Predicting Temperature and Differential Pressure in Data Centers Using Physical Modeling

Lisa Rivalin[1], *Andrew* Grier[1], *Tobias* Tiecke[1], *Chi* Zhou[1], *Doris* Gao[1], *Prakriti* Choudhury[1] *and John* Fabian[1]

[1]Meta Platforms, Infrastructure department, Menlo Park, California

Abstract. Data centers account for 1% of global energy consumption, and optimizing their energy use is a priority in the industry. Since cooling contributes several tens of percent of the total energy consumption in many data centers, the industry is moving towards very energy-efficient data center buildings. An effective cooling strategy is the containment strategy, which consists of separating the cold air going into the servers (supply air) and the hot air going out of the servers (return air), creating physically secluded cold and hot aisles. One important source of parasitic heat in data halls is the hot-air recirculation, flowing back from the hot to the cold aisle at the rack level and having multiple negative consequences, such as the creation of hot spots damaging the hardware and increasing the data center cooling needs. To prevent this phenomenon and improve the data centers' energy efficiency, the ability to predictively model conditions and events, even those not seen before, in data centers is increasingly important. We have identified three main reasons leading to recirculation in a Data Hall: baseline Recirculation characterized by design Hot Aisle Containment and racks leakage, recirculation caused by negative differential pressure locally at the rack, and flow deficit at the end of an aisle when there is not enough supply air to cool the racks. We developed physical-law-based models to predict thermal conditions in a data hall accurately and quantify the three types of recirculation airflow. This paper presents an application of these models in an actual data center in the United States with a successful prediction of the thermal behaviour within 1°F of MAE (mean absolute error) for cold aisles in normal conditions. It achieves more accurate predictions than a data-science-only-based model in under-provisioned conditions.

1 Introduction

As the demand for a digitized world is exploding, efficient operations of data centers are a core tool in limiting the energy demands of the growing fleet [1]. Meta has worked to reduce the power consumed in cooling our data centers, achieving 9% overhead, though the industry standard is 50% energy overhead on server power for cooling of the building [2]. Data center containment is a technique that reduces the energy used to cool the servers by separating the cold room-level supply air from the hot exhaust air from IT equipment at the server level [3]. Previous studies have shown the high efficiency of this strategy [4][5]. In this paper, we consider a hot aisle containment configuration in which the airflow path is the following: the cold air is supplied at the room level from the supply shaft and diffused through the rows into the front side of the racks - this is the cold aisle. The servers heat the airflow and then exhaust it in the hot aisle at a high temperature. The hot air is secluded in the hot aisle by containment walls. We call "recirculation" the heat remaining in the cold aisle despite the hot aisle containment.

The server inlet temperatures depend on the ratio between the cold air supply and server airflow [6]. When the total cooling supply airflow entering the data hall is lower than the total airflow consumed by the servers, the rack's inlet and outlet air temperatures significantly increase [7] and may put the hardware at risk. In contrast, a higher airflow rate leads to very inefficient cooling since it amplifies the bypass of cold around and above the servers and not through it [8], leading to more operating costs and less energy-efficient data centers. The usual ratio between supply airflow and servers' airflow is usually above 1.2 to reach acceptable operating conditions [9].

To prevent recirculation and improve the data centers' energy efficiency, the ability to predictively model conditions and events, even those not seen before, in data centers is increasingly important. This paper proposes a physics-based method to assess data centers' recirculated heat to predict the data hall cold aisle temperature. The main novelties of this paper are the introduction of the "effective" recirculation heat fraction that characterizes the heat remaining in the cold aisle despite containment and the simple method to assess this fraction, with a fast-computing time to accurately

predict the cold aisle temperature both at the data hall and the rack levels with minimum input data needed.

2 Methodology

Our approach is to create a simple first-principle model with a short runtime and very few inputs and parameters to get the average cold aisle temperature within a data hall. We have built a steady-state single-zone numerical model in Modelica language using the Modelica Buildings Library [10], where the inputs are the supplied and exhausted airflow, the supply air temperature, and the effective recirculation heat ratio (Fig 1). The "effective" recirculated heat is the portion of heat remaining in the cold aisle, calculated as the fraction of the heat generated by the servers (IT load) required to raise the total air supplied to the cold aisles to the temperature measured at the cold aisle temperature sensors. We are not modeling the exact airflow rate flowing back from the hot to the cold aisle; hence we say 'effective.' This parasitic heat is most likely caused by air flowing back from the hot aisle to the cold aisle but can also be caused by other sources of unwanted heat, such as radiation or leakages from the air recycling plenum above the data hall that is redirecting the air exhausted by the servers back to the economizer system.

