

<http://dx.doi.org/10.17703/JCCT.2023.9.6.883>

JCCT 2023-11-106

## 데이터센터 냉각 시스템의 에너지 절약을 위한 인공신경망 기반 열환경 예측 모델

### Artificial Neural Network-based Thermal Environment Prediction Model for Energy Saving of Data Center Cooling Systems

임채영\*, 여채은\*\*, 안성울\*\*\*, 이상현\*\*\*\*

Chae-Young Lim\*, Chae-Eun Yeo\*\*, Seong-Yool Ahn\*\*\*, Sang-Hyun Lee\*\*\*\*

**요약** 데이터센터는 24시간 365일 IT 서비스를 제공하는 곳이기 때문에, 2030년에는 데이터센터의 전력 소비량은 약 10%로 증가될 것으로 예측되고, 고밀도 IT장비들의 도입이 점차 증가하면서, IT장비가 안정적으로 운영될 수 있도록 냉방 에너지 절감 및 이를 위한 에너지 관리가 갖춰져야 하기에 다양한 연구가 요구되고 있는 상황이다.

본 연구는 데이터센터의 에너지 절약을 위해 다음과 같은 과정을 제안한다. 데이터센터를 CFD 모델링하고, 인공지능기반 열환경 예측 모델을 제안하였으며, 실측 데이터와 예측 모델 그리고 CFD 결과를 비교하여 최종적으로 데이터 센터의 열관리 성능을 평한 결과 전처리 방식은 정규화 방식으로 사용되었고, 정규화에 따른 RCI, RTI 및 PUE의 예측 값 또한 유사한 것을 확인할 수 있다. 따라서 본 연구에서 제안하는 알고리즘으로 데이터센터에 적용될 열환경 예측 모델로 적용 및 제공할 수 있을 것으로 판단된다.

**주요어** : 데이터 센터, 에너지 절약, CFD, 열환경 예측 모델, CNN-LSTM, 열관리 성능평가

**Abstract** Since data centers are places that provide IT services 24 hours a day, 365 days a year, data center power consumption is expected to increase to approximately 10% by 2030, and the introduction of high-density IT equipment will gradually increase. In order to ensure the stable operation of IT equipment, various types of research are required to conserve energy in cooling and improve energy management.

This study proposes the following process for energy saving in data centers. We conducted CFD modeling of the data center, proposed an artificial intelligence-based thermal environment prediction model, compared actual measured data, the predicted model, and the CFD results, and finally evaluated the data center's thermal management performance. It can be seen that the predicted values of RCI, RTI, and PUE are also similar according to the normalization used in the normalization method.

Therefore, it is judged that the algorithm proposed in this study can be applied and provided as a thermal environment prediction model applied to data centers.

**Key words** : Data Center, Energy Saving, CFD, Thermal Environment Prediction Model, CNN-LSTM, Thermal Management Performance Evaluation

\*정회원, 고등기술연구원 에너지환경IT융합그룹 선임연구원 (제1저자) Received: October 2, 2023 / Revised: October 20, 2023

\*\*정회원, 고등기술연구원 에너지환경IT융합그룹 연구원 (참여저자) Accepted: November 10, 2023

\*\*\*정회원, 고등기술연구원 에너지환경IT융합그룹 책임연구원 (참여저자)\*\*\*\*Corresponding Author: leesang64@honam.ac.kr

\*\*\*\*정회원, 호남대학교 컴퓨터공학과 부교수 (교신저자) Dept. of Computer Engineering, Honam University, Korea

접수일: 2023년 10월 2일, 수정완료일: 2023년 10월 20일

게재확정일: 2023년 11월 10일

## I. 서 론

데이터센터는 고밀도 에너지 다소비 건물군에 속하며, 2030년에는 데이터센터의 전력 소비량은 약 10%로 증가될 것으로 예측되고 있기에 데이터센터의 에너지 소비량 절감은 필수적이다. 데이터센터의 전체 전력소비량 중 약 50%를 차지하고 있는 부문은 냉각 시스템으로, 냉방 에너지 절감을 위한 에너지 관리 연구가 활발히 진행되고 있다 [1][2].

