Data driven fault tolerant thermal management of data centers

Masoud Kheradmandi
Dept. of Computing and Software
McMaster University
Hamilton, Canada
kheradm@mcmaster.ca

Douglas G. Down
Dept. of Computing and Software
McMaster University
Hamilton, Canada
downd@mcmaster.ca

***Abstract*—In this work we address the problem of fault-tolerant operation of data center thermal management systems. A data driven fault tolerant control framework is proposed. First, a data driven dynamic predictive model is trained using subspace identification. This model is utilized in a model predictive controller to predict dynamic behavior of a data center. An economic model predictive control is designed to maintain temperatures within an allowable range while minimizing operational costs of the cooling system. An observer based fault diagnosis method is then developed to detect and isolate sensor faults. When a fault occurs, it is diagnosed and the corresponding measured signal is replaced with the reconstructed signal from a state estimator. The effectiveness of the proposed method is illustrated through simulations on a mechanistic data center model.**

***Index Terms*—Data centers; Thermal management; Zonal control; MPC; EMPC; Data driven predictive model; Time series analysis; Machine learning.**

I. Introduction

Growing demand for cloud based services requires reliable operation and monitoring of data centers (DCs). These operational systems require proven capabilities in scale, speed and security to support costumer demand. At the same time, DCs are currently drawing up to 5% of the world's electricity [1], [2], with this proportion growing because of increasing demand for cloud computing infrastructures. About 40% of this energy is usually provided for thermal management of IT equipment (ITE) [3]–[5]. The aim of DC cooling systems is to maintain server temperatures within a safe temperature region, where servers can be reliably operated. The safe temperature ranges for different types of DC hardware are available in guidelines such as those of ASHRAE [6]. This can be achieved by utilizing modern and advanced control and monitoring techniques. The challenges in the closed-loop operation of DCs are compounded by uncertainties manifested as variations in demands or faults. DC performance can be negatively affected by abnormal operating conditions and faults, which can lead to undesired consequences including unsatisfactory quality of service, economic loss and hardware failure (or any combination of these).

A data center's control and monitoring framework must be capable of fault detection and isolation (FDI) and fault tolerant control (FTC). All model-based FDI frameworks consist of two main components, an underlying model to predict healthy system behavior and FDI logic to detect and isolate faults. A wide variety of methods have been proposed assuming the availability of a good model (mechanistic or data driven). Theses approaches are based on comparing measured/available data with estimated/reconstructed signals. In the presence of a fault a considerable difference typically exists between some of the available signals and reconstructed signals estimated using the model. The logical part of an FDI system utilizes the residual values to diagnose fault(s). With respect to controller design, there are several approaches employed for dynamic management of data centers.

Violation of safe temperature guidelines, even by 1 or 2 degrees, can result in server failures [7], [8]. It can also result in poor performance [9]. Therefore a controller is required that can maintain safe temperatures while minimizing operating cost. In order to control temperatures within a DC, we find that a control theoretic viewpoint is applicable.

Recently we proposed a data driven controller for data center thermal management [10]. Using a Model Predictive Control (MPC) approach, our technique solves an optimization problem at each sampling instant over a finite time horizon, subject to a dynamic model of the system and general constraints, in order to calculate the control action. The proposed framework minimizes operating costs while ensuring temperatures are within a given safe temperature zone. In order to enable fault tolerant operation the FDI system should be integrated with the controller. There are different types of data driven fault diagnosis frameworks that can be utilized. In [11] a black-box fault detection method for rack-level fault detection is proposed. In [12] and [13] the problem of workload management under faults is considered. However, the problem of fault tolerant control of a cooling unit is not addressed in these works.

Motivated by the above considerations, in this work we address the problem of designing a fault tolerant control framework for ensuring that temperature is within a safe range. To this end first a fault detection and isolation framework is proposed, which is then integrated with an economic model predictive control framework. The rest of the manuscript is organized as follows: First, the general description for the data center considered in this work, a mechanistic model for DCs as a test bed, a machine learning approach (subspace identification) and a representative formulation for Economic MPC are reviewed. Then the proposed FDI approach for sensor faults is presented and is integrated with the controller. The efficiency of the proposed FDI and FTC methods is illustrated by simulating a DC with two rack mounted cooling units. Finally, concluding remarks are presented.