Our modeling approach is to gather all the cold aisles in one single volume of air of the same dimension as the totality of the data hall, including the hot and cold aisles, and to assume the volume of air is well-mixed. We consider the totality of the airflow entering the volume is also leaving this volume to avoid over or under pressure in the data center. Our goal in this paper is to find a way to characterize this parasitic heat with very few data measurements and without using fitting methods. The "recirculation ratio" that is calculated is then multiplied by the IT load. When the recirculation ratio is equal to 0, the containment and insulation of the building are perfect. When the recirculation ratio is 100%, all the heat in the data hall is well mixed, and there is no containment at all.

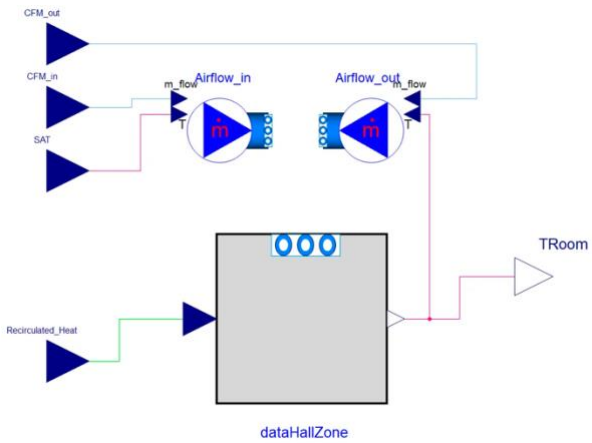

Fig. 1. Data hall model estimating the room temperature using the recirculated heat, supply air temperature, and supply, and exhaust airflows as inputs

A similar logic is applied to the row level to get the temperature at the center of a cold aisle for each rack position.

We have identified three main reasons leading to recirculation in a Data Hall (Fig 2):

- Baseline (or background) Recirculation: characterized by the designed hot Aisle Containment and racks leakage and possible other sources of radiated or unwanted leakage,
- Differential-pressure induced recirculation: recirculation caused by negative differential pressure locally at the rack,
- Flow deficit recirculation at the end of an aisle when there is not enough supply air to cool the racks

