynamics, for a first assessment of the different control strategies. In order to setup a useful simulation-centric control system design project, however, there is a preliminary step to be taken into account, namely the derivation of a dynamic model, which translates certain interesting properties of the real system into mathematical equations of the to-be-controlled system.

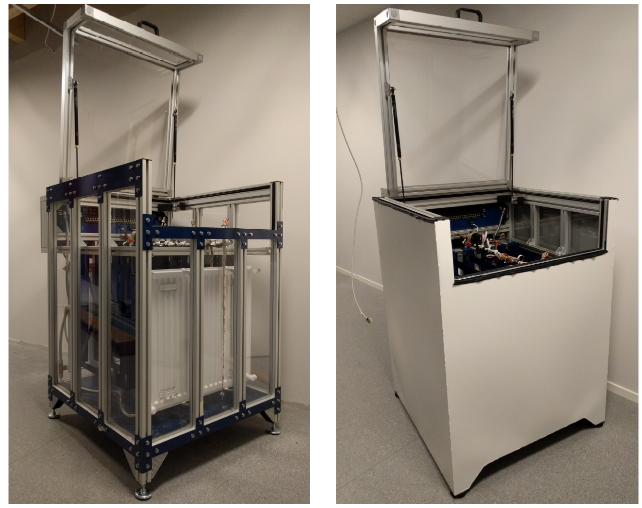
In the considered scenario, to further maximize the energetic efficiency of the heat recovery process through advanced control techniques it is of paramount importance to have detailed quantitative models of the thermal dynamics within the system, e.g., to implement model predictive control approaches to decide where to geographically allocate IT loads within the data center, how to run the various pumps and valves, etc. In other words, quantitative control-oriented modelling is instrumental to the maximization of the exergetic content of the coolant and of the efficiency of the heat recovery step.

However, to the best of our knowledge there are very few contributions so far that deal with control-oriented liquid immersion cooling system modeling and analysis. In [9] the cooling performance of an immersion cooling system with natural convection for high power servers is evaluated by Computational Fluid Dynamics (CFD) simulations and actual experiments. By combining CFD simulations and Matlab & COMSOL-based models [10] analyzes the cooling system of servers which are immersed in a dielectric liquid where water is used to transport the heat outside of the data center. [11] proposes a model of a cooling system of computational devices which are in direct contact with the coolant. It thus seems that there exists a lack of analyses that have *control* as the final user of the results.

In this paper we thus consider the specific problem of modelling DLC systems from a thermal dynamics perspective [12]. More precisely, we aim at first drawing general considerations through analyzing the specific experimental vessel described in Section II, then at deriving a flexible strategy for the gray-box modelling of general DLC systems inspired by energy-based modelling techniques like Power Oriented Graph (POG) [13], (Section III). In other words, we exploit the concept of power flows within a physical system to create a lumped parameters quantitative description of the heat dynamics within a vessel. To validate the model we then perform an ad-hoc parameters identification step, and verify its approximation capabilities in Section IV. In the same section we also compare the statistical performance of our calibrated gray box against classical black-box models that have been learned through Prediction Error Method (PEM) system identification approaches. As a short anticipation of our results, we report that the tuned gray box model can simulate test datasets with fit indexes up to 91.08%, while we haven't been able to identify black box models with fit-performances on the same test datasets higher than 72.56%.

II. The considered experimental DLC testbed

The experimental testbed used in this manuscript for both inspiration and field tests purposes is physically located within the Infrastructure and Cloud research & test Environment (ICE) facility at Research Institutes of Sweden (RISE) Swedish Institute of Computer Science North (SICS) in Luleå, Sweden, a facility dedicated to testing innovative technologies for data centers [14]. The considered DLC setup is composed of a $0.84\,\mathrm{m}$ side cubic vessel, built with an aluminum frame with sealed acrylic glass walls as in Figure 1a, and comprising an external and fully welded $2\,\mathrm{mm}$ thick metal-sheet shell to eliminate the risk of leaks as in Figure 1b.