II. Preliminaries

In this section, a description of a recently developed data center with rack mounted cooling units is reviewed. This is utilized as our test-bed model. A machine learning based predictive model that is used to predict the dynamic behavior of the system is described. Finally a description of the economic model predictive (EMPC) framework is presented. This material is a summary of the work presented in [10] and is included here for completeness.

A. Data Center

There are different types of cooling methods in data centers: room based, row based and rack mounted cooling units. In this work we consider a single rack data center with two rack mounted cooling units (RMCUs). The schematic for a data center IT rack within an enclosure that is cooled by two RMCUs is presented in Fig. 1.

In this work we utilize a mechanistic model as a test bed. The model was recently developed in [14]. The model development uses heat transfer and fluid mechanics principles to capture dynamic behavior of the system. A detailed explanation of the model development can be found in [14]. The final formulation is a set of differential, algebraic equations.

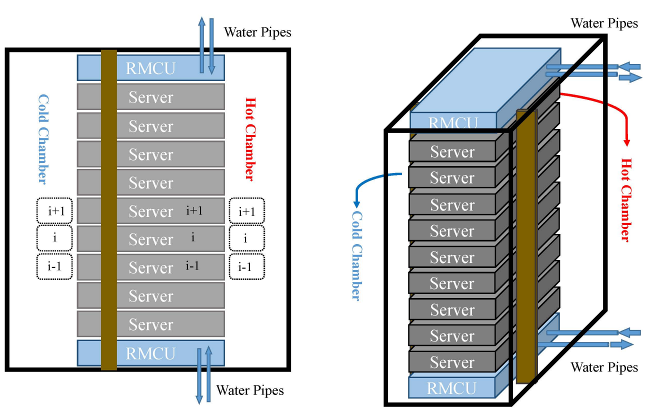

Fig. 1. Schematic of the IT enclosure integrated with a single rack and two RMCUs with separated cold and hot chambers. The zones (control volumes) in the front and back chambers are shown.

B. Dynamic Predictive Modeling

In this section we briefly review a data-driven (machine learning) method for dynamic modeling. We will utilize a subspace-based identification method to train a discrete time state space model. The goal in model identification is to compute the linear time invariant (LTI) model parameters with the following form:

$$x_{k+1} = Ax_k + Bu_k + w_k \tag{1}$$

$$y_k = Cx_k + Du_k + v_k \tag{2}$$

where $x \in \mathbb{R}^{n_x}$ and $y \in \mathbb{R}^{n_y}$ denote the vectors of state variables and measured outputs, $w \in \mathbb{R}^{n_x}$ and $v \in \mathbb{R}^{n_y}$ are zero mean, white vectors of process noise and measurement noise with the following covariance matrices:

$$E[\begin{pmatrix} w_i \\ v_j \end{pmatrix} \begin{pmatrix} w_i^T & v_j^T \end{pmatrix}] = \begin{pmatrix} Q & S \\ S^T & R \end{pmatrix} \delta_{ij} \tag{3}$$

where $Q \in \mathbb{R}^{n_x \times n_x}$, $S \in \mathbb{R}^{n_x \times n_y}$ and $R \in \mathbb{R}^{n_y \times n_y}$ are covariance matrices, and δ_{ij} is the Kronecker delta function.

In order to calculate the model parameters from given input and output data, the training procedure from [15] is adopted. A detailed description of the data-driven modeling, which involves orthogonal projection and least squares, is explained in [15] and [16].

Note that the order of the LTI system n_x is selected in a manner such that the identified system is controllable and the prediction of validation data is acceptable.