$$R_{total} = R_{Baseline} + R_{DP} + R_{flow} \quad (1)$$

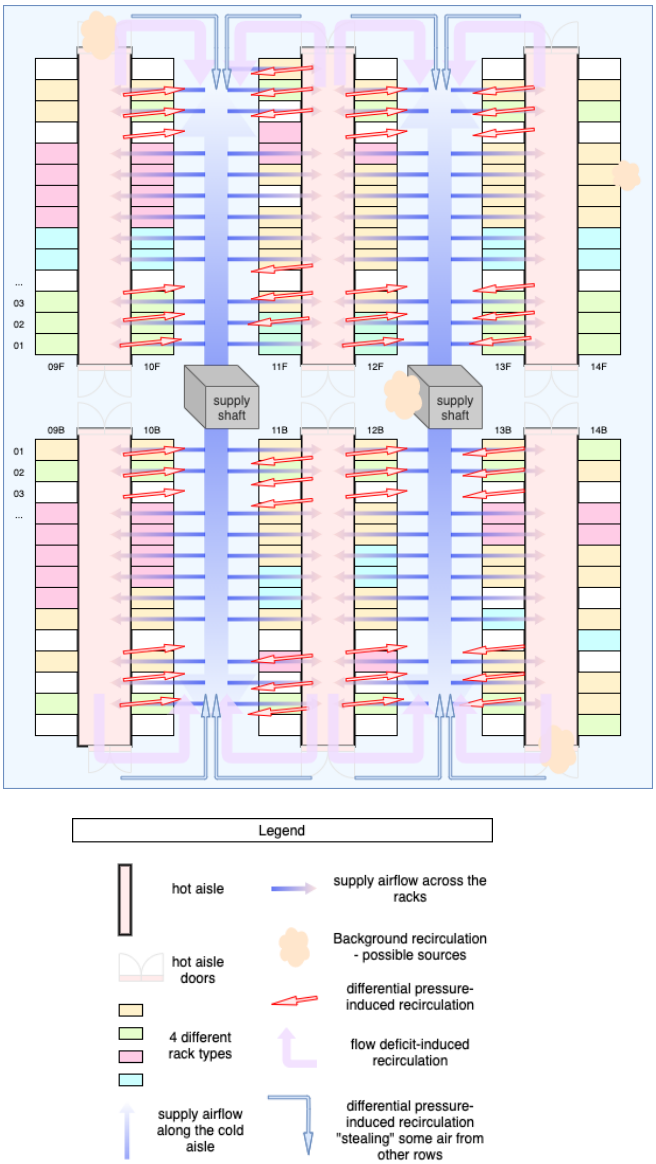

Fig. 2. Schematic of the airflow on a hot aisle contained data center

2.1 Baseline recirculation

We call baseline recirculation the minimum amount of heat remaining in the cold aisle even when the supply and total server demand ratio is above 1.2. The hot aisle

containment has been designed to meet a requirement of 5% of leakage. The containment system has additional sources of leakages, for instance, when the servers need maintenance or when the maintenance team opens the doors to access the aisles. The baseline recirculation is specific to each data center because it translates the particular installation and possible gaps between containment walls, the exact setup of the racks, and the operation of the data hall. Therefore, the baseline recirculation ratio is estimated with measured operational data. To do so, we have created a thermal model in Modelica language that can calculate the average data hall zone cold aisle temperature as a function of the supply airflow entering and leaving the room, the supply air temperature, and the IT load. The recirculated heat is defined as a percent of the IT load remaining in the server rooms. The measured average cold aisle temperature is then compared to the simulated one in normal operating conditions by varying the recirculation percent. We then select the recirculation percent with the lower mean average error.

2.2 Differential-Pressure-induced Recirculation

For any rack position where the differential pressure between cold and hot aisles is negative (i.e., the cold aisle pressure is lower than the one in the hot aisle), the flow is naturally coiling back through the rack gaps. The total amount of air due to this effect is accounted for as the differential-pressure-induced recirculation.
The initial velocity at the shaft is estimated using the shaft dimensions and airflows, according to [11].

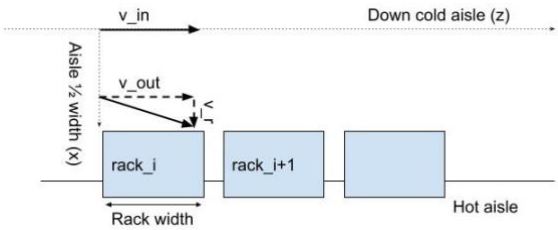

Fig. 3. Velocities along half of a cold aisle

For rack position i, we assume some flow coming into the slice from the left-hand side with forward velocity v_in in the middle of the aisle. The fans in $rack\_i$ pull some air through them into the hot aisle behind the racks with an average velocity v_r in the volume of the cold aisle being pulled towards the rack. The remaining airflow v_out passes to the next slice as the center aisle velocity. With this model flow field in hand, we calculate the pressure profile by using momentum conservation as described by the time-averaged momentum Navier-Stokes equation below:

$$\frac{\partial \rho V_i}{\partial t} + \frac{\partial \rho V_i V_j}{\partial x_j} = -\frac{\partial p}{\partial x_i} + \frac{\partial}{\partial x_j}\left[\mu_{eff}\left(\frac{\partial V_i}{\partial x_j} + \frac{\partial V_j}{\partial x_i}\right)\right] + \rho\beta(T_0 - T)g_i \quad (2)$$

where

$x_{i,j}$ are the coordinates
ρ is the air density
V_j is the velocity component in x_j – direction
p is the pressure
μ_{eff} is the effective viscosity
β is the thermal expansion coefficient of air
T_0 is the thermal expansion coefficient of air
T is the temperature
g is the gravity acceleration

We are interested in the z-dependence of pressure, so we focus on the $i = z$ case. We assume no variation in the flow field along the y-direction (vertical; assumption confirmed by numerical estimate of buoyancy term) and that the solution is steady-state. Next, we eliminate the term arising from viscosity by noting that $\mu_{eff} \leq 0.01$ in SI units for the conditions of the datahall, while $\rho = 1.225$ in SI units, providing a 100x suppression of that term relative to the others which is not overcome by the extra derivative in that term for these conditions. Finally, we are left with

$$\frac{\partial}{\partial z}p \approx -\rho\frac{\partial}{\partial x}(v_z v_x) - \rho\frac{\partial}{\partial z}(v_z^2) \quad (3)$$