또한  $CO_2$  배출 증가량은 IT분야에서 가장 높은 것으로 조사되었다 [3]. 데이터센터의 에너지 소비 비율은 냉각 시스템 50%, IT 장비 26%로 구성되며 IT 장비보다 냉각 시스템에 사용되는 에너지 비율이 높다 [4]. 최근 데이터센터 내 냉각시스템의 에너지 절감을 위한 에너지 관리 및 냉각 시스템 제어 등의 연구가 국내외로 활발히 진행되고 있다 [5].

데이터센터의 냉방 에너지 절감을 위해 데이터센터 CFD 모델링을 통해 에너지 절감 연구가 존재하였지만, 복잡한 수식에 의해 설계된다. 근래 국내외로 Deep Learning, Machine Learning 및 Fuzzy Control 등 인공지능영역을 활용한 냉방 에너지 절감 시스템 연구는 활발히 진행되고 있으며 점차 증가하는 추세로 나타났다. 인공지능영역 모델은 복잡한 수식에 의한 설계가 필요하지 않고 데이터 기반 모델로써 간편하게 예측을 실시할 수 있다는 장점이 있다 [6][7].

그리고 CFD 모델을 포함한 인공지능영역기반의 예측모델 방식은 학습에 따라 높은 정확도를 확보하는 장점으로, 예측 정확도가 높은 성능을 발휘할 수 있다.

따라서 본 연구에서는 데이터센터의 에너지 절감을 위해 다음과 같은 과정으로 진행하고자 한다. 데이터센터를 CFD 모델링을 하고, 인공지능기반 열환경 예측 모델을 제안하였으며, 실측 데이터와 열환경 예측 모델 및 CFD 결과를 비교하여 최종적으로 데이터 센터의 열관리 성능을 평가하였다.

## II. 관련연구

### 1. 인터넷 데이터 센터(Internet Data Center, IDC)

데이터센터협의회는 IDC를 ICT 장비를 집적시키고 통합·관리 가능한 솔루션, 시설, 인프라를 구축하여 24시간 365일 정상적인 기능을 제공하며 무중단으로 운영

하는 시설로 정의하고 있다 [8].

### 2. IDC 열 환경 기준

데이터센터에서 열 환경 조건은 랙 유입 공기의 온도로 데이터센터 실내 온도를 의미하지만, 열복도와 냉복도를 격리 시킬 경우에는 냉복도의 온도가 열 환경 조건에 해당된다.

### 3. 인공지능 기반 예측 알고리즘

#### 1) 가우시안 프로세스 회귀 알고리즘(GPR)

GPR(Gaussian Process Regression)는 Covariance Function, Mean Function, Gaussian Distribution 및 Bayesian Probability에 기반을 둔 지도학습 알고리즘으로 Non-Parameter 추론을 구현한다 [9]. GPR는 결정된 모델의 가중치에 의하여 새로운 입출력 데이터에 대해 결과를 예측하며 예측값에 대한 변화율을 확인할 수 있는 장점이 있다.

#### 2) 서포트 벡터 머신(SVM)

SVM(Support Vector Machine)은 선형 및 비선형 분류와 회귀 등에 사용되는 다목적 ML(Machine Learning) 알고리즘이다 [10]. Support Vector와 Margin을 이용하여 최대 마진 및 서포트 벡터를 찾아가는 과정은 Karush Kunh Tucker 조건과 Lagrange방법을 사용하여 완성된다 [11]. 서포트 벡터 머신은 주어진 문제의 정의 및 형태에 대해 추론이나 견해가 반영되기에 모델 성능이 크게 좌우될 수 있다 [12].

#### 3) CNN-LSTM 혼합복합 모델

CNN(Convolutional Neural Network)은 합성곱 신경망으로서 최근 데이터마이닝의 한 종류로서 사용되는 기계 학습 방법으로, CNN은 입력 데이터를 학습하고 패턴을 분석하며 특징을 추출하는 Feature Extraction 역할을 수행한다 [13]. CNN Layer는 Max Pooling Layer를 통해 최대값을 리턴하여 데이터의 공간적인 특징을 유지하면서 크기를 줄여준다 [14].

LSTM(Long Short Term Memory)은 기존의 RNN의 Vanishing Gradient Problem을 해결하기 위해 장/단기 기억을 가능하도록 설계한 신경망의 구조로 CNN-LSTM 혼합복합 모델을 제안하기 위하여 활용한다.