(a) Vessel without the external metallic shell.

(b) Vessel with the external metallic shell.

Fig. 1: Pictures of the considered experimental testbed.

The vessel has been filled with a dielectric oil with low viscosity and good heat transfer properties, and has been used to cool four Open Compute Windmill V2 servers donated from Facebook. The waste heat generated by the servers is removed from the tank through a dedicated and water-based closed cooling circuit. More precisely, this cooling circuit comprises two heat exchangers immersed in the dielectric oil (one being visible in Figure 1a) and connected through dedicated copper pipes to an external plate fin water-to-air heat exchanger. A dedicated fan then forces external air through the fins of this exchanger, so that the heat produced by the Windmill servers is eventually dispersed into the air. A schematic representation of the considered experimental DLC testbed is reported in Fig. 2.

The oil-filled vessel is designed so to exploit natural convection to enhance the cooling effectiveness of the system. In practice the heating components of the servers (mainly their CPUs) heat the immediately surrounding oil, that will naturally rise; simultaneously, the immersed heat exchangers cool the immediately surrounding oil, that will naturally fall. This induces a natural oil flow within the vessel, something beneficial since with this configuration the oil never passes through a pump, something that can cause pump-degradations over time.

To understand better the phenomena involved in the servers cooling process, the vessel comprises a set of temperature

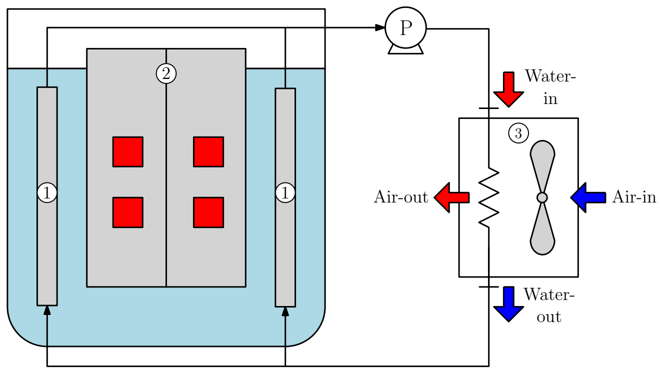


Fig. 2: A schematic representation of the considered experimental DLC testbed, including: immersed heat exchangers (1), servers (2) with CPUs a red squares, the pump (P), and the external water-to-air heat exchanger (3).

sensors and a dedicated Supervisory Control And Data Acquisition (SCADA) system to acquire and manage the collected information. More precisely, the measurements consist in:

- temperatures of the cooling water flowing into the vessel and out from the vessel;
- mass flow rate of this water;
- air temperature just before and after the plate fin heat exchanger (3 in Figure 2);
- flow rate of the air forced by the external air fan associated to the plate fin heat exchanger;
- power metering of each server's electrical power usage (red squares in Figure 2);
- a matrix of thermocouples was immersed in the oil to find the temperature profile along a desired vertical section of the vessel.

III. Gray-box modelling of DLC systems

We consider a strategy to quantitatively model DLC systems through an approach that lumps parameters that builds on an energetic modeling approach. More precisely, we propose to exploit the concept of power flows within a physical system, and thus assume that the overall physical system is composed of elements (nodes) that: *a)* shall be considered geographical portions of the system, *b)* shall be represented by a thermal capacity and a local temperature, and *c)* can exchange heat with other nodes or with the external world. The total mathematical model can then be derived by the set of energy conservation equations that hold for each node of the system.

The benefits of using this modeling technique are variegate: first, it guarantees satisfying the conservation of energy within the system, and thus embedding in the model a law that the system is expected to follow. Moreover, once the system topology has been defined, the model can be written in a straightforward and automatic way. This also ensures computationally fast numerical simulations. Finally, adding new elements to the model does not require to rewrite the model but only to increase the size of the matrices of the model while keeping the same logical structure.

