Detecting and modelling air flow overprovisioning / underprovisioning in air-cooled datacenters

Emanuele Simonazzi*, Miguel Ramos Galrinho†, Damiano Varagnolo‡, Jonas Gustafsson§, Winston Garcia-Gabin¶

* Department of Information Engineering, University of Padova, Padova, Italy.
Email: `emanuele.simonazzi@studenti.unipd.it`

† Department of Automatic Control, KTH, Stockholm, Sweden.
Email: `galrinho@kth.se`

‡ Department of Computer Science, Electrical and Space Engineering Luleå University of Technology, Luleå, Sweden.
Emails: `damiano.varagnolo@ltu.se`

§ RISE ICT/SICS North, Luleå, Sweden.
Email: `jonas.gustafsson@ri.se`

¶ ABB Corporate Research, Västerås, Sweden.
Email: `winston.garcia-gabin@se.abb.com`

Abstract—**When cooling and exhaust air flows in air-cooled datacenters mix, the energetic efficiency of the cooling operations drops. One way to prevent this mixing of happening is by augmenting the air tightness of the hot and cold aisles; this, however, requires installing opportune hardware that may be expensive and require time consuming installations. Alternatively, one may try to minimize cooling and exhaust air flows mixing by opportunely controlling the speeds of the fans of the Computer Room Air Handling (CRAH) units so that the distribution of the air pressure field within the computer room is favorable.**

Implementing this type of flow control requires both detecting when there actually is some type of flow mixing somewhere, plus understanding how to operate the cooling infrastructure so that these mixings do not happen. To this aim, there is the need for models that can both help deciding whether these mixing events occur, plus designing automatic control strategies for reducing the risks that they will happen.

In this manuscript, we propose an ad-hoc methodology for the data-driven derivation of control-oriented models that serve the purposes above. The methodology is built on classical Prediction Error Method (PEM) approaches to the system identification problem, and on laddering on the peculiarities of the physics of the phenomena under consideration. Moreover, we test and assess the methodology on a industrial-scale air-cooled datacenter with an installed capacity of 240 kW, and verify that the obtained models are able to capture the dynamics of the system in all its potential regimes.

Index Terms—**Datacenters cooling, statistical learning, energy efficiency, switching systems**

I. Introduction

In datacenters´ cooling equipment has an energy consumption comparable to the electrical power fed in the computing part [1]. A sound strategy for improving air-cooling based datacenters' efficiency is to implement cooling schemes that avoid overprovisioning cooling flows to the servers [2]. However, detecting the occurrence of these phenomena and modeling their effects on the thermal dynamics of the computer room are both non-trivial tasks.

We then notice that, to the best of our knowledge, there are no structured and datacenter-oriented methodologies for finding models of the airflows that simultaneously:

- do not require programming Computational Fluid Dynamics (CFD) simulations or implementing tests that use air flows speed measurement systems;
- return control-oriented models, so that the results will enable implementing model predictive control strategies for the operation of the datacenters' CRAH units;
- are flexible enough to be implementable in a large variety of datacenters designs.

A. Literature review

The EU began making efforts in data center power efficiency by the end of 2010 through the launch of the projects *All4Green* and *CoolEmAll*. Since then, 6 more projects aiming to reduce the environmental impact of data centers have been launched [3].

The high number of extensive projects reflects the complexity of the issues associated to the modeling of the thermal dynamics in datacenters, which involves large models based on physics first principles [4] or ad-hoc CFD simulations. Thermal principles have also different

temporal and dimensional scales: from the ones referring to the internal thermal behavior of servers, as described in [5], the ones referring to the dynamics within single racks, as described in [6], to the ones of the air flows within the entire data center room, as in [7] or [8]. Complex thermal dynamics are also present within the individual cooling units, as confirmed in [9], [10], and in association with the humidity levels of the coolant flows [11].

Dynamics can be also found in the management of the cloud (i.e., of the various IT requests depending on the queues of services that the compute infrastructure has to serve). For example, [12] proposes a model for estimating in real time the amount of resources (e.g., CPU and memory) needed for satisfying a given service. These models are very useful for service deployment, for real-time identification of resource bottlenecks or with the objective to maintain smooth operation of different services and minimize downtime (e.g., [13]).

B. Statement of contributions

Despite the abundant literature on the modeling of datacenters' thermal dynamics, to the best of our knowledge it seems that there is a lack of publicly available studies on the general detection and data-driven modeling of flow overprovisioning or underprovisioning phenomena. In this manuscript, we address this issue, and more specifically consider the specific flow modeling problem connected to the identification of different air mixing regimes at the computer room level in general room configuration scenarios. In other words, we propose a strategy that is in our intentions at least theoretically applicable to any air-cooling based setup.