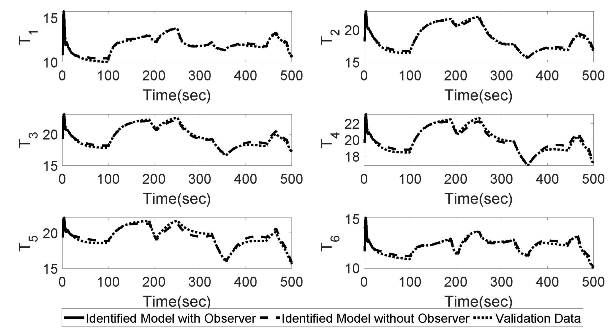

Fig. 2. Model validation results (outputs: six temperature values in front of servers)

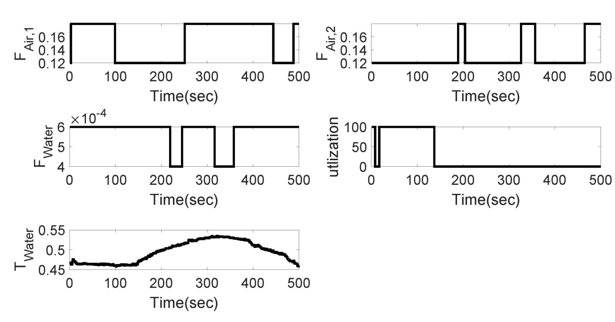

Fig. 3. Model validation results inputs and measured disturbances

Remark 1: The order of the model or the number of states to be used for the model (the only model parameter choice) is decided from properties of the training data fitting. This is done at the singular value decomposition step, where the choice of the number of states is equal to the number of dominant singular values. Note also that for the case that data are collected in closed-loop with noise or disturbances in the plant, open-loop identification methods are theoretically biased for data driven modeling, and closed-loop identification must be used, for example the closed-loop identification method utilized in [17].

C. Economic Model Predictive Control

In this section an economic model predictive controller based on the identified model is described. The goal of this controller is to keep all the temperatures in front of the servers within a desired range. The proposed controller is found as a solution of the following optimization problem, which we call the EMPC optimization problem:

$$\min_{\tilde{u}(j),\epsilon_-(j),\epsilon_+(j)} \sum_{j=1}^{P} c_{\epsilon_-}^T \epsilon_-(j) + c_{\epsilon_+}^T \epsilon_+(j) + c_u^T \tilde{u}(j) \tag{4}$$

subject to:

$$\tilde{x}(k+1) = A\tilde{x}(k) + B\tilde{u}(k) + B_d\tilde{d}(k) \tag{5}$$
$$\tilde{d}(k+1) = \tilde{d}(k) \tag{6}$$
$$\tilde{y}(k) = C\tilde{x}(k) + D\tilde{u}(k) + D_d\tilde{d}(k) \tag{7}$$
$$\tilde{u} \in \mathcal{U}, \quad \epsilon_- \geq 0, \quad \epsilon_+ \geq 0 \tag{8}$$
$$\tilde{x}(k) = \hat{x}_k, \quad \tilde{d}(k) = d_k \tag{9}$$
$$y_{min} - \epsilon_- \leq \tilde{y} \leq y_{max} - \epsilon_+, \tag{10}$$

where $\tilde{y}_k$ and $\tilde{u}_k$ are the predicted output trajectory and input at the kth sampling instant. The vectors $\tilde{x}$ and $\hat{x}$ are the predicted value and the estimate of the subspace state, obtained utilizing a state estimator. The vectors $\tilde{d}$ and d are the predicted value and the estimate of the measured disturbances, equal to the latest measurement. The vectors c_{ϵ_-} and c_{ϵ_+} denote penalties associated with slack variables. The vector c_u denotes the cost of the manipulated input variables. The slack variables ϵ_- and ϵ_+ are added to the lower and upper bounds on the outputs to ensure the feasibility of the optimization problem (slack variables). The values y_{min} and y_{max} are lower and upper bounds on the output (temperatures).

