



## 컨테이너먼트형 데이터센터 최적 제어 알고리즘을 위한 열환경 예측모델 개발

### Development of Supply Air Temperature Prediction Model for Optimal Control Algorithm of Containment Data Center

최영재\* · 박보람\*\* · 조지현\*\*\* · 문진우\*\*\*\*

Young Jae Choi\* · Bo Rang Park\*\* · Ji Hyeon Cho\*\*\* · Jin Woo Moon\*\*\*\*

\* Main author, Graduate Student, School of Architecture and Building Science, Chung-Ang University, South Korea (chyoungwo@gmail.com)

\*\* Coauthor, Graduate Student, School of Architecture and Building Science, Chung-Ang University, South Korea (pbr\_1123@naver.com)

\*\*\* Coauthor, Graduate Student, School of Architecture and Building Science, Chung-Ang University, South Korea (selmainger326@gmail.com)

\*\*\*\* Corresponding author, Professor, School of Architecture and Building Science, Chung-Ang University, South Korea (gilerbert73@cau.ac.kr)

#### ABSTRACT

**Purpose:** This study aimed at developing a temperature prediction model for a containment data center. The predictive model must be guaranteed with stability and accuracy in order to be used for real-time control. Therefore, statistical evaluation was conducted to verify the prediction performance of the proposed model. **Method:** The predictive models were developed using four representative machine learning algorithms. A thermodynamic based containment data center and cooling system were modeled by MATLAB & Simulink software. The initial and optimized models were evaluated by  $R^2$  and Cv(RMSE), and the model with the highest performance was applied to the simulation. **Result:** In the initial models, RF and ANN presented highest accuracy on  $R^2$  (0.89) and Cv(RMSE) (17.85%), respectively. After the optimization, ANN presented the best prediction performance on both  $R^2$  (0.99) and Cv(RMSE) (0.94%). The result supports the accuracy and stability of the ANN model to be used for real-time control, and based on which the optimal control algorithm will be developed on further study.

#### KEYWORD

데이터센터  
기계학습  
최적 제어알고리즘

Data Center  
Machine Learning  
Optimal Control Algorithm

#### ACCEPTANCE INFO

Received Sep. 28, 2020

Final revision received Oct. 6, 2020

Accepted Oct. 12, 2020

© 2020. KIEAE all rights reserved.

## 1. 서론

최근 활발한 기술개발이 이루어지고 있는 인공지능, 자율주행, 5G 산업 등은 빅데이터(Big Data) 구축과 원활한 데이터 처리 및 송수신이 필수적이다. 데이터 규모와 사용처가 확대됨에 따라 이를 전문적으로 저장 및 처리할 수 있는 데이터센터에 대한 수요 또한 증가하고 있으며, 세계 데이터센터 시장 규모는 2019년 3억 1,400만 달러에서 2025년 5억 8,300만 달러로 증가할 전망이다[1].

데이터센터에서 소비하는 전력량은 2010년 전 세계 전기에너지 소비량의 약 1%를 차지하였으며, 2030년에는 3~13%까지 확대될 것으로 예측되었다[2]. 에너지소비량 절감을 위해 각국은 그린데이터센터에 대한 표준 및 등급을 배포하였으며 국내 또한 2012년부터 인증제도를 도입하여 시행 중이다. 데이터센터 에너지 효율 측정 기준은 PUE (Power Usage Effectiveness)로 표현된다. PUE는 총 전력 사용량 대비 IT 장비 전력사용량으로, 1에 가까울수록 에너지 효율이 높은 것을 의미한다. 국내 데이터센터 평균 PUE는 2.66으로 해외 선진국 평균 PUE 1.2에 비해 낮은 실정이며[3], 에너지 효율 향상을 위해서는 데이터센터 전력사용량의 약 50.5%를 차지하는 냉방 시스템의 에너지 절감이 필수적이다.