To effectively illustrate this general methodology, we consider a standing example, graphically represented in Fig. 1, where the CRAH is (arbitrarily) drawn on the left and the servers racks are on the right.

In the setup shown in Fig. 1, the best situation from a cooling flow provisioning point of view is when the direct cold flow and the return hot air flow do not mix. Flow overprovisioning can be then graphically described though the left panel in Fig. 2; the right panel, obviously, graphically represents a flow underprovisioning situation.

CRAHs flow overprovisioning or underprovisioning can theoretically occur in all that plants that do not employ dedicated hardware that ensure air tightness. This means that, our standing example deals with a very specific room configuration, the need for detecting and modeling these phenomena is present also in other configurations (e.g., raised floor setups).

Among all the potential strategies for identifying (and in a later stage controlling), these flow mixing phenomena, we here consider the category of methodologies that want to solve the issue without requiring the usage of additional sensors or implementing ad-hoc CFD simulations, since this is economically appealing.
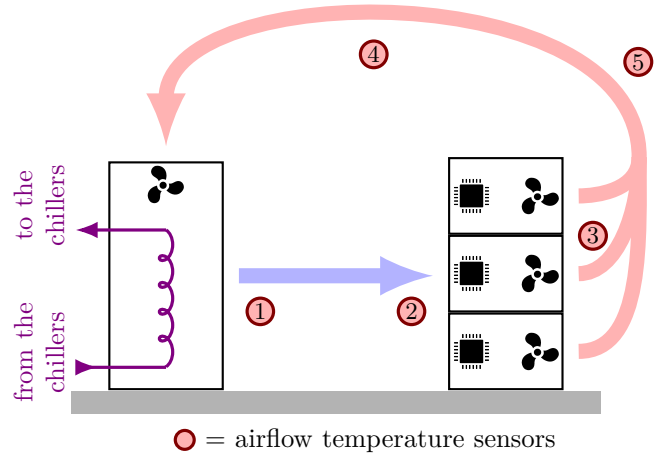

Figure 1. Schematic representation of the computer room that will serve as a standing example for illustrating the methodology proposed in this manuscript. Note that this scheme corresponds to the physical testbed considered in our field tests and described in Sec. II.
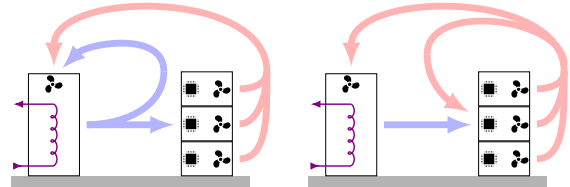

Figure 2. Schematic representation of the situations corresponding to CRAH units overprovisioning cooling flows (in the left panel) or underprovisioning them (right panel).

In this paper we thus propose a data-driven methodology to:

- detect if there exist different mixing regimes in a generic computer rooms;
- create data-driven models that can account and describe these different mixing regimes if they exist and that can be used for designing tailored control systems;
- implement and assess this methodology on field data from a real system, using a standing example that can help clarify how the strategy could be implemented in other settings.

C. Organization of the manuscript

Sec. II describes the physical system where we did our experiments. Sec. III describes the ad-hoc model describing different mixing regimes and to be identified using data driven techniques. Sec. IV describes the methodology for detecting different mixing regimes using data-driven techniques plus reports our results from field data. Sec. V concludes the manuscript reporting what we learned from developing the methodology and applying it to field sit-

uations, and proposes which research and development directions are opened by our contributions.

II. Testbed description

The experimental tests were performed in module 2 at Research Institutes of Sweden (RISE) Swedish Institute of Computer Science (SICS) North SICS-ICE facility, an experimental 240kW wide datacenter intended for testing innovative datacenters management strategies and situated in Luleå, Sweden [14]. The considered computer room is composed by ten server-racks placed in two parallel rows of five racks each. As is also represented in Figures 3, these two rows are in their turn placed so to create a hot aisle between them and two cold aisles between the racks and the four CRAH units (for a total of two CRAHs in each cold aisle). The setup can be thus graphically represented through the scheme shown in Fig. 1. For more information about the facility please see https://ice.sics.se/.
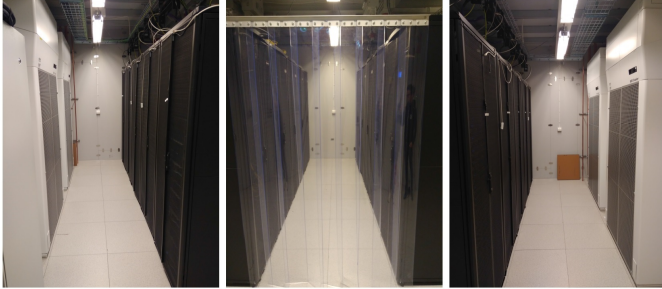