We numerically approximate under our assumed flow field with

$$\frac{\Delta p_i}{w_{rack}} \approx -\rho\frac{(v_{out,i}v_{rack,i})}{w_{aisle}} - \rho\frac{(v_{out,i}^2 - v_{in,i}^2)}{w_{rack}} \quad (4)$$

where

Δp_i (Pa) is the pressure change across rack position i along the cold aisle,
w_{rack} and w_{aisle} are the width of a rack and the half-width of cold aisle, respectively
ρ is the density of air,
$v_{in,i}$ is the air velocity entering rack position i,
$v_{rack,i}$ is the air velocity flowing through the rack at position i into the hot aisle,
$v_{out,i}$ is the velocity of the air leaving rack position i due to the missing flow going through the rack.

Both terms on the RHS can be thought of as static pressure changes due to differing bulk flow conditions (i.e. dynamic pressures) for a constant density gas at constant temperature.

This formula allows us to compute the pressure change along the cold aisle due to the flow out through the racks at position i with knowledge of the cold aisle dimension, the rack width, the flow entering the beginning of the cold aisle, and the flow being pushed into the hot aisle at each rack position (sum of the rack on the left and the right of the cold aisle at that position). We then combine the pressure profile along the row with the pressure spike at the beginning of the cold aisle to obtain the final pressure profile.

The previous process leads us to a relative differential pressure profile down the cold aisle. An important last step of the process is to offset this profile to make sure we are reflecting the actual profile in the data center. To do so, we use the unique differential pressure sensor per cold aisle that is available in our data centers (usually around the middle of the row).

Finally, the last step of the calculation is to convert the calculated pressure profile into recirculated air flow from the hot aisle to the cold aisle. For this, we need a model of the conductance of the racks. To assess the recirculated heat flow, we use two inputs: a model of the flow through a standard rack face, and a model of flow through the opening on the side of the racks near the back due to the lack of side panels. Both these models have been built and validated with measured data collected during in-situ experiments.

For any rack position with a negative differential pressure, we apply the following logic: if a rack is populated at that position, we apply the pressure model to account for recirculation through the face. If that rack is adjacent to an empty rack position, we apply the crack leakage conductance once for each empty adjacent rack position, i.e., a crack at the edge of each group of racks. If there are two populated racks next to one another, we do not compute flow into the cold aisle through the side-of-rack gaps between the two racks considering (this has been verified with CFD studies) that the two racks close off effectively that volume from the cold aisle.

2.3 Flow deficit Recirculation

For cases when the data hall is being supplied with less airflow than the racks are calling for, we argue on mass-balance terms that the missing airflow must come from recirculation. In these severe and obviously unwanted cases, we add in an additional recirculation airflow term which is the difference between the total supply airflow and the servers' airflow.

3 Results and discussion

3.1 Baseline recirculation estimation

The baseline recirculation is calculated as described in 2.1. The result is expected to be different from one data center to another because it is highly dependent on the geometry and physical installation of the racks and containment walls of the data halls. It translates the minimum effective recirculation rate that we can achieve at best in a data center without any structural changes to the insulation or containment walls. Fig 4 shows the result for one modeled data center. The baseline recirculation is estimated to be around 12.5% with a mean average error of 0.7 F for temperatures around 65°F. This study corroborates previous measurement tests that was conducted in our data centers for a specific type of racks.

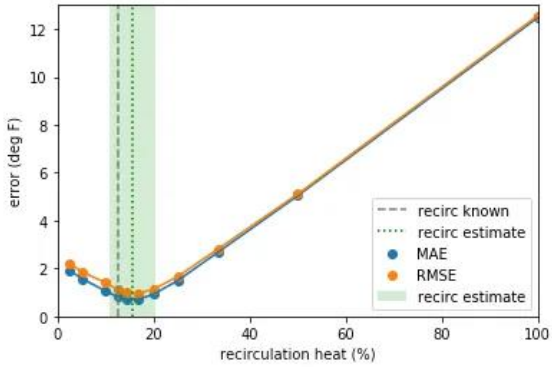

Fig. 4. Error in Fahrenheit between the simulated and measured cold aisle temperature as a function of the recirculation heat ratio. The green dotted line/area shows the fitted value +/- 1 standard deviation.