The proposed lumped parameters model is characterized by a low model order. The limited number of parameters can thus be easily identified from experimental data, and for this reason the model is suitable for real-time simulations and for control design purposes. It can indeed be used to analyze the system dynamics, to predict non-measured variables, to study the effect of parameters variations, to test different control strategies, and as a model-based design tool.

A. Constructing the thermal model

Oil nodes: we start by considering that, due to temperature stratification phenomena, there exist a natural stratification process that leads to horizontal oil layers with different temperatures. We thus consider a generic division of the tank into O *oil nodes* that represent and capture the transient behavior of these horizontal oil layers. We assume that each oil node $o = 1, \dots, O$ is characterized by an average thermal capacity C_o^{oil} and an average temperature T_o^{oil}. We also consider that there exist thermal heat transfer parameters $g_{o,o+1}^{\mathrm{oil}}$ between connected pairs of oil nodes, i.e., assume that these oil nodes can exchange heat with the neighbor ones. Defining as $d_{o,o+1}^{oil}$ the distance between the center of nodes o and $o+1$, parameter $g_{o,o+1}^{\mathrm{oil}}$ can be expressed as $g_{o,o+1}^{\mathrm{oil}} = g^{\mathrm{oil}} d_{o,o+1}^{oil}$, g^{oil} is the thermal conductivity of the oil. Also, each oil node is expected to have some interactions with the external environment through the walls of the vessel; for this reason we consider heat transfer coefficients $g_o^{\mathrm{loss}} = g_o^{\mathrm{walls}} A_o^{\mathrm{walls}}$ between each oil node and the walls of the vessel in contact with this oil layer, where g_o^{walls} is the convective heat trasnfer coefficient through the walls and A_o^{walls} is the corresponding heat exchange areas. For the experimental setup considered in Section II we consider $O = 3$ layers, corresponding to the logical top, middle, and bottom of the tank.

CPU nodes: we moreover consider the presence of heating sources (i.e., the CPUs of the servers) that are immersed in the vessel. We thus model these sources with a number of *CPU nodes*, each described by its temperature $T_{s,c}^{\mathrm{cpu}}$. Here the subscripts s and c denote respectively the server index and the index of the CPU within the server s. Moreover, if S is the total number of servers immersed in the vessel and C is the number of CPUs per server, $s = 1, \dots, S$, and $c = 1, \dots, C$. In the field case described in Section II the testbed comprises $S = 4$ servers, each one including $C = 2$ CPUs. Related to these nodes, we consider that it is in general possible to measure the electrical power usage $\dot{Q}_s^{\mathrm{cpu}}$ of each server s, and that typically it is not possible to break down this power usage to each individual CPU. This implies that the general model needs to include a parameter that describes how the server power usage $\dot{Q}_s^{\mathrm{cpu}}$ should be broken down on the various individual CPUs.

Heat exchanger nodes: we also consider the presence of a number of heat exchangers that are immersed within the oil. Through each heat exchanger, cold water flows at flow rate ϕ^{hx} in order to extract heat from the coolant within the DLC system. For this type of nodes we consider that the within-the-exchanger coolant enters (typically from the bottom) the heat exchanger at a temperature $T_{\mathrm{in}}^{\mathrm{hx}}$, and exits it (typically from the top) at a higher temperature $T_{\mathrm{out}}^{\mathrm{hx}}$. For simplicity, we assume that the oil is in contact with a fictitious immersed heat exchanger having external area equal to the total area of both the heat exchangers considered in the experimental setup.