Figure 3. Photos of the cold aisles (left and right panels) and the hot aisle (center panel) of the computer room used to test and validate our methodology. In this setup the cold air exits from the various CRAH units, enters into the servers, passes through them and cools them down, then exits from the severs, rises and recirculates back to the CRAHs by entering at their top. The air is then cooled down by passing through an opportune cooling grid, so that the cycle can be repeated.
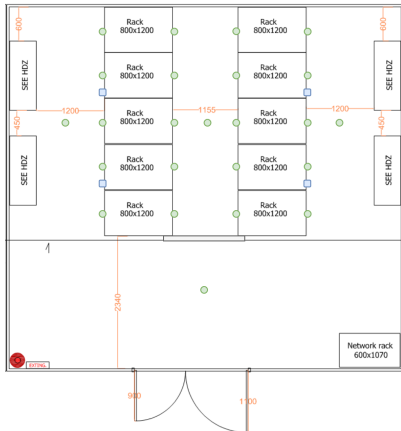

Figure 4. Instrumentation diagram of the computer room used for our experiments, together with the placement of two different types of aisles temperature sensors (respectively green circles and blue squares).

III. The proposed model

We consider control-oriented models of the combined thermal and flow dynamics within computer rooms that can be expressed through opportune time-invariant differential equations leading to nonlinear dynamics. his can be generically expressed as

$$\dot{y} = f(y, u) \tag{1}$$

where the order of the differential equations may actually be higher than one, y typically denotes temperatures, and u denotes the operating condition of the computer room (typically in terms of fan speeds and IT load percentages). More precisely, and given the intuitions about the existence of flow overprovisioning / underprovisioning phenomena illustrated in Figures 1 and 2, we consider that the general thermal dynamics (1) can be expressed with a switching Multi Input Multi Output (MIMO) Linear Time Invariant (LTI) system defined by

$$y = \begin{cases} \dfrac{B_{\mathrm{u}}(q)}{A_{\mathrm{u}}(q)} u + v_{\mathrm{u}} & \text{if } u \in \Omega_{\mathrm{u}} \\[2ex] \dfrac{B_{\mathrm{o}}(q)}{A_{\mathrm{o}}(q)} u + v_{\mathrm{o}} & \text{if } u \in \Omega_{\mathrm{o}} \end{cases} \tag{2}$$

where:
- the subscripts u and o follow the mnemonics

  u ⇔ underprovisioning   o ⇔ overprovisioning

  so that, for example, the situation $(x, u) \in \Omega_{\mathrm{u}}$ indicates a generic flow underprovisioning regime (like, for example, the one in the right panel of Fig. 2). Note that here for simplicity we consider only two models; however, the procedure is actually general, so that if desired we may consider a higher number of different air flows regimes;
- the input and output vectors u and y are to be opportunely determined starting from a-priori information about the plant and measured data (an operation that will be described in details in Sections IV-B and IV-D);
- the transfer matrices $\dfrac{B_\star}{A_\star}$, the domains $\Omega_\star$ and the spectra of the noises $v_\star$ (with $\star = \mathrm{u, o}$) are to be identified using field data (an operation that will be described in details in Sections IV-C, IV-E, and IV-F).

In words, our modeling problem is thus twofold: *i)* finding which u and y are appropriate for describing potential flow mismatches occurrences in the system; *ii)* identifying the thermal dynamics of the flow mixing in a data-driven fashion. The second point serves the special purpose of obtaining models that can help operating the CRAHs so that they do not overprovision or underprovision coolant.

IV. Methodology

This section presents a methodology for obtaining models that have the aim of aiding to detect of if a system

is operating in a flow overprovisioning or flow underprovisioning regime. More precisely, the methodology consists of the following steps, each with its own aims, and each independently described in detail in the various subsections below:

1) Identifying in which zones it is relevant to detect flows mixing phenomena.
2) Listing the important sensors and actuators available in the plant.
3) Designing and executing air flow experiments.
4) Identifying the most relevant signals through correlation analysis.
5) Determining the regions where the flow models are approximately linear.
6) Identifying and validating the models using data-driven approaches.

A. Identifying in which zones it is relevant to detect flows mixing phenomena

Given the wide variability of potential computer room designs, there is a long list of places where one may experience flow mixings. For example, in a computer room designed like the one in Fig. 1 it is safe to assume that these phenomena may happen as schematized in Fig. 2. Situations may in any case greatly change depending on the specific computer room under consideration. Other configurations (e.g., raised floor or backdoor cooling) would obviously lead to different situations. Our suggestion is thus to do, as a first step, a visual analysis of the plant using Piping and Instrumentation Diagram (P&ID) to perform a first guess of what may happen from an air flows mixing point of view.