데이터센터는 안정적인 IT 장비 운용을 위해 365일 24시간 냉방

시스템을 가동한다. 따라서 냉방 시스템의 종류, 효율 및 운영방식은 에너지소비량에 큰 영향을 미친다. 선행 연구 조사결과 국내의 경우 외기냉방시스템 도입 및 시스템 운영방식에 따른 에너지소비량 분석에 관한 연구가 주를 이루고 있었다. Cho (2009)는 공조파트یشن 시스템 적용을 통해 급기온도를 상향함으로써 연간 약 4%의 에너지 소비량을 절감할 수 있음을 확인하였으며, 중간기 및 동절기에 외기 냉수냉방과 외기냉방 시스템 도입을 통해 일반 냉동 시스템 대비 각각 16.6%, 42.2%의 에너지 성능향상을 이루었다[4]. Kim, et al. (2013)은 시뮬레이션을 통해 CRAC(Computer Room Air Conditioner), 외기냉방 VAV, 증발냉각, 고온수 냉각 및 제습 증발 냉각 시스템을 적용하여 PUE를 비교하였다. 그 결과 CRAC, 외기 냉방 VAV, 증발냉각, 고온수 냉각 및 제습증발냉각 시스템 순으로 각각 1.8, 1.5, 1.4, 1.3의 PUE가 산출되었다. 하지만 고온수 냉각수는 랙(Rack)의 표면에 약 60°C의 고온수가 공급되는 시스템으로써 물의 누출로 인한 안전성의 위험이 있다[5]. Kim, et al. (2019)은 중앙냉수시스템에서 외기냉방모드 적용 시 CRAH (Computer Room Air Handling Unit) 유닛의 설정온도, 풍량 변화에 따른 MLC (Mechanical Load Component)를 비교 분석하였다. 그 결과 설정온도를 상승시킬 경우 외기냉방모드 가동 시간이 늘어나고 CRAH의 풍량을 감소시킬 경우 팬 소비 전력이 감소하여 에너지가 절감되었다[6].

인공지능을 통한 소프트웨어적 접근 시도 또한 이루어지고 있으

나 아직은 기초적인 연구에 머무르고 있는 실정이다. Song(2019)은 자체 e-IoT 플랫폼을 통해 취득한 데이터를 활용하여 머신러닝(Machine Learning) 기반 냉방 시스템 자동제어 솔루션 개발에 대한 기초계획을 수립하였다. 예측모델은 설정온도에 따른 항온항습기의 전력량을 예측하여 권장 설정온도를 도출하며, 이를 시스템에 적용해 약 30% 이상의 냉방 에너지 절감을 달성할 수 있을 것으로 전망하고 있다[7]. 구체적 적용에 의한 결과도출은 이루어지지 않았으나 최신 기술 도입을 시도했다는 사실에 의의가 있다.

반면 국외의 경우 시스템 구성 및 효율에 대한 연구보다는 DCIM(Data Center Infrastructure Management) 레벨에서 소프트웨어적인 접근으로 에너지를 절감하는 연구가 활발히 진행되고 있었다. 전산실 서버는 주 발열원이자 시스템 측면에서는 외란으로 간주된다. CRAC/H는 해당 발열을 제거하기 위해 부분부하에 대응할 수 있어야 하며 주 제어 대상은 풍량, 냉수 온도, 냉수 유량이다. 일반적으로 냉수 온도, 실내 설정온도를 상승시키면 시스템 가동률이 감소하여 에너지가 절감되는 것으로 알려져 있으나, 이는 팬 가동률을 증가시킬 수 있어 전체적인 에너지소비량은 상승할 수 있다[8]. 따라서 에너지 절감을 위한 시스템 요소들의 최적화 과정이 반드시 필요하며, 이를 위해 MPC(Model Predictive Control) 기반의 제어 방법 및 최적화에 관한 연구가 다수 진행되고 있다. 연구 초기에는 수학적 모델링을 통한 MPC가 주를 이루었으나, 이는 변수의 다양화, 비선형적 모델링의 어려움으로 인해 전문적인 지식을 요구한다. 최근에는 모델링의 용이성 및 정확도 향상을 위해 Data-driven 방식의 모델링을 채택하고 있다. 예측 대상은 연구별로 상이하나 대부분 기계학습(Machine Learning) 기법을 활용하고 있었다. 연구 내용은 다음과 같다.