Let moreover $T^{\mathrm{ext.Air}}$ be the temperature of the external air. Denoting the state and input vectors $\boldsymbol{x}$ and $\boldsymbol{u}$ with

$$\boldsymbol{x} = \left[T_1^{\mathrm{oil}},\ \ldots,\ T_O^{\mathrm{oil}},\ T_{1,1}^{\mathrm{cpu}},\ \ldots,\ T_{S,C}^{\mathrm{cpu}}\right]^{\mathrm{T}}, \tag{1}$$

$$\boldsymbol{u} = \left[T^{\mathrm{ext.Air}},\ T_{\mathrm{in}}^{\mathrm{hx}},\ \phi^{\mathrm{hx}},\ Q_1^{\mathrm{cpu}},\ \ldots,\ Q_S^{\mathrm{cpu}}\right]^{\mathrm{T}}, \tag{2}$$

our general DLC system model becomes

$$L(\boldsymbol{\theta})\dot{\boldsymbol{x}} = A(\boldsymbol{\theta})\boldsymbol{x} + B(\boldsymbol{\theta})\boldsymbol{u} + \boldsymbol{\ell}\left(\boldsymbol{x}, \boldsymbol{u}; \boldsymbol{\theta}\right) \tag{3}$$

where $\boldsymbol{\theta}$ is the vector of all the parameters describing the model through the opportune matrices L, A, B, and nonlinear map $\boldsymbol{\ell}$. Note that in the remainder of the manuscript we will omit the dependency of these quantity on θ, assuming it tacit. More precisely, moreover,

- the square matrix L is diagonal and contains all the thermal capacities of the respective nodes;
- the square matrix A collects the various heat exchange coefficients among the various nodes and between the nodes and the external environment. As an example, and for simplicity, assume that for the field case described in Section II we let all the heating sources (i.e., the CPUs of the servers) be immersed in the top oil node. Then A becomes

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \tag{4}$$

  where the various blocks A_{ij} are as in (5);
- the matrix B captures the effects of the outdoor temperature and the servers input powers on the temperatures of the various oil and CPUs nodes. For the field case described in Section II this matrix becomes

$$B = \begin{bmatrix} B_{11} & \mathbf{0}_{3\times 2} & \mathbf{0}_{3\times 4} \\ \mathbf{0}_{8\times 1} & \mathbf{0}_{8\times 2} & B_{22} \end{bmatrix} \tag{6}$$

  with $B_{11} = \begin{bmatrix} g_1^{\mathrm{loss}} & g_2^{\mathrm{loss}} & g_3^{\mathrm{loss}} \end{bmatrix}^T$ and

$$B_{22} = \begin{bmatrix} \alpha & 0 & 0 & 0 \\ 1-\alpha & 0 & 0 & 0 \\ 0 & \alpha & 0 & 0 \\ 0 & 1-\alpha & 0 & 0 \\ 0 & 0 & \alpha & 0 \\ 0 & 0 & 1-\alpha & 0 \\ 0 & 0 & 0 & \alpha \\ 0 & 0 & 0 & 1-\alpha \end{bmatrix}. \tag{7}$$

  Note that the coefficient α in B_{22} corresponds to the parameter describing how the server power usage $\dot{Q}_s^{\mathrm{cpu}}$ should be broken down on the various individual CPUs;
- the non-linear vector field $\boldsymbol{\ell}\left(\boldsymbol{x}, \boldsymbol{u}\right)$ represents the heat exchanged by the oil nodes and the heat exchanger nodes. Given that our aim is to obtain a control-oriented numerical representation of the thermal dynamics within the DLC vessel, we model this heat exchange phenomenon exploiting the ε-Number of Transfer Units (NTU), [15]. This method is used typically to characterize the heat transfer in heat exchangers when inlet temperatures of the fluids and the heat transfer coefficient are available.

More precisely, we define temperatures T_i^{hx} of water in the heat exchanger portions in contact with the $i-th$ oil layer. Then, we solve iteratively the following backward equation:

$$T_i^{\mathrm{hx}} = T_i^{oil} + (T_{i+1}^{\mathrm{hx}}) \exp\left(-NTU_i\right),\ i = O, \ldots, 1, \tag{8}$$

by initializing $T_{O+1}^{\mathrm{hx}} = T_{in}^{\mathrm{hx}}$. It is worth noting that water exiting the heat exchanger is at temperature T_1^{hx}, i.e. $T_{out}^{\mathrm{hx}} = T_1^{\mathrm{hx}}$. The quantity NTU in (8) is the number of transfer units, i.e.

$$\mathrm{NTU}_i = \frac{g^{\mathrm{hx}} A_i^{\mathrm{hx}}}{\phi^{\mathrm{hx}} c_p^{water}}, \tag{9}$$

where $c_p^{water} = 4185\,\mathrm{J\,kg^{-1}\,K^{-1}}$ is the specific heat of water.