Example: for our standing example of the room shown in Figures 3 and 4, the most noticeable potential mixing effects are intuitively as the ones in Fig. 2.

B. Listing the important sensors and actuators available in the plant

After performing the step described in Sec. IV-A the user should decide the structure of (1), i.e., decide what shall be the composition of the input and output vectors u and y of the model. This shall be performed once again by visually inspecting the P&ID diagrams and the tabulates from the Supervisory Control And Data Acquisition (SCADA) systems so to understand which sources of information are available in the plant and that may either cause or be correlated to overflow or underflow phenomena. More precisely, we suggest to consider:

- all the air temperature and air flow sensors that are available in the computer room and that are at least suspected to be measuring something related to the flow overprovisioning / underprovisioning physical phenomena that one wants to model;

- all the sources of information related to the thermal and mass exchanges induced by the CRAH units within the considered computer room (e.g., liquid flows and temperatures at the inlet and outlet, temperatures of the air at the inlet and outlet, etc.);
- all the ventilation actuation signals that are most responsible for the air mixing phenomena, and thus all the instantaneous rotational speeds of the various air handling units within the considered computer room or sufficiently close to the point where one may suspect flow overprovisioning / underprovisioning to happen (and thus the signals relatives to the fans of the servers, of the CRAH units, etc., that are close to that geographical zone under consideration). Note that, from a practical perspective point of view, fans speeds of geographically close servers should be averaged into a unique signal, to reduce the number of inputs to the system.

Regarding the last point, note that information about the rotational speeds of the various fans of the various servers is not always available. In this case, it is meaningful to use as a proxy (and of course if measured) the instantaneous power consumptions levels of the various fans. It is indeed known both from physics-based laws and quantitative evidence that the power consumptions increase cubically with the rotational speed, with the coefficients of the cubic polynomial to be identified through either opportune experiments or the fans' datasheets. To this aim, see also the experimental setup in [5]. If also this information is not available then one may resort to use as a proxy the IT loads levels of the various servers.

Remark 1 Not all the identified inputs will be controllable by the user. For example, the rotational speed of the fans of the servers are virtually always determined by the internal cooling control systems of each individual server. For this reason, in the following we will distinguish between controllable inputs and non-controllable (but still measurable) inputs.

Example: in our field case the relevant signals are: the various CRAHs output temperatures (say y_1), the racks input temperatures (say y_2), the racks output temperatures (say y_3), the cold aisle and hot aisle temperatures (say respectively y_4 and y_5, shown with green circles in Fig. 4), the temperature sensors in front of the racks (say y_6, shown with blue squares in Fig. 4), the CRAH fans speeds (say u_1), and the average IT loads of the servers within the computer room (say u_2). The outputs y_* with their corresponding references are also illustrated in Fig. 1.

C. Designing and executing air flow experiments

To be descriptive, data-driven models need to be trained on datasets that represent all the various working condi-

tions in which the system is expected to operate. This implies that the user shall collect the dataset $\mathcal{D} = \{u_1, \ldots, u_m, y_1, \ldots, y_p\}$ of those signals corresponding to the inputs and outputs listed in Sec. IV-B while running the datacenter around all the various potential set-points of its operating conditions. For our specific problem of identifying flow overprovisioning or underprovisioning situations we suggest to design the controllable inputs as follows:

1) divide the inputs in two different types: *i)* the CRAHs fans speeds, and *ii)* all the rest of the various potential controllable inputs (e.g., overall IT load within the datacenter, temperature of the inlet water to the CRAHs, etc.);
2) as for the controllable inputs of type *ii*, generate a number of fixed operating conditions using a Latin hypercube sampling approach (with a number of samples that depends on how much experimental time is available);
3) for each of the operating conditions above, design the CRAHs fan speeds as descending stair signals that start from the maximum fan speed supported by the CRAH units to arrive to that minimal fans speed that guarantees the servers to remain within their thermal comfort limits. We suggest to let these CRAHs fans speeds be composed of approximately 10 steps in amplitude, with each step lasting sufficiently long so that the datacenter reaches or is near to reach a thermal equilibrium.

Letting the datacenter run for each of the operating conditions (and thus also stairs signal on the CRAHs fans speeds) above, we get a specific part of the overall dataset, say $\mathcal{D}_c$. We actually suggest to perform each of these experiments twice, and visually inspect that there is repeatability, (i.e., the same inputs lead to the same outputs). This is something that one may think is guaranteed to hold, but that in our experience actually does not. When repeatability does not hold, one may have failed to take into consideration all the controllable inputs that have some effect on the to-be-modelled dynamics, so that step IV-B shall be repeated.