In order to compute the control action, the EMPC optimization problem is solved at each sampling instant. The first input in the computed sequence is then implemented. The controller optimization problem is a linear programming problem. Therefore, using available solvers, the global optimum (if constraints are consistent) can be calculated. A Kalman filter is employed for state estimation in this study.

III. Data-Driven Fault Diagnosis and Fault-Tolerant Control of Data Centers

A. Fault Diagnosis for Data Centers

In this section a data-driven fault detection and isolation technique for sensor faults is proposed. The method is then utilized to develop a fault tolerant control framework for data centers. In order to diagnose sensor faults consider the identified model of the system subject to the faults:

$$x_{k+1} = Ax_k + Bu_k + w_k, \tag{11}$$
$$y_k = Cx_k + Du_k + v_k + f_k, \tag{12}$$

where $f_k \in \mathbb{R}^{n_y}$ denotes the sensor fault vector and the rest of the variables are as defined earlier. In order to detect and isolate these faults a bank of observers is designed. The number of observers is equal to the number of measured outputs (n_y). The ith observer, O_i, is designed to detect faults in the sensor for the ith output. The observers have the following form:

$$O_i : \begin{cases} \hat{x}_{k+1}^{O_i} &= A\hat{x}_k^{O_i} + Bu_k + K_k^{O_i}(y_k^{[i]} - \hat{y}_k^{[i]}) \\ \hat{y}_k^{O_i} &= C\hat{x}_k^{O_i} + Du_k \end{cases}, \quad i = 1, \ldots, n_y \tag{13}$$

where $\hat{x}^{O_i}$ and $\hat{y}^{O_i}$ denote the predicted state and output calculated with the ith observer. The observer gain is calculated similar to (**??**)-(**??**). The difference is that the ith observer excludes y_i from its measured and predicted outputs. The vectors $y_k^{[i]} \in \mathbb{R}^{n_y-1}$ and $\hat{y}_k^{[i]} \in \mathbb{R}^{n_y-1}$ denote the measured and predicted outputs excluding the ith output of the system ($y_k^{[i]} := [y_{k,1}, \ldots, y_{k,i-1}, y_{k,i+1}, \ldots, y_{k,n_y}]^T$).

The ith observer uses the most recent measurement $y_k^{O_i}$ which includes all the outputs except the ith output ($y_{k,i}$) to estimate the state of the system ($x_k^{O_i}$). Then $x_k^{O_i}$ is utilized in (2) to estimate the outputs of the system $\hat{y}_k^{O_i}$. The estimated (reconstructed) output $\hat{y}_k^{O_i}$ is then compared with the corresponding measured value (y_k) to calculate the residual value of the signals. Also, note that in the update step of state estimation, the ith output is not utilized, therefore this estimation would be insensitive to a fault in the ith sensor. In the absence of a fault the residual value should be negligible. The residual values may be non-zero due to plant-model mismatch. In order to determine a threshold on the residual caused by the plant-model mismatch, training data are utilized. For each sensor, the threshold is the maximum value of the residual for each output in healthy (fault free) training data.

In order to detect sensor faults, the residuals are continuously calculated by comparing measured and reconstructed output signals. In the case that there exists a sensor fault in the ith sensor, the residual between the reconstructed signal from the ith observer for the ith output with measured value exceeds the threshold for that output. More precisely:

$$\begin{cases} r_{i,k} & > \eta_i \quad \text{There is a fault in the } i\text{th sensor} \\ r_{i,k} & \leq \eta_i \quad \text{There is not a fault in the } i\text{th sensor,} \end{cases} \tag{14}$$

where r_i denotes the residual vector and is computed as follows:

$$r_{i,k} := |\hat{y}_{i,k}^{O_i} - y_{i,k}|. \tag{15}$$

The vector η is the threshold vector.

B. Fault-Tolerant Control of Data Centers

In order to ensure safe and efficient operation, after fault detection and isolation, fault tolerant control should be utilized. In this section the developed framework is augmented with a fault tolerant mechanism.