Once obtained water temperatures T_i^{hx}, each component of field $\boldsymbol{\ell}$ can be obtained as:

$$\ell_i = \begin{cases} g^{\mathrm{hx}} A_i^{\mathrm{hx}} \Delta T_{\mathrm{LMTD},i}, & i = 1, \ldots, O \\ 0 & i = O+1, \ldots, O + S\cdot C \end{cases}, \tag{10}$$

where $\Delta T_{\mathrm{LMTD},i}$ is the logarithmic mean temperature difference, defined as:

$$\Delta T_{\mathrm{LMTD},i} = \frac{T_{i+1}^{\mathrm{hx}} - T_i^{\mathrm{hx}}}{\log \dfrac{T_{i+1}^{\mathrm{hx}} - T_i^{\mathrm{oil}}}{T_i^{\mathrm{hx}} - T_i^{\mathrm{oil}}}}. \tag{11}$$

IV. Model calibration and validation

We present two different types of results in two separate subsections:

1) the outcomes of estimating the parameters of the proposed gray box model through a least-squares approach on top of datasets obtained from the testbed described in Section II;
2) quantitative and qualitative comparisons of the approximation capabilities of the estimated model against the ones of classical black box models whose parameters are

$$A_{11} = \begin{bmatrix} -g_{1,2}^{\mathrm{oil}} - \sum\limits_{s,c} g_{s,c}^{\mathrm{cpu}} - g_1^{\mathrm{loss}} & g_{1,2}^{\mathrm{oil}} & 0 \\ g_{1,2}^{\mathrm{oil}} & -g_{1,2}^{\mathrm{oil}} - g_{2,3}^{\mathrm{oil}} - g_2^{\mathrm{loss}} & g_{2,3}^{\mathrm{oil}} \\ 0 & g_{2,3}^{\mathrm{oil}} & -g_{2,3}^{\mathrm{oil}} - g_3^{\mathrm{loss}} \end{bmatrix} \quad A_{12} = \begin{bmatrix} g_{1,1}^{\mathrm{cpu}} & g_{1,2}^{\mathrm{cpu}} & \ldots & g_{4,2}^{\mathrm{cpu}} \\ 0 & 0 & \ldots & 0 \\ 0 & 0 & \ldots & 0 \end{bmatrix} = A_{21}^T \quad A_{22} = -\mathrm{diag}\left(g_{1,1}^{\mathrm{cpu}}, g_{1,2}^{\mathrm{cpu}}, \ldots, g_{4,2}^{\mathrm{cpu}}\right) \tag{5}$$

identified through PEM approaches on top of the same datasets used for the gray box case.

A. Estimating the parameters of the proposed gray box model

We consider a Fisherian approach, and thus the model parameters to be deterministic. We then assume to be endowed with a training set $\mathcal{D} = \left\{ (\boldsymbol{x}(k), \boldsymbol{u}(k))_{k=1,\dots,K} \right\}$ and estimate θ through a classical Least-Squares (LS) approach

$$\widehat{\theta}_{\mathrm{LS}} = \arg \min_{\theta \in \mathbb{R}^8_+} \sum_{k=1}^{K} \left\| \boldsymbol{x}(k) - \widehat{\boldsymbol{x}}(k; \theta) \right\|^2 \qquad (12)$$

where the simulated $\widehat{x}(k; \theta)$ states are obtained by propagating the forward-Euler discrete-time counterpart of the continuous-time dynamics in (3) initialized with the first measurement in the training dataset, i.e., with $\widehat{x}(1; \theta) = \boldsymbol{x}(1)$. Note that with this notation we assume the dependency of $\widehat{x}$ on the measured inputs $\boldsymbol{u}(k)$ as tacit. Finally, we notice that the hypothesis space is a positive quadrant.