Example: for our field case, we consider the descending stair signal shown in Fig. 5. Here, one of the CRAH units fan speed is progressively decreased from 80% to 30% while the other CRAH units are kept running at their maximum level (in this specific case, moreover, the overall IT load within the computer room was kept at 50%). In this case, smaller steps have been made around the zone where the system is expected to transition from a flow overprovisioning situation to a flow underprovisioning one.
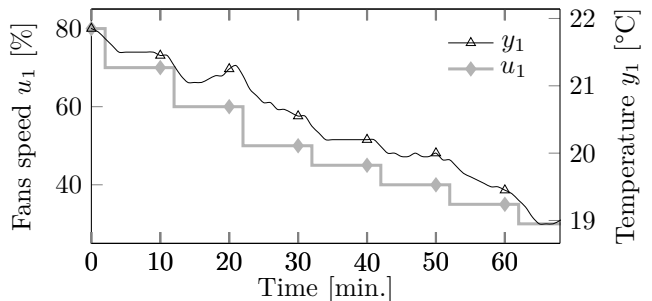

Figure 5. Example of an input signal used to command the CRAH units in the datacenter room used in our field tests during the collection of the datasets. For the sake of completeness, the figure shows also the temperature of the air flow produced by that CRAH unit.

D. Identifying the most relevant signals through correlation analysis

By executing the steps listed in Sections IV-A and IV-B, the user will get an intuitions-based listing of all the signals involved in the dynamics (1). Importantly, the process of creating this list:

- cannot be easily automated, and must thus rely on human choices;
- may end up with a potentially overly large list of inputs u_i that are *suspected* to cause or correlate to the flow overprovisioning / underprovisioning phenomena, and outputs y_j that are informative for detecting these events.

We thus let the methodology include a step of pruning this set of potential u's and y's, so that the user will retain only that components that are having an actual information content. Although more advanced approaches may be performed, we suggest to start this pruning step by performing a correlation analysis. More precisely, assume to have merged all the datasets $\mathcal{D}_c$ described in Sec. IV-C into a dataset $\mathcal{D} = \{u_1, \ldots, u_m, y_1, \ldots, y_p\}$ of synchronized time-series with every component in this dataset being a specific signal. What can happen is then that:

- two distinct inputs u_i and u_j are highly correlated, say, e.g.,

$$\left| \frac{\mathbb{E}\left[(u_i - \mathbb{E}\left[u_i\right])(u_j - \mathbb{E}\left[u_j\right])\right]}{\text{st.dev}\left(u_i\right)\text{st.dev}\left(u_j\right)} \right| \geq 0.95. \qquad (3)$$

Then it may be beneficial to eliminate the input that is, from numerical perspectives, least correlated to the various outputs y_l. Even if obvious, it is important to remark that it is of paramount importance that the dataset $\mathcal{D}$ captures all the operating conditions that the datacenter is expected to experience in its operations. This means that if one detects that two controllable inputs u_i and u_j are highly correlated then there is some mistake in the step of designing the controllable inputs;

- similar concepts applies to couples of outputs y_i and y_j: if there exist highly correlated outputs then it may be meaningful to ignore one of the two;
- if instead an input u_i is uncorrelated to all the various outputs y_j then this *may* be an indication that this input is superfluous and that it *may* be discarded too. Notice that using the conditional tense is mandatory here, since the potential non-linearity of the dynamics may lead to empirically uncorrelated signals even if under the presence of deterministic causation;
- a similar concept applies when an output y_i is simultaneously uncorrelated to all the various inputs u_j and other outputs y_s. In this case this may be seen as an indication that there *may* be no extractable information from the sensor (something that in any case can be double-checked during the data-driven modelling step). If this happens, it means that the sensor is, from a modelling perspective, badly placed. It may thus be worth to investigate if in this case it is meaningful to move the sensors into an other geographical location.

Example: in our standing example, the first output that we discarded was y_6, because it was highly correlated to y_2 (mainly because of their geographical vicinity, see also Fig. 4).

As for identifying which sensors provide most information for the modelling purpose, we then compare the readings from the various temperature sensors against the CRAH fan speed signal. In this way, we note that almost all these temperatures increase when the CRAH fan speed drops. The unique exception is the temperature of the air cooling flow in output from the CRAH, (i.e., y_1, which decreases as the fan speed decrease, as shown in Fig. 5).