Beitelmal, et al.(2006)은 중앙냉수시스템의 정상상태 모델링을 통해 허용 오차 내에서 에너지소비량 예측모델을 구축하였으며 정유량, 변유량 모드에 에너지 소비량을 예측하여 관리자가 선택할 수 있도록 하였다[9]. Chen, et al.(2010), Fang, et al.(2019)는 서버의 프로비저닝(Provisioning) 최적화를 통해 각 랙의 발열량을 예측하고 이를 기반으로 최적 온도를 예측해 최적 제어를 실시하는 연구를 실시하였으며, Chen의 경우 IT 부문은 35%, 냉방은 15%의 에너지 절감을 달성하였다[8,10]. Wang, et al.(2011)은 인공신경망(Artificial Neural Network, ANN)을 활용하여 전산실 온도를 예측함으로써 균일한 온도 분포를 위해 프로비저닝을 실시하였고 이 또한 13.34%의 에너지를 절감하여 온도 예측적인 접근을 우선시 하더라도 에너지를 절감할 수 있음을 증명하였다[11]. Song, et al.(2013)은 유전알고리즘(Genetic Algorithm)에 ANN 온도 예측모델을 탑재하여 최적 온도를 산출하였다. Gao(2014)는 실제 데이터센터 데이터를 활용한 ANN 기반의 PUE 예측모델을 적용함으로써 오차 0.4% 이내의 높은 정확도를 가진 모델을 구축하였고 실제 상황에서 기계학습을 활용한 예측 및 제어가 에너지 절감에 기여할 수 있음을 증명하였다[12]. 이의 연장선으로 Lazić, et al.(2018)은 ANN의 단점을 보완하기 위해 강화학습(Reinforcement Learning)을 도입하였으며 해당 모델은 실제 데이터센터에 적용되어 몇 시간 내에 안정적인 온도 제어를 실시하였다. 팬 가동률과 냉수 펌프 개폐도를 조절하였으며 기존의 PID 제어기와 비교하여 약 9%의 냉방 시스템 비용 절감을 달성하였다[13].

기존 제어 시스템을 개선하는 방향의 연구 또한 진행되고 있었다. Bash, et al.(2006)은 센서 설치를 늘리고 이를 정량화하여 전산실 내 서버 위치에 따라 개별 제어를 하는 것만으로도 에너지소비량을 절감시킬 수 있음을 증명하였으며 이는 최근 확대 설치 중인 IoT 기반의 DCIM 만으로도 에너지 효율 향상에 기여할 수 있음을 시사한다[14]. Dend, et al.(2014)는 기존 PID 제어의 개선 방안으로 PID-type Fuzzy Adaptive control 모델을 개발하였다. 해당 모델은 수동으로 조절해야 하는 PID 계수를 자가보정하는 모델로 기존 모델 대비 적은 오버슈팅과 짧은 과도응답을 거쳐 안정적으로 운영이 가능하였다[15]. 이 외에도 외기냉수시스템 스케줄, 냉각수 회수 온도 예측, 데이터센터 모델링 기법 등에 대한 연구가 다방면으로 진행되고 있다[16-19].