For completeness, we plot the traces of the measured input signals that have been used for both training and testing the proposed model parameters in Figures 3 and 4. Plots of the outputs and a discussion on the generalization capabilities of the model are instead delayed to Section IV-B.

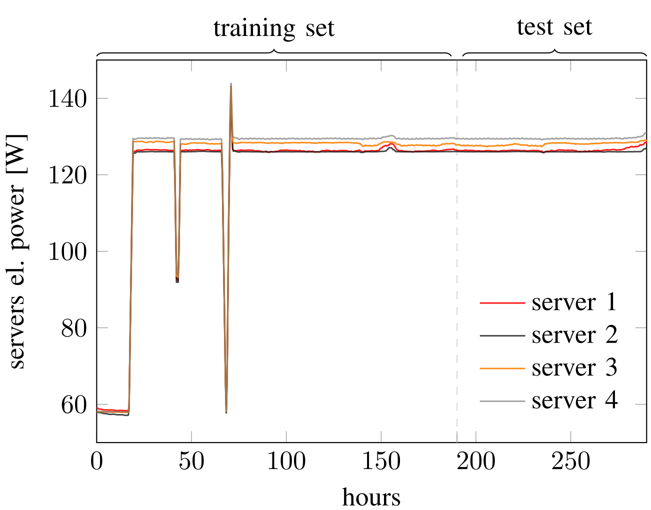


Fig. 3: Temporal evolution of the measured electrical power usage of the 4 servers immersed in the considered DLC vessel. The spikes in the training set correspond to servers shutdowns and restarts.

For completeness, point-estimates of the values of the parameters are reported in Table I.

B. Comparing the proposed gray box against classical black box models

A natural question that shall be asked is whether it is worth to put the effort of deriving a gray box for this type of system, or if it was better to take a black box instead.

To answer this methodological question we thus performed black box PEM identification steps, and then we compared

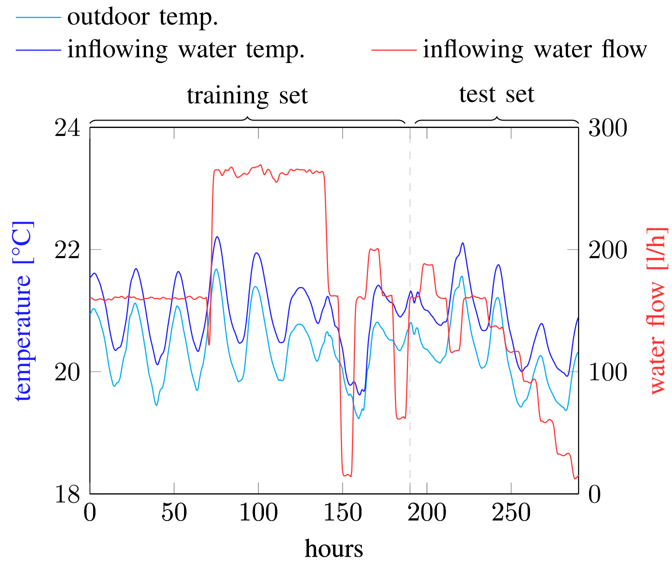


Fig. 4: Temporal evolution of the other signals considered as inputs for the considered DLC vessel. Note how the inputs in the training set are designed to excite the system around different working conditions.

parameter	*meaning*	*point estimate*
g^{oil}	oil conductivity	$64.96\,\mathrm{W\,m^{-1}\,K^{-1}}$
g^{walls}	walls heat transf. coeff.	$\{7.46, 2.54, 2.51\}\mathrm{W\,m^{-2}\,K^{-1}}$
g^{hx}	HX heat transf. coeff.	$40.48\,\mathrm{W\,m^{-2}\,K^{-1}}$
g^{cpu}	CPUs conductance	$3.84\,\mathrm{W\,K^{-1}}$
α	thermal load partition	0.59