Being the associated sensor directly connected to the CRAH output, the signal y_1 does not contain information that is useful for detecting flow mixing. It can however work as a normalization factor, in the sense that it can be used to compute the temperature differences

$$\Delta_{21} = y_2 - y_1,$$
$$\Delta_{31} = y_3 - y_1,$$
$$\Delta_{41} = y_4 - y_1,$$
$$\Delta_{51} = y_5 - y_1,$$

that represent the change of temperature in the room from the CRAH output to an other zone in the computer room (see Fig. 1 for a physical interpretation of the various Δ's).

Among the various Δ's, the signal that is most affected by flow overprovisioning or underprovisioning phenomena is Δ_{21}. This can be seen both from intuitive

perspectives, by checking Fig. 1, but also by analyzing quantitatively the collected evidence. In other words, and as also Fig. 6 graphically confirms, Δ_{21} is the Δ signal that is most correlated with u_1. This means that this will be the signal that we will want to model when we will perform our system identification step.
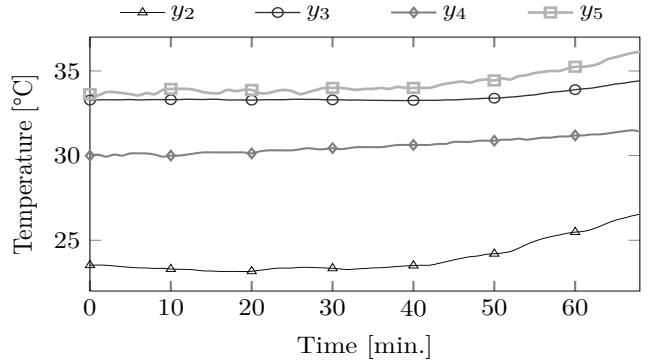

Figure 6. Comparison of the temperatures measured by the various sensors within the computer room used in our experiments while the CRAH fans speed u_1 was as in Fig. 5.

E. Determining the regions where the flow models are approximately linear

The key point of this manuscript is to identify when the overprovisioning or the underprovisioning of the cooling flows happen, and obtain data-driven models that can forecast the thermal dynamics of the room when these phenomena occur. Referring to Figures 1 and 2 for an intuitive explanation, these events happen depending on the values of the air pressure field within the computer room (something that is in turn affected by the speeds of the various fans rotating within the computer room). This intuitive explanation can then be mathematically expressed by letting the domains Ω_o and Ω_u in model (2) depend only on the values of the various fans speeds, and not on the various temperatures measured within the computer room. Incidentally, we also notice that typically there are no or very little flow speed measurements sensors within modern datacenters computer rooms. This means that flows mixing can be detected only through the proxy of checking how much the temperatures of these flows mix, as suggested also in Sec. IV-D.

To this aim, we propose to determine Ω_o and Ω_u using the following ad-hoc strategy: consider the controllable inputs of type *ii* defined in step 2 in Sec. IV-C, (i.e., consider separately every set of fixed operating conditions on the various controllable inputs). For each of the sets above, consider the associated CRAHs fan speeds, consisting of descending stair signal. Recall that every step of this signal lasts enough to let the system reach or almost reach its thermal equilibrium. It thus immediately follows that the

various output signals y_j will be a series of step responses. Consider then the physics of the flows overprovisioning / underprovisioning phenomena: the intuition suggests that the gains will typically be much smaller when the CRAHs are overprovisioning the cooling flows. This implies that these input-output gains are expected to be clearly different, depending on the flows region. This eventually implies that from the step responses computed above it should be immediately possible to not only compute the input-output gains of the system for the various operating conditions, but also verify for which operating conditions the system experiences a shift from flow overprovisioning to flow underprovisioning (i.e., determining the regions Ω_o and Ω_u in model (2)).

Example: Fig. 7 shows one of the experiments described above for our field case. From the figure we see that the input-output gains from u_1 to Δ_{21} are smaller when considering the system when operating between $t = 0$ and $t = 20$ with respect to the gains from $t = 20$ to $t = 70$. This is a clear indication that the system dynamics transitions from one regime to another. Then, from an automatic control perspective there is the need to describe the system using two different models. These two models will therefore be called $\mathcal{M}_o$ and $\mathcal{M}_u$ and they will transition around the operating conditions experienced by the system around $t = 20$.
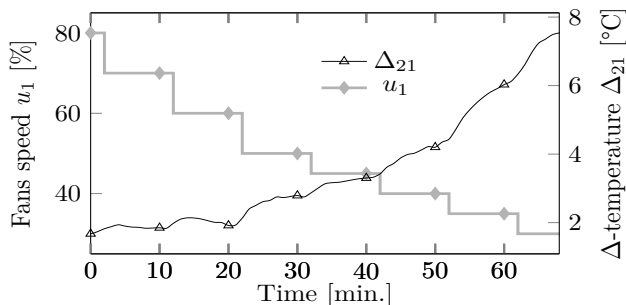

Figure 7. Comparison of the temporal evolutions of the temperature difference Δ_{21} against the CRAH fans speed signal u_1. Δ_{21} corresponds to a set of step responses.