In the case of a measurement fault in the ith sensor, the fault is diagnosed by the proposed fault detector. The measured signal for the ith sensor is replaced with the reconstructed value of the signal by the observer O_i, also the observer is replaced with the state estimator O_i. The proposed framework is illustrated in Fig. 4:

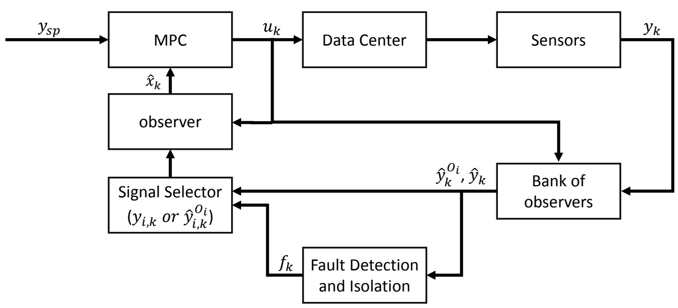

Fig. 4. Fault tolerant control framework

IV. Illustrative Simulation Results

A. Data Driven Modeling

The model is computed using the proposed identification method. The manipulated variables are two air flow rates in front of the fans and overall chilled water flow rate. The measured disturbances are the server utilizations (assumed to be uniform among servers) and chilled water temperature. The controlled outputs are the six temperatures in front of the servers.

In order to identify a model, inputs and utilizations are perturbed using pseudo-random binary sequences, except for the chilled water temperature. The chilled water is typically taken from a utility service, and its temperature is subject to changes due to the time varying load. Here, we assume a sinusoidal signal. The identified model consists of twenty states, three manipulated input variables, two measured disturbances and six measured outputs.

The identified model has the following form:

$$x_{k+1} = Ax_k + Bu_k + B_d d_k + w_k, \tag{16}$$

$$y_k = Cx_k + Du_k + D_d d_k + v_k, \tag{17}$$

$$E[\begin{pmatrix} w_i \\ v_j \end{pmatrix} \begin{pmatrix} w_i^T & v_j^T \end{pmatrix}] = \begin{pmatrix} Q & S \\ S^T & R \end{pmatrix} \delta_{ij}, \tag{18}$$

where d is vector of measured variables. The matrices B_d and D_d are gains for measured disturbances in the state and output equations. In order to use our identification approach, d is augmented with u to create the overall input, while the rest of the steps are as illustrated.

For model validation, a different training data batch was used. The model validation results are presented in Fig. 2 and Fig. 3. Since the initial state of the state space model cannot be determined from the training data, a Kalman filter (a state estimator) is utilized in order to achieve output convergence, then open-loop prediction is used for model validation in order to evaluate model prediction performance. At time 300 seconds, the output of the predictive model has converged to the plant output, and the identified LTI model in open-loop (without state update), along with the known input trajectory, is utilized for output prediction with the remainder of the data. The results show that after convergence of the model states, the prediction performance of the model for simulating the process behavior is reasonable, and is appropriate for a predictive control implementation.

B. Fault Detection under Open Loop

In order to test the proposed FDI framework a fault as a constant bias is added to temperature sensor 5. The results are presented in Fig. 5 and Fig. 6.

TABLE I
CONTROLLER PARAMETERS

Variable	Value
c_{ϵ_-}	$\begin{bmatrix}10^6 & 10^6 & 10^6\end{bmatrix}^T$
c_{ϵ_+}	$\begin{bmatrix}10^6 & 10^6 & 10^6\end{bmatrix}^T$
c_u	$\begin{bmatrix}10^{-2} & 10^{-2} & 1\end{bmatrix}^T$
P	8
y_{max}	16