### III. 인공지능 예측 모델링 설계 및 구현

#### 1. 시뮬레이션 CFD 모델링

본 연구의 CFD 시뮬레이션 모델링은 그림 1과같은 기류 분석을 기반으로 진행하였으며, CFD 프로그램과 컴퓨터 프로그래밍 언어로 실시간 분석을 위해 통합 환경 구축하였다. 시뮬레이션의 3D 그래픽은 OpenGL을 이용하였으며, 유동해석 솔버의 경우 OpenFOAM(Version 0.7)으로 채택하였고, 실제 데이터 센터의 실내 환경을 모사하기 위하여 서버의 발열량(Q, [kcal/hr]), 풍량(m, [kg/hr]), 공기비열(c, [kcal/kg°C]), 냉복도 및 열복도의 온도차( $\Delta T$ , [°C])를 이용하여 풍량을 식 1과 같이 산출하였고, 후 처리는 ParaView를 사용하였다.

$$Q = m_{air} c \Delta T \quad (\text{식 1})$$

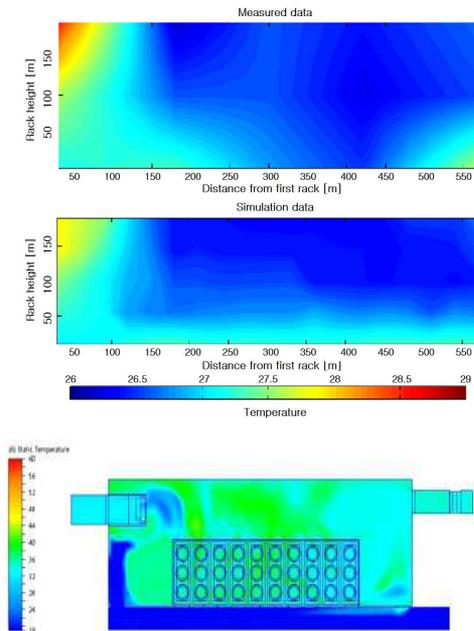


그림 1. 데이터 센터 및 랙 구성에 따른 CFD 기류 구성  
 Figure 1. Configure CFD airflow with data center and rack configurations

#### 2. 열환경 예측 모델 알고리즘 설계

본 연구에서 제안하는 모델 개발을 위한 데이터의 획득은 열역학 법칙을 활용하여 데이터센터의 수직적 열환경 모델을 이용하였다. IT 부하는 Rack 당 최대 6.6kW의 부분부하로, 4시간 간격으로 구간별 20% ~

50%로 변경하도록 설정하고, 급기온도 설정점은 1시간 간격으로 18°C ~ 27°C 사이로 설정하였다.

#### 1) 데이터셋 구성 및 전처리 알고리즘

데이터셋 구성을 하기위해 데이터 전처리 과정을 수행한다. 데이터 전처리 방식은 데이터셋을 정규 분포에 근사하게 가공하는 표준화(Standardization)방식과 정규화(Normalization)방식이 있으며, 최적 선정을 위하여 두 가지 전처리 방식에 따른 초기 예측모델의 성능을 비교분석 하였고, 전처리 방식 선정을 위하여 두 가지 전처리 방식에 따른 GPR, SVM의 초기 예측모델의 성능을 비교분석하였다. 알고리즘 최적화 단계에는 가우시안 과정(Gaussian Process)에 따른 오차 추정을 기반한 확률론적 알고리즘인 베이지안 최적화(Bayesian Optimization) 알고리즘이 사용하였고, 알고리즘으로는 Levenberg-Marquardt Algorithm이 사용하여 예측 모델 구성한다.

#### 2) 열환경 예측 모델 설계

예측모델을 비교하기 위하여 GPR, SVM, CNN, LSTM(Long Short-Term Memory)으로 설정하였다. ANN 구조 모델로 GPR과 SVM 모델은 입출력 변수, 데이터 전처리, 학습 데이터 설정 등은 초기 Default로 표 1과 같이 설정하였다. 본 연구에서는 복잡한 데이터셋을 효과적으로 구성할 수 있도록 CNN과 RNN 구조의 LSTM 알고리즘을 Flatten 함수를 활용하여 CNN-LSTM 모델을 제안하였다.