F. Identifying and validating the models using data-driven approaches

Once the regions Ω_o and Ω_u in model (2) have been determined, it is possible to identify the various transfer functions for the various regimes in (2) using classical PEM identification methods. To this aim, it is necessary to extract, from the datasets determined in Sec. IV-C, those parts that are relative to the various regions Ω_o and Ω_u.

Example: applying classical LTI system identification strategies, we can at this point learn three distinct models: $\mathcal{M}_o$ and $\mathcal{M}_u$, which correspond to the models described before, and $\mathcal{M}_j$, which is a model that does not assume that there exist two different and distinct regimes in the system, but rather assumes that the behavior of the system is uniform across all the potential operating conditions.

Figures 8, 9 and 10 instead graphically describe the simulation capabilities of the different models in reproducing some test sets (i.e., data that has not been used during the model learning steps). It is important to notice how model $\mathcal{M}_j$ totally fails in capturing the dynamics of the system in all its potential regimes, while the other two models $\mathcal{M}_o$ and $\mathcal{M}_u$ have better approximation capabilities. To improve the prediction capability of $\mathcal{M}_u$ a future work can be to subdivide Figures 9 in two or more areas with their respective linear models.

V. Conclusions

We considered how to construct data-driven models for both detecting when the cooling flows from the CRAH units mix with the exhaust air flows, and for understanding how to operate the cooling infrastructure so that this type of events does not happen.

The proposed methodology combines two main ingredients:

- first, it considers the peculiarities of the physics of the phenomena under consideration: indeed the identification of the regions of the operating conditions that lead to either flows overprovisioning or underprovisioning is based on a input-output gains analysis. In turn, this analysis is inspired by the consideration that the input-output gains of the system heavily depend on the type of flow mixing that occur within the computer room;
- moreover, we apply classical PEM approaches to the system identification problem, which are well-established learning strategies for building the models from the collected measurements once the different mixing regions have been identified.

The methodology has been then tested and assessed on a full scale field case. We compared two sets of models: *i)* $\mathcal{M}_o$ and $\mathcal{M}_u$, i.e., the models that should describe the thermal behavior of the system under respectively flow overprovisioning and underprovisioning regimes; and *ii)* $\mathcal{M}_j$, i.e., a model that was learned assuming that the system does not experience changes of regime.

The results in Figures 8, 9 and 10 clearly indicate that the identified $\mathcal{M}_o$ and $\mathcal{M}_u$ have better generalization capabilities than $\mathcal{M}_j$, suggesting thus that our assumption that the system experiences two distinct and rather different flow regimes was correct.
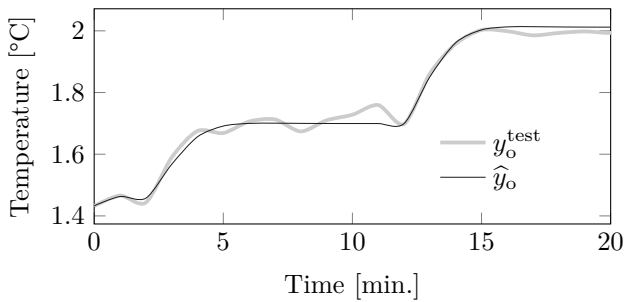
Figure 8. Experiment for the validation of the identified model $\mathcal{M}_\text{o}$ (in our case having an ARIX structure with orders 2 and 2). First section of the experiment, between $t = 0$ and $t = 20$.
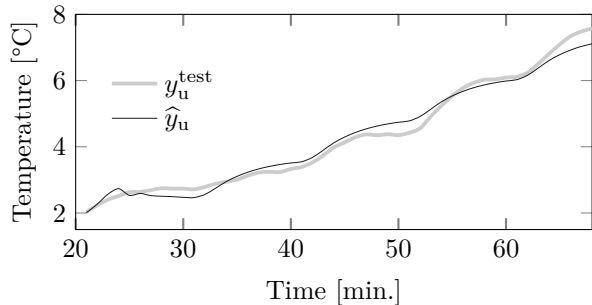