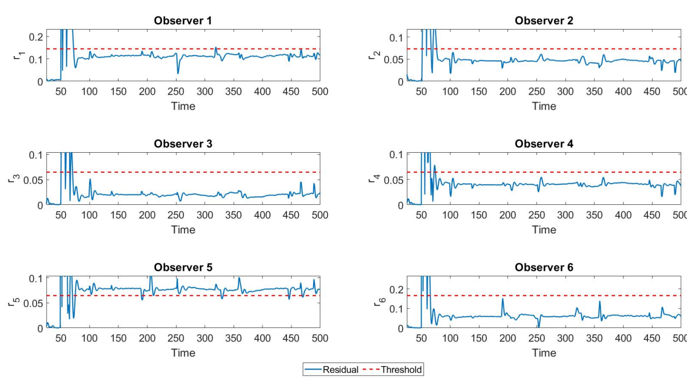

Fig. 5. Residuals of observers under sensor fault of sensor 5

C. Fault Tolerant control

The closed-loop trajectory with the proposed fault tolerant control given a fault in sensor number 4 (the hottest point) is simulated and is compared with the system without added fault tolerance. The fault tolerant control framework, after diagnosing the fault, steers the system back to the required zone. The results are presented in Fig. 7 and Fig. 8.

Remark 2: Although a data center with rack mounted cooling units is used to demonstrate our approach, the proposed framework can be used for other types of cooling units and for any size of DC. We have not made any assumptions for

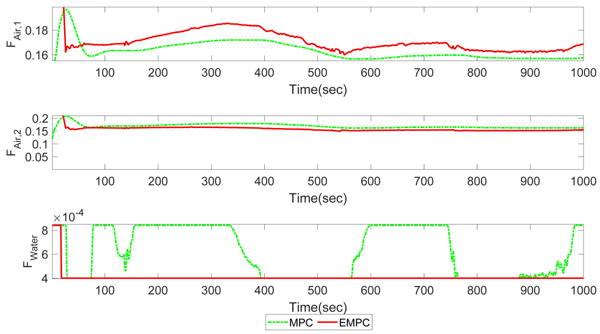

Fig. 6. Closed-loop profile (input)

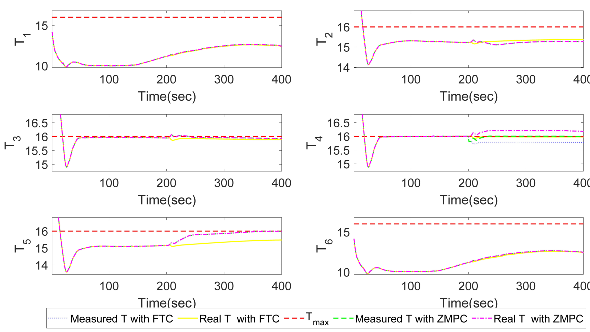

Fig. 7. Comparison of closed-loop trajectory between FTC and controller (Outputs)

the structure of the DC system and a data driven model is used in the predictive controller, so there is no reason that this approach should not scale well.

Remark 3: Our results assumed a negative bias in a sensor, which results in the temperature exceeding the desired threshold at the corresponding zone in the data center. Note that on the other hand, a positive bias would not result in threshold violation, but it would increase the operation cost of the system if FTC is not employed.

V. Conclusions

In this study, a novel fault detection and isolation method is developed for DC temperature sensor faults. In order to diagnose sensor faults a bank of observers is utilized to reconstruct measured signals and by comparing them to measured values faults are detected. The framework is integrated with a zone model predictive control approach to create a fault tolerant framework. This approach is based on replacing the signal of a faulty sensor with an estimated signal. The proposed approaches are implemented on a simulated data center and shown to be able to provide improved closed-loop behavior.

References

[1] K. C. Armel, A. Gupta, G. Shrimali, and A. Albert, "Is disaggregation the holy grail of energy efficiency? The case of electricity," *Energy Policy*, vol. 52, pp. 213–234, 2013.

[2] M. Deru, K. Field, D. Studer, K. Benne, B. Griffith, P. Torcellini, B. Liu, M. Halverson, D. Winiarski, M. Rosenberg *et al.*, "US Department of Energy commercial reference building models of the national building stock," 2011.

[3] J. Dai, M. M. Ohadi, D. Das, and M. G. Pecht, *Optimum cooling of data centers*. Springer, 2016.