표 1. 예측 모델의 초기 설정  
 Table 1. Initial settings of predictive model

학습알고리즘	Hyperparameters	Value
GPR	Sigma	6.77
	Basis Function	pure quadratic
	Kernel Function	Squared Exponential
SVM	Epsilon	1.17
	Box Constraint	11.7
	Kernel Scale	1

GPR의 Hyper-Parameters 설정시 예측값의 표준 편차가 작을 때는 예측 구간이 매우 협소하고, 신호에 잡음이 포함되어 있을 때는 예측 구간이 그림 2와 같이 넓어진다. 본 연구의 제안하는 열환경 예측모델의 최적

화 변수 뉴런수는 20 개, 은닉층은 4층으로 구성하였다.

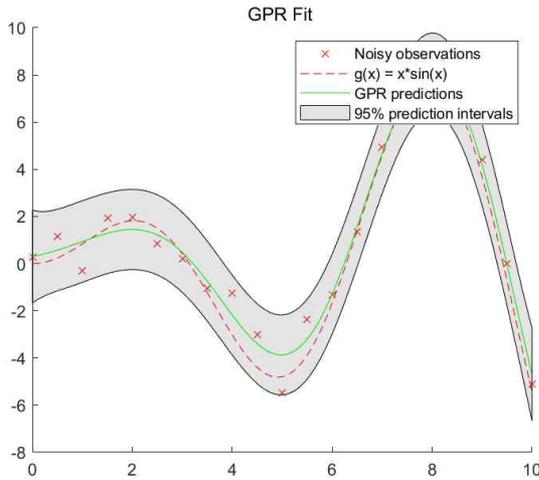


그림 2. 가우시안 프로세스 회귀 알고리즘 모델  
Figure 2. Gaussian Process Regression Models

제한한 LSTM은 Cell state와 Forget gate, Input gate, Output gate로 구성하며, Forget gate는 Cell state로부터 정보의 보존 여부를 Sigmoid layer에 의해 결정하고, 입력하는 값에 환경변수를 추가하여 학습시키는 양을 증가시켰다. 추가한 환경 변수의 경우 스케일을 동일하게 구성했다. 데이터셋은 수식적 모델을 포함하여 총 2,254개의 데이터 세트들로 구성하여, 6:2:2 비율로 Train, Validation, Test 분류하였다. CNN은 합성곱 레이어와 Max Pooling을 이용하여 구성하였으며, 활성화 함수는 ReLU값으로 지정하였다. 시간의 흐름에 따른 정확도를 향상시키기 위해 Moving Window 알고리즘을 적용하였다.

#### IV. 실험결과 및 성능평가

##### 1. RTI(Return Temperature Index) 예측결과

본 연구에서 제한하는 알고리즘에서는 온도는 0~120으로 설정하여 데이터셋을 구성하였으며, RTI의 예측값은 107.89±4.95%이고 CFD의 RTI 값은 105.5%로 나타내었는데, 오차 범위를 고려한다면 예측값과 CFD 결과가 유사한 것을 확인할 수 있다.

공조에너지 지표 RTI(Return Temperature Index)

예측실험을 진행하였다. 그림 3은 서버랙 입출구 온도와 CRAH 입출구 온도값으로, 환기온도지수 RTI의 예측범위를 최대 105%에서 최소 78%로 평균 91.1%로 예측 정확도를 보였다.

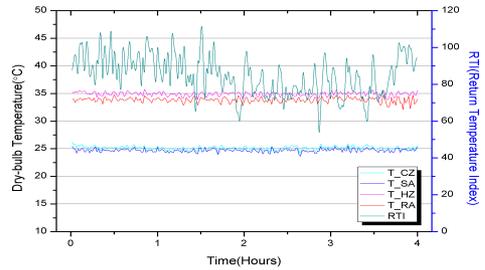


그림 3. RTI 예측 결과 그래프  
Figure 3. RTI prediction result graph

##### 2. 열관리 성능 : 온도 분포 RCI 예측결과

서버랙 전단 5cm의 위치에 상중하 온도센서 30Point를 부착하여 온도를 측정 및 분석하였고, RCI(Rack Cooling Index\_Hight) 예측결과 그림 4와 같이 95.4%로 예측되었다.