이상 선행 연구 분석에 따르면 국내 데이터센터 관련 연구는 대부분 시스템 효율 향상 및 외기냉방시스템에 치중되어 있는 편이며, 통합 관리 시스템 측면에서의 에너지 절감 노력은 부족한 실정이다. 반면 국외의 경우 최신 기술 도입을 통한 소프트웨어적 접근으로 에너지 절감을 시도하고 있다. 기계학습, MPC, 최적화 알고리즘 등을 적극 활용하고 있는 추세이며 에너지와 관련된 여러 인자들을 동시에 고려하여 최적 열환경 제어를 목표로 하고 있다. 이에 국내 또한 데이터센터 에너지 절감을 위하여 소프트웨어적 접근이 필수적인 것으로 사료된다. 따라서 본 연구에서는 데이터센터 최적 열환경 제어 알고리즘 개발을 위해 머신러닝 기반의 실내 온도 예측모델을 개발하고 성능평가를 통해 실시간 제어 모델로서의 적용 가능성을 확인하고자 한다.

데이터센터는 보안이 중요한 시설로서 실제 데이터 획득에 어려움이 있기 때문에 예측모델 구축 및 성능평가의 용이성을 위하여 적용 대상은 컨테이너형 데이터센터 1개 모듈과 일련의 냉방시스템으로 한정한다. 열역학에 기반하여 MATLAB & Simulink로 모델링을 실시하였으며 해당 모델은 데이터 취득 및 개발되는 예측모델의 적용 대상으로 사용된다. 예측모델은 MATLAB Library를 통해 4가지 종류의 기계학습 알고리즘으로 개발되며 최적화 및 성능평가를 통해 최고 성능을 나타내는 예측모델을 채택한다.

## 2. 기계학습 알고리즘

기계학습(Machine Learning)은 주어진 데이터를 컴퓨터 스스로 학습할 수 있게 하는 알고리즘을 의미하며 학습 방법에 따라 지도학습, 비지도학습, 강화학습으로 분류된다. 지도학습의 경우 학습 데이터셋에 정답이 포함되기 때문에 주로 분류(Classification) 및 회귀(Regression) 모델 구축에 사용된다. 학습 방법 및 목적에 따라 다양한 학습 알고리즘이 존재하며 본 연구에서는 회귀에 사용되는 대표적인 4가지 학습 알고리즘에 대한 개발을 실시한다. 각 학습 알고리즘에 대한 개념은 다음과 같다.

첫째는 가우시안 프로세스 회귀(Gaussian Process Regression, GPR) 모델이다. 가우시안 프로세스 회귀는 공분산 함수(Covariance Function), 평균 함수(Mean Function), 가우시안 분포(Gaussian Distribution) 및 베이저안 확률론(Bayesian Probability)에 기반한 지도학습 알고리즘으로 비모수적(Non-parameter) 추론을 구현한다

[20]. 가우시안 프로세스(Gaussian Process)는 비모수적 추론으로써 다변량 정규분포를 무한차원으로 확장한다. 공분산 함수에 의해 입출력 값의 공분산을 정의하고 그 값에 따라 임의의 차원을 가지는 가중치 사후분포를 결정한다. 이때 공분산 함수에 의한 가중치 사후분포는 초모수 값에 따라 달라진다. 본 연구에서는 BasisFunction, KernelFunction, Sigma의 베이지안 최적화(Bayesian Optimization)를 실시한다.

둘째, 서포트 벡터 머신(Support Vector Machine, SVM)은 서포트 벡터와 마진(Margin)을 이용하여 결정 경계를 형성하는 기계학습 알고리즘이다. 각 데이터들의 마진을 최대화 하는 방향으로 학습을 진행하여 분류에 많이 사용되나, 목표를 반대로 하면 회귀 모델로도 사용이 가능하다. 즉, 마진에 최대한 많은 데이터가 포함되도록 학습하는 식이다. 이때 마진의 정도를 결정하는 Epsilon은 초모수으로써 최적화의 대상이며 데이터 분포에 따라 선형 커널, 다항 커널, 가우시안 커널 등을 사용해 데이터 분포 특성을 반영한다. 본 연구에서는 KernelScale, BoxConstraint, Epsilon에 대한 베이지안 최적화를 실시한다.