The baseline recirculation ratio is a useful metric to characterize the performance of the containment installation or to identify data centers with a higher maintenance rate that needs specific attention. The fact that the baseline recirculation ratio is around 12.5% for this data hall explains why the ratio between supply and demand is around 1.2. Typically, the baseline recirculation ratio is the only metric that would be used in a data-science-based model since usually no data are available to describe the thermal behavior of a data hall under several under provisioned conditions.

3.2 Differential Pressure induced Recirculation Estimation

The model described in 2.2 is applied for several data halls. Fig. 5 shows characteristic distributions of total recirculation for an ensemble of racks estimated due to differential pressure, one for typical running conditions with a small overabundance of supply air relative to servers' airflow, and an undersupply case during a test period where the supply to server airflow ratio was intentionally lowered slightly below parity:

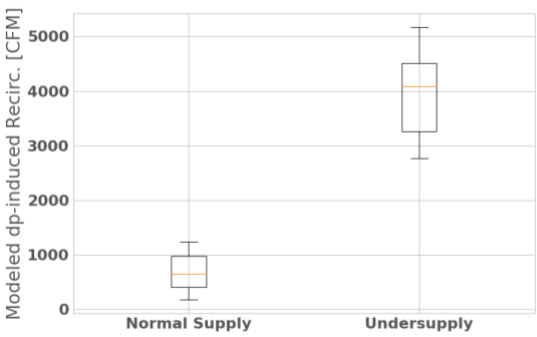

Fig. 5. dp-induced recirculation flow totals per aisle in normal case and severe undersupply conditions. Means shown as the red bar, Q3 and Q1 as box limits, and p5 and p95 as whiskers.

We note that dp-induced recirculation is predicted to exist by this model even in standard or over supplied conditions. In the most severe undersupply cases, this effect can grow to be equivalent to > 15% of the total supply airflow from a few % under standard conditions..

The data centers are equipped with a single differential-pressure sensor in the cold aisle to measure against the hot aisle static pressure. To validate the pressure profile against sensor data, we have deployed test differential-pressure sensors at 6 additional rack positions.

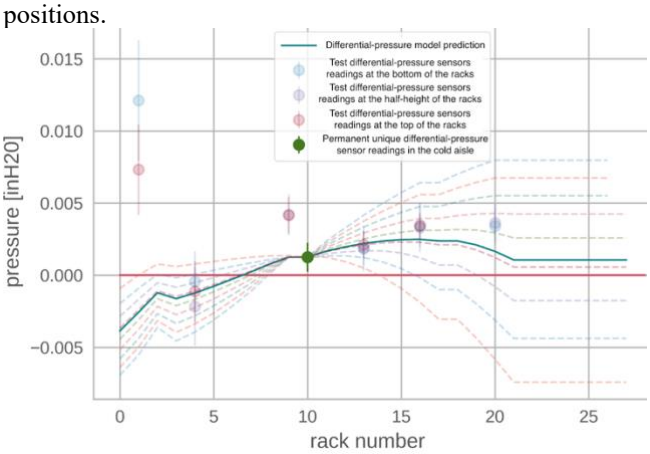

Fig. 6. Pressure profile predictions and comparison with sensor data. Solid line is the prediction from the model using measured input data and dashed lines show the range of predictions for the expected range of input data.

Although the overall simulated form of the pressure profile is consistent with measurements, we observe a static pressure rise at the first rack position not captured in the modelled profile. This is due to a 90° redirection of the airflow flowing from the shaft located on the ceiling to the aisles at the head of the rows that our current model, focused on z-direction motion, does not capture.

3.3 Flow deficit Recirculation

This case occurs in severe conditions where the supply airflow added to the differential pressure-induced recirculation airflow is less than the server's airflow demand. Even though this is impossible in normal operating conditions - we have not encountered such a case in stable data center historical data, it is possible and seems to occur in normal operating conditions for highly populated rows, meaning that a single row can be in this regime while the other rows function normally. In that case, the sum of airflows pulled by the servers in that aisle is larger than the supplied airflow provided by the supply shaft at this aisle.

We handle this case in our dp-induced term by allowing for the surplus supply airflow from other aisles to enter an individual aisle through the backside of the aisle. In those cases, the airflow down the cold aisle changes direction at a rack position and begins to increase. This increase in flow up the cold aisle means the static pressure of the cold aisle begins to decrease again as the airflow moves towards the end of the row, leading to more differential pressure-induced recirculated airflow if the local pressure goes below the one of the hot aisles (Fig 8).