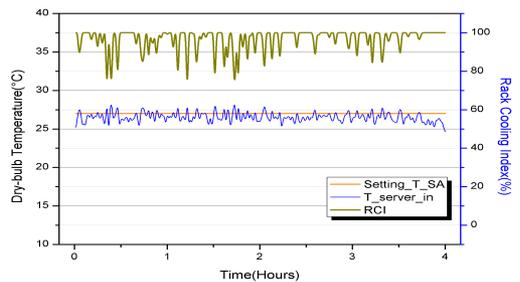
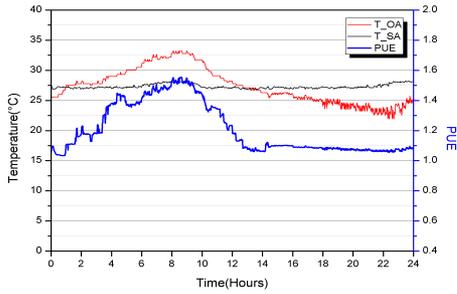


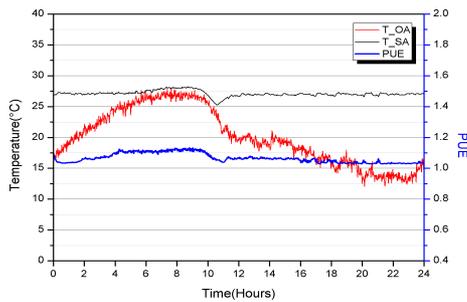
그림 4. 서버 입구 온도예측 및 RCI 예측 결과  
Figure 4. Server inlet temperature prediction and RCI prediction results

##### 3. 데이터 센터의 PUE 예측결과

데이터센터의 IT부하 200kW 운전 조건에서 실제 소비전력을 측정하여 PUE를 예측하였다. 여름철은 그림 5(a)와 같이 T\_OA의 trend로 PUE를 예측하여, 평균 1.21로 예측하였고, 환절기는 같은 운전 조건에서 실제 소비전력을 측정하여, 그림 5(b)와 같이 SA의 Trend로 학습시 PUE를 평균 1.07로 예측하였다.



a. 여름철 PUE 예측 결과 그래프  
 a. Summer season chart



b. 간절기 PUE 예측 결과 그래프  
 b. The change of seasons Chart

그림 5. 소비 전력 및 PUE 예측 결과 그래프  
 Figure 5. Power consumption and PUE prediction result graph

#### 4. 제안한 모델의 성능평가

데이터 전처리는 정규화 및 표준화에 따른 예측모델 성능을 변동계수(Cv)와 상관계수값으로 분석하여 약 1.02 이하로 기준을 만족하였으며, 데이터 전처리 방식에 따른 예측성능은 표 2와 같다.

표 2. 데이터 전처리 모델의 예측성능  
 Table 2. Prediction Performance of pre-processing Methods

Method		Norm./Stand.	Normalization/Standardization
Cv [%]	Train	1.01	0.84/0.83
	Validation	1.02	0.81/0.79
	Test	1.01	0.80/0.79

		Norm./Stand.	Normalization/Standardization
correlation coefficient	Train	0.99	16.26/16.41
	Validation	0.92	21.29/23.23
	Test	0.90	20.68/22.96

최적화의 목적함수는 예측모델에 대한 변동계수를 산출 함수로 선정하여 열환경 예측모델의 최적 변동계수를 산출하였다. 본 연구의 제안하는 열환경 예측모델의 최적 구조는 변동계수(Cv) 0.85일때 뉴런수는 20 개, 은닉층은 4층으로 구성한다.

## V. 결론 및 향후과제

본 연구에서는 데이터센터를 CFD 해석을 통하여, 데이터센터의 열 관리 성능을 평가하였고, 인공지능 기반 열환경 최적 예측 모델을 제안 및 분석하였으며, 최종적으로 이 모델링과 실측 데이터 및 CFD 해석을 비교하여 데이터 센터의 열관리 성능을 평가하였다. 열환경 예측모델에 대한 결과는 다음과 같다. 초기 예측모델의 구조로 데이터 전처리 방식은 성능평가를 통해 정규화 방식이 표준화 방식 대비 변동계수는 0.84, 81, 80로 각각 1~3%p 높게 산출되었다. 정규화에 따른 상관계수는 0.85로 도출되었으며, 예측 모델의 최적 뉴런은 20 Neuron으로 구성시에 열환경 예측 모델에 적합성을 갖는 것으로 판단되었다.