[4] K. Ebrahimi, G. F. Jones, and A. S. Fleischer, "A review of data center cooling technology, operating conditions and the corresponding low-grade waste heat recovery opportunities," *Renewable and Sustainable Energy Reviews*, vol. 31, pp. 622–638, 2014.

[5] H. M. Daraghmeh and C.-C. Wang, "A review of current status of free cooling in datacenters," *Applied Thermal Engineering*, vol. 114, pp. 1224–1239, 2017.

[6] R. K. Sharma, C. E. Bash, C. D. Patel, R. J. Friedrich, and J. S. Chase, "Balance of power: Dynamic thermal management for internet data centers," *IEEE Internet Computing*, vol. 9, no. 1, pp. 42–49, 2005.

[7] N. El-Sayed, I. A. Stefanovici, G. Amvrosiadis, A. A. Hwang, and B. Schroeder, "Temperature management in data centers: why some (might) like it hot," *ACM SIGMETRICS Performance Evaluation Review*, vol. 40, no. 1, pp. 163–174, 2012.

[8] M. K. Patterson, "The effect of data center temperature on energy efficiency," in *2008 11th Intersociety Conference on Thermal and Thermomechanical Phenomena in Electronic Systems*. IEEE, 2008, pp. 1167–1174.

[9] W. Torell, K. Brown, and V. Avelar, "The unexpected impact of raising data center temperatures," *Write paper 221, Revision*, 2015.

[10] M. Kheradmandi, D. G. Down, and H. Moazamigoodarzi, "Energy-efficient data-based zonal control of temperature for data centers," 2019, submitted for publication.

[11] C. Ciccotelli, L. Aniello, F. Lombardi, L. Montanari, L. Querzoni, and R. Baldoni, "Nirvana: A non-intrusive black-box monitoring framework for rack-level fault detection," in *2015 IEEE 21st Pacific Rim International Symposium on Dependable Computing (PRDC)*. IEEE, 2015, pp. 11–20.

[12] M. Isard, "Autopilot: automatic data center management," *ACM SIGOPS Operating Systems Review*, vol. 41, no. 2, pp. 60–67, 2007.

[13] R. Niranjan Mysore, A. Pamboris, N. Farrington, N. Huang, P. Miri, S. Radhakrishnan, V. Subramanya, and A. Vahdat, "Portland: a scalable fault-tolerant layer 2 data center network fabric," in *ACM SIGCOMM Computer Communication Review*, vol. 39, no. 4. ACM, 2009, pp. 39–50.

[14] H. Moazamigoodarzi, S. Pal, S. Ghosh, and I. K. Puri, "Real-time temperature predictions in it server enclosures," *International Journal of Heat and Mass Transfer*, vol. 127, pp. 890–900, 2018.

[15] M. Kheradmandi and P. Mhaskar, "Model predictive control with closed-loop re-identification," *Computers & Chemical Engineering*, vol. 109, pp. 249–260, 2018.

[16] M. Khseradmandi and P. Mhaskar, "Adaptive model predictive batch process monitoring and control," *Industrial & Engineering Chemistry Research*, vol. 57, no. 43, pp. 14 628–14 636, 2018.

[17] M. Kheradmandi and P. Mhaskar, "Data driven economic model predictive control," *Mathematics*, vol. 6, no. 4, p. 51, 2018.

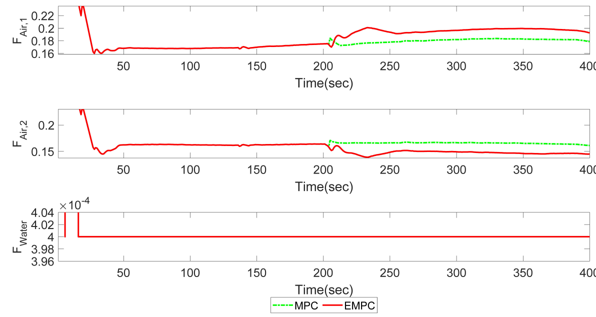

Fig. 8. Comparison of closed-loop trajectory between FTC and controller (Inputs)