Figure 9. Experiment for the validation of the identified model $\mathcal{M}_\text{u}$ (in our case having again an ARIX structure with orders 4 and 4). Second section of the experiment, between $t = 20$ and $t = 70$.
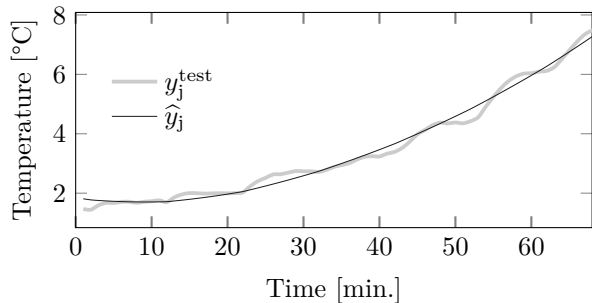

Figure 10. Experiment for the validation of the identified model $\mathcal{M}_\text{j}$(in our case having a BJ structure with orders 1, 1, 1 and 1). Complete experiment, between $t = 0$ and $t = 70$.

The here presented methodology opens up some potential new research directions: first, deriving suitable model-predictive CRAH control strategies that exploit these models. Moreover, here we limited ourselves to the case of two distinct flows regimes. It is however meaningful to assume that there may be more regimes; there is the need then for procedures that also automatically identify the number of different regimes that the system experiences.

TABLE I
FIT OF THE VALIDATION TESTS IN FIGURES ON THE LEFT SIDE

$\mathcal{M}_\text{o}$	$\mathcal{M}_\text{u}$	$\mathcal{M}_\text{j}$
89%	88%	86%

REFERENCES

[1] M. Dayarathna, Y. Wen, and R. Fan, "Data center energy consumption modeling: A survey," *IEEE Communications Surveys & Tutorials*, vol. 18, no. 1, pp. 732–794, 2016.
[2] L. Parolini, "Models and control strategies for data center energy efficiency," Ph.D. dissertation, Carnegie Mellon University Pittsburgh, PA, 2012.
[3] C. M., "The EU's many data center projects," 16-Feb-2015.
[4] W. Garcia-Gabin, E. Berglund, H. Zhang, K. Mishchenko, and X. Zhang, "A data center model for testing control and optimization algorithms," in *IECON 2017 - 43rd Annual Conference of the IEEE Industrial Electronics Society*, 2017, pp. 7165–7170.
[5] R. Lucchese, J. Olsson, A.-L. Ljung, W. Garcia-Gabin, and D. Varagnolo, "Energy savings in data centers: A framework for modelling and control of servers' cooling," *IFAC World Congress*, 2017.
[6] Z. Song, B. T. Murray, and B. Sammakia, "A dynamic compact thermal model for data center analysis and control using the zonal method and artificial neural networks," *Applied thermal engineering*, vol. 62, no. 1, pp. 48–57, 2014.
[7] R. Zhou, Z. Wang, C. E. Bash, A. McReynolds, C. Hoover, R. Shih, N. Kumari, and R. K. Sharma, "A holistic and optimal approach for data center cooling management," in *American Control Conference (ACC), 2011*. IEEE, 2011, pp. 1346–1351.
[8] E. Wibron, A.-L. Ljung, and T. S. Lundström, "Computational fluid dynamics modeling and validating experiments of airflow in a data center," *Energies*, vol. 11, no. 3, p. 644, 2018.
[9] A. Beghi, G. D. Mana, M. Lionello, M. Rampazzo, and E. Sisti, "Energy-efficient operation of an indirect adiabatic cooling system for data centers," in *The 2017 American Control Conference*. The 2017 American Control Conference, May 24 2017.
[10] A. Beghi, L. Cecchinato, G. D. Mana, M. Lionello, M. Rampazzo, and E. Sisti, "Modelling and control of a free cooling system for data centers," *Energy Procedia*, 2017.
[11] Y. Berezovskaya, A. Mousavi, V. Vyatkin, X. Zhang, and T. B. Minde, "Improvement of energy efficiency in data centers via flexible humidity control," in *Industrial Electronics Society, IECON 2016-42nd Annual Conference of the IEEE*. IEEE, 2016, pp. 5585–5590.
[12] C. Flinta, A. Johnsson, J. Ahmed, F. Moradi, R. Pasquini, and R. Stadler, "Real-time resource prediction engine for cloud management," in *Integrated Network and Service Management (IM), 2017 IFIP/IEEE Symposium on*, 2017, pp. 877–878.
[13] J. Ahmed, A. Johnsson, F. Moradi, R. Pasquini, C. Flinta, and R. Stadler, "Online approach to performance fault localization for cloud and datacenter services," in *Integrated Network and Service Management (IM), 2017 IFIP/IEEE Symposium on*, 2017, pp. 873–874.
[14] "Swedish institute of computer science, infrastructure and cloud datacenter test environment," https://ice.sics.se/.