본 연구에서의 온도 및 RTI 지표등을 통해 예측 결과 및 CFD 결과를 비교하여 CFD 모델링 결과와 실측 결과를 바탕으로 열 관리 성능에 대하여 평가 및 예측하였다. RCI 및 PUE 예측값 또한 유사한 것을 확인할 수 있다. 따라서 본 연구에서 제안하는 알고리즘으로 데이터센터에 적용될 열환경 예측 모델로 적용 및 제공할 수 있을 것으로 판단된다. 향후, 데이터 센터의 에너지 효율성을 향상 하기위해 더 다양한 예측 모델을 연구하여 냉각 제어 시스템에 적용하고자 한다.

## References

- [1] Research and Markets, "Data Center Colocation Market - Growth, Trends, and Forecast (2020-2025)", <https://www.businesswire.com/news/home/20200525005100/en/Global-Data-Center-Colocation-Market-2020-2025>, 2020.11.13.
- [2] A.S.G. Andrae & T. Edler, On Global Electricity Usage of Communication Technology: Trends to 2030, challenges, 6, pp.117-157, 2015
- [3] Whitehead, B. & Andrews, D., Shah, A., & Maidment, G. Assessing the Environmental impact of data centres. Part 1: Background, energy use and metrics, Building and Environment, Vol. 82, 151-159. 2014
- [4] Y.J. Choi et al., Development of Supply Air Temperature Prediction Model for Optimal Control Algorithm of Containment Data Center, KIEAE Journal 20(5), pp.159-164, 2020.
- [5] J. Nathan Kutz, Data-Driven Modelling & Scientific Computation : Methods for Complex Systems & Big Data, OXFORD, 2013
- [6] C.Y. Lim, C.E Yeo, S.Y. Ahn, M.O Lee & H.J Sung "Design and Performance Evaluation of Digital Twin Prototype Based on Biomass Plant" The Journal of the Convergence on Culture Technology (JCCT) 9, no.5 935-940, 2023
- [7] C.Y. Lim, C.E Yeo, W.J. Cho, J.H Gu & S.H. Lee "Design and Implementation of IEC 62541 based Industry Internet of Things Simulator for Meta-Factory" The Journal of the Convergence on Culture Technology (JCCT) 9, no.3 789-795, 2023
- [8] J. D. Yang & X. C. W. Liu, Self-tuning PID-type Fuzzy Adaptive Control for CRAC in Datacenters, International Conference on Computer and Computing Technologies in Agriculture, 2014.
- [9] Y.J. Kim & C.S. Park, Gaussian Process Model for Real-Time Optimal Control of Chiller System, Journal of the Architectural Institute of Korea, vol.30, no.7, pp.211-220, 2014. DOI : 10.5659/JAIK\_PD.2014.30.7.211
- [10] Aurellen Geron & Hands-On Machine Learning with Scikit-Learn & TensorFlow, Oreilly
- [11] J.A. Suykens & J. Vandewalle, Least squares support vector machine classifiers, Neural Processing letters, 9(3), pp.293-300, 1999.
- [12] M. Martin, On-line support vector machine regression, In European Conference on Machine Learning, Springer, Berlin, Heidelberg, 2430, pp.282-294, 2002.
- [13] Zhao, B., Lu, H., Chen, S., Liu, J & Wu, D., "Convolutional neural networks for time series classification", Journal of Systems Engineering and Electronics, 28(1), 162-169, 2017
- [14] Kim, T.Y., & Cho, S.B., "Predicting residential energy consumption using CNN-LSTM neural networks", Energy, 182, 72-81, 2019.

※ This work was supported by Institute of Information & communications Technology Planning & Evaluation (IITP) grant funded by the Korea government(MSIT)(No,RS-2023-0023374 5, Development and performance evaluation of data center cooling system using liquefied gas cooling heat)