TABLE I: Summary of the point-estimation results for the parameters of the proposed gray box model. Note that all the geometrical parameters, e.g., A^{hx} and d^{oil}, have been measured through direct inspection of the vessel.

the capabilities of both gray box and the best black box in estimating the outflowing water temperature T^{hx}_{out}, that is the most relevant quantity from a heat recovery perspective. In particular, we consider both classical linear models (specifically, ARX, ARMAX, OE and BJ) and nonlinear ones (specifically, Hammerstein-Wiener and wavelet networks with different standard choices of the input and output nonlinearities and wavelet functions), and in practice mimicked the standard steps that what one would follow when facing the problem of building black box models of the system under consideration. Inputs of the black box models are all the components of vectors $\boldsymbol{x}$ and $\boldsymbol{u}$ in (2), while the output is the temperature of water exiting the heat exchanger, i.e. T^{hx}_{out}.

We then compare the generalization capabilities of both gray and (best) black box models through the qualitative and quantitative means of Figure 5 and Table II respectively. We reports only the results obtained on the test set, but the differences observed between training and test sets are small, suggesting a good generalization capability.

A bit surprisingly, despite expecting to identify a nonlinear model as the best black box one, among the several tests that

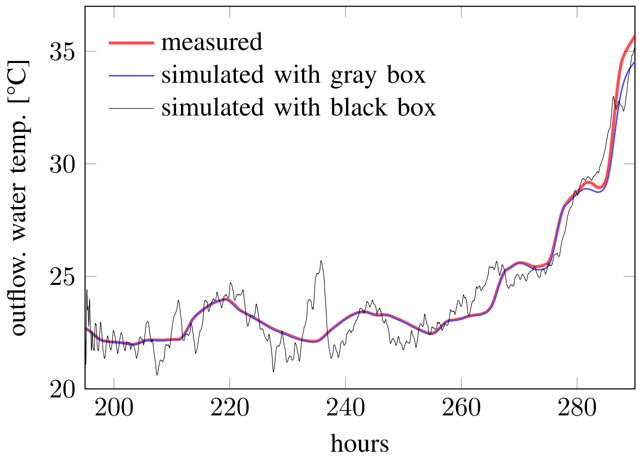


Fig. 5: Measured vs. simulated output data obtained through the identified gray and black box models on a test set.

	gray box	best black box
RMSE	0.27 °C	0.91 °C
fit	91.08 %%	72.56 %

TABLE II: Statistical indexes on the test set associated to the identified gray and black box estimators.

we performed the best black box model that we have identified is an Output Error model with 3 as the maximum order of the polynomials in its transfer functions.

As Table II indicates, learning the models using the training set shown in Figures 3 and 4 led the gray box model to have far better simulation performance than the black box one on the test set shown in the same figures. Obviously collecting more data is expected to lead to better statistical performances, specially for the black box modelling approach, so that the conclusions above cannot be considered as general. However, we also noticed that collecting that specific training set took 8 days, and we have no clear intuitions of how many days of experimenting one needs to arrive at the breaching point of equal predictive performance between the two different strategies. Our conclusion is thus that from practical perspectives the proposed gray box model should be favored.

V. Conclusions

Taking a control-oriented approach to the problem of modelling liquid cooled systems is important, since it enables developing model-based control algorithms (e.g. predictive) for maximizing the efficiency of the associated heat recovery blocks. Specifically, due to the nature of the phenomena that are involved in the DLC system (e.g. the heat generation and thermal heat transfer with very slow and nonlinear dynamics) and the available a priori expert knowledge, we chosen to model the system by means of a gray box approach. First we derived the functional structure of the dynamics of the system starting from physical principles, then considering the limited number of model parameters, we applied the classical frequentist parameters estimation approaches on training sets corresponding to field experimental campaigns lasting very few days. We thus verified that the proposed gray box modelling approach enables to find models with generalization capabilities that are better than the ones of black box models that are identified with the very same data. In conclusion, our field case seems thus that the proposed gray box modelling approach has practical sense.