3.4 Temperature prediction at the data hall level

We used the model described in part 2 to predict the average cold aisle temperature. The effective recirculation radio input of the model has been calculated following eq (1) where each term is calculated according to the methods described part 2.1 to 2.3.

We compare this approach to a model that would only consider the baseline recirculation term, usually what data-science only based models typically predict, since this factor has been calculated with measured data, largely available in a well provisioned data center. The goal of this comparison is to validate why the analytic portion of the model add values to the data science-based only prediction accuracy.

Both models were compared to measured data during a test conducted in data halls where the supply airflow has been drastically reduced to 90% of the server airflows for several hours (Fig.9). The cold aisle temperature sensors have a 0.5°C uncertainty.

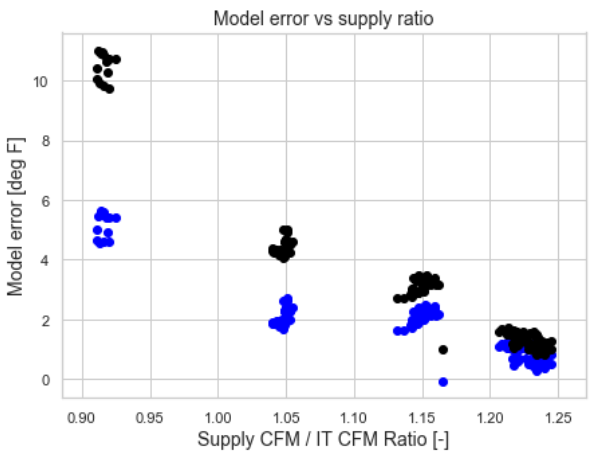

Fig. 9. Measured cold aisle temperature vs over(under)supply ratio for both analytical and historical data inference-based model for a sample of timespans with stable temperatures during the period including the stress test in a data hall. Blue is the analytical model and black is the inference-based model

We can see that both models successfully predict the average cold aisle temperature of the data hall under well provisioned conditions, with less than 1 F of mean absolute error or about 1% of relative error. The analytical model presents an acceptable accuracy: 5.1°F of mean absolute error in very undersupplied conditions, capturing at least half of the missing effect from the model ignoring the dp-induced recirculation.

Fig. 10 presents the breakdown of recirculation terms for three supply and server airflows' ratio categories: normal operating conditions (above 1.2), undersupplied (between 1 and 1.2) and severe undersupplied (less than 1). Fig. 10 confirms that the recirculation phenomenon is accentuated in undersupplied conditions.

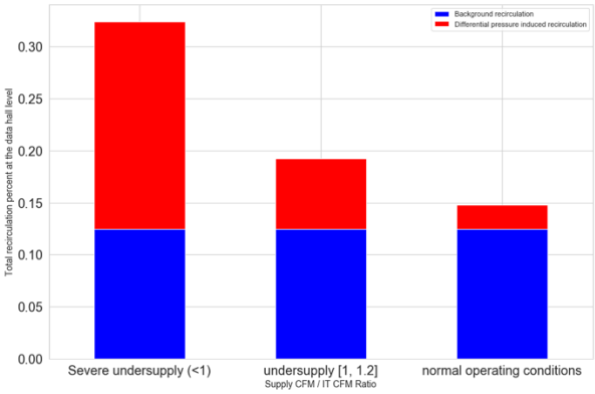

Fig. 10: Recirculation breakdown as a function of the supply ratio category

3.5 Temperature prediction at the row hall level

The row-level temperature is estimated with the same approach as the one applied to the room level, but to predict the average temperature in between two racks for each rack position. Note that we consider the baseline recirculation to be evenly distributed along the aisle. We stop our temperature profile prediction at the point along the cold aisle where flow of supply air changes direction from coming from the front of the row to coming up the back of the row due to the equations having more unknown than inputs.

Fig. 11 shows the result of our prediction both in normal and severe undersupplied conditions on the same cold aisle.

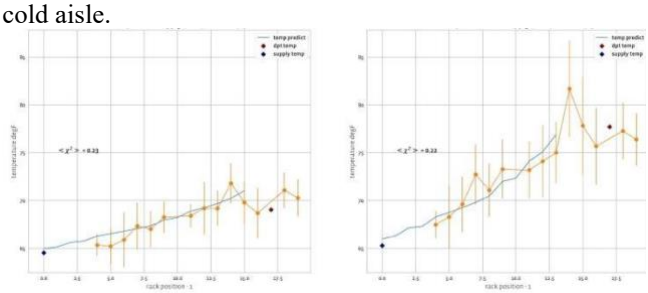

Fig. 11: Temperature profile prediction for each rack position for a test aisle for normal conditions (left) and undersupplied conditions (right)

The predicted slope of the temperature as a function of the rack position is higher for the undersupply data, this reflects the recirculation for each rack position.

If we consider all the rows of a data hall, this model's accuracy is 4.7°F in normal conditions and 6.3°F in undersupply conditions. However, the server inlet temperature sensor have a 2.5°C uncertainty which leads us to some reservations to the use of MAE as a right metric to assess this model, for that reason we use the χ^2 metric, defined as follows:

$$\chi^2 = \frac{(prediction - data)^2}{Uncertainty\ in\ the\ data}$$

χ^2 is a measure of the difference between the model and data normalized by the uncertainty in the data. The model shows a χ^2 of 2.15 in normal conditions and 2.9 in undersupply conditions, which is satisfying for our prediction needs.

4 Conclusion and next steps

This paper presents a simple, first-order analytical model with very few inputs and parameters to assess the average cold aisle temperature at the data hall level. This model successfully predicted the cold aisle temperature of a data hall with less than 1F of mean absolute error for normal conditions and 5F for severe undersupply conditions during a test period. To achieve these results, we considered the effective recirculation ratio as the portion of IT load remaining in the cold aisle despite the containment and broke it down into three terms: baseline recirculation that describes the containment and building insulation leaks and maintenance operations in the data hall, differential pressure-induced recirculation and flow limitation recirculation.

The differential-pressure recirculation model provides a general understanding of the pressure profile along the cold aisle, which is an important input into calculating the pressure induced recirculation. The model outputs a trend along the whole row that is then offset by an actual measurement of a single rack. We have used this pressure model to accurately model the recirculation.

We then applied this model at the row level to get a prediction of the cold aisle temperature in between each rack along the data center aisles.

While our prediction at the row level is satisfying, we would like to improve our understanding of the local recirculation effects at the rack level. We would also like to understand whether we are including some local recirculation in our estimate of background recirculation, i.e. whether some of the blue in the bars in Fig 10 should actually be attributed to the red. To achieve this, a more detailed airflow model could be helpful to identify the recirculated air flow flowing back from the hot aisle to the cold aisle but may be at the expense of computing time.

References

[1] E. Masanet, A. Shehabi, N. Lei, S. Smith, & J. Koomey, Recalibrating global data center energy-use estimates in Science (2020)

[2] Meta 2021 Sustainability Report. Accessed: https://sustainability.fb.com/wp-content/uploads/2022/06/Meta-2021-Sustainability-Report.pdf (2022)

[3] J. Wan, X. Gui, S. Kasahara, Y. Zhang & R. Zhang Air Flow Measurement and Management for Improving Cooling and Energy Efficiency in Raised-Floor Data Centers: A Survey in IEEE Access, 6 (2018)

[4] C. H. Wang, Y. Y. Tsui & C. C. Wang, On cold-aisle containment of a container datacenter in Applied Thermal Engineering 112 (2017)

[5] J. Cho, C. Park & W. Choi Numerical and experimental study of air containment systems in legacy data centers focusing on thermal performance and air leakage in Case Studies in Thermal Engineering, 26 (2021)

[6] W. X. Chu & C. C. Wang A review on airflow management in data centers. In Applied Energy 240 (2019)

[7] H. Lu, Z. Zhang Numerical and experimental investigations on the thermal performance of a data center in Applied Thermal Engineering, 180 (2020).
[8] S. A. Nada, M.A. Said & M.A. Rady Numerical investigation and parametric study for thermal and energy management enhancements in data centers' buildings in Applied Thermal Engineering, 98 (2016)
[9] L. Li, C. Jin, & X. Bai A determination method on the supply airflow rate with the impact of leakage and bypass airflow of computer room air conditioning in Journal of Building Engineering 59 (2022)
[10] M. Wetter, Z. Wangda, T. S. Nouidui and X. Pang. Modelica buildings library, Journal of Building Performance Simulation 7 4 (2014)
[11] A. Rubel Computations of jet impingement on a flat surface. 16th Aerospace Sciences Meeting, Huntsville, Alabama, January 16-18, (1978)