The next and main important steps that we foresee as continuation of this research line are essentially two: first, couple our thermal model with models of different heat recovery systems, so to have a complete and holistic quantitative picture of the data center plus heat harvester infrastructure. Second, validate the whole control-oriented approach by developing control strategies for maximizing the efficiency of the heat recovery and test them in real life conditions.

Acknowledgment

We would like to thank all the staff at RISE SICS North for supporting this work with their time and technical expertise.

References

[1] J. Dai, M. M. Ohadi, D. Das, and M. G. Pecht, *OPTIMUM COOLING OF DATA CENTERS*. Springer, 2016.

[2] M. Avgerinou, P. Bertoldi, and L. Castellazzi, "Trends in data centre energy consumption under the european code of conduct for data centre energy efficiency," *Energies*, vol. 10, no. 10, p. 1470, 2017.

[3] A. H. Khalaj and S. K. Halgamuge, "A review on efficient thermal management of air-and liquid-cooled data centers: From chip to the cooling system," *Applied Energy*, vol. 205, pp. 1165–1188, 2017.

[4] M. Saini and R. L. Webb, "Heat rejection limits of air cooled plane fin heat sinks for computer cooling," in *Thermal and Thermomechanical Phenomena in Electronic Systems, 2002. ITHERM 2002. The Eighth Intersociety Conference on*. IEEE, 2002, pp. 1–8.

[5] M. J. Ellsworth, G. F. Goth, R. J. Zoodsma, A. Arvelo, L. A. Campbell, and W. J. Anderl, "An overview of the ibm power 775 supercomputer water cooling system," *Journal of Electronic Packaging*, vol. 134, no. 2, p. 020906, 2012.

[6] P. E. Tuma, "The merits of open bath immersion cooling of datacom equipment," in *Semiconductor Thermal Measurement and Management Symposium, 2010. SEMI-THERM 2010. 26th Annual IEEE*. IEEE, 2010, pp. 123–131.

[7] J. Lind and E. Rundgren, "Industrial symbiosis in heat recovery collaborations between data centers and district heating and cooling companies," Master's thesis, KTH, Stockholm, 2017.

[8] C. S. Chin, *Computer-Aided Control Systems Design: Practical Applications Using MATLAB and Simulink*. CRC Press Taylor and Francis Group, 2017.

[9] M. Matsuoka, K. Matsuda, and H. Kubo, "Liquid immersion cooling technology with natural convection in data center," in *2017 IEEE 6th International Conference on Cloud Networking (CloudNet)*, 2017, pp. 1–7.

[10] A. Almaneea, H. Thompson, J. Summers, and N. Kapur, "Cooling system analysis for a data center using liquid immersed servers," *International Journal of Thermal Technologies*, vol. 4, no. 3, pp. 200 – 207, 2014.

[11] A. Alexander and T. Anatoliy, "Mathematical model of liquid immersion cooling system for supercomputer (in Russian)," *Program Systems: Theory and Applications*, vol. 1, no. 28, pp. 187–199, 2016.

[12] J. Cadafalch, D. Carbonell, R. Consul, and R. Ruiz, "Modelling of storage tanks with immersed heat exchangers," *Solar Energy*, vol. 112, pp. 154 – 162, 2015.

[13] F. Grossi and R. Zanasi, "Dynamic models of thermal systems using an energy-based modeling approach," *ASME J Heat Transfer*, vol. 138, no. 10, pp. 102 801 – 12, 2016.

[14] "Swedish institute of computer science, infrastructure and cloud datacenter test environment," https://ice.sics.se/.

[15] R. K. Shah and D. P. Sekulic, *Fundamentals of heat exchanger design*. John Wiley & Sons, 2003.