셋째, 랜덤 포레스트(Random Forest, RF)는 의사 결정트리(Decision Tree)를 기반으로 하는 앙상블(Ensemble) 기계학습 알고리즘이다. 데이터 특성을 임의로 추출하는 배깅(Bagging) 혹은 기타 메소드를 통해 다수의 의사 결정트리를 생성하고 트리로 구성된 포레스트를 형성한다. 입력값에 대해 여러 예측값이 도출되며, 회귀의 경우 최종 예측값은 각 트리의 앙상블 시뮬레이션을 통해 예측값들의 평균으로 결정된다. 본 연구에서는 Method, Learning Rate, NumCycling, MinLeafSize, MaxNumSplits에 대한 베이지안 최적화를 실시한다.

마지막으로 인공신경망(Artificial Neural Network, ANN)은 인간의 뇌에서 일어나는 신경의 상호작용을 모방한 기계학습 알고리즘이다. 입력층, 은닉층, 출력층으로 구성되며 각 층은 노드(Node)로 구성된다. 노드는 가중치(Weight)와 편의 값(Bias)으로 구성되며 여러 노드들은 다중퍼셉트론 구조를 형성한다. 입력층에 입력되는 입력값에 따라 가중치와 편이 값은 업데이트되며 이 과정은 출력층에서 도출되는 예측값과 주어진 정답의 에러가 기준 이하를 만족할 때까지 반복된다. 본 연구에서는 베이지안 최적화를 통해 은닉층(Hidden Layer), 노드(Hidden Neuron) 수, Mu Value에 대한 최적화를 실시한다.

### 3. 컨테이너형 데이터센터 및 시스템 모델링

컨테이너형 데이터센터는 핫스팟 방지, 냉방 시스템 효율 향상 등을 목적으로 냉복도(Cold Aisle) 혹은 열복도(Hot Aisle)를 차폐한 구조를 의미한다. 냉복도 차폐 시 CRAC (Computer Room Air Conditioner)나 CRAH (Computer Room Air Handling Unit)에서 토출되는 찬 공기가 전산실의 공기와 혼합되지 않아 랙(Rack)에 인입되는 과정에서 온도 변화가 적다. 또한 열복도 차폐 시 CRAC/H에 회수되는 공기 온도가 높아 냉방 시스템 효율을 향상시킨다[21]. 본 연구에서는 MATLAB & Simulink 소프트웨어를 활용하여 냉복도와 열복도가 모두 차폐된 컨테이너형 1개와 냉방시스템에 대한 모델링을 실시하며 구성도는 Fig. 1, 과 같다.

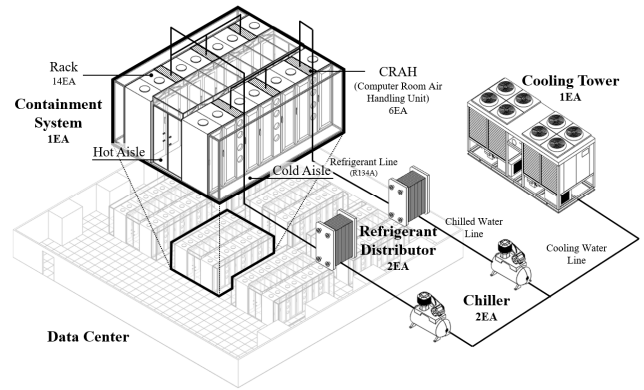


Fig. 1. Composition of Containment Data Center and Cooling System

컨테이너 내부에는 14개의 랙과 6개의 CRAH로 구성되어 있다. 각 랙의 최대 IT부하는 6.6kW이며 CRAH는 In-row 형으로 각 랙 사이에 배치된다. 컨테이너는 주변환경과 침기가 이루어지지 않고 완전 단열되는 것으로 가정하며, 모든 열교환은 100% 비율로 이루어진다. IT 부하는 kW로 정의되며 1kW = 860kcal로 환산하여 계산한다.

일반적인 랙 전후면의 온도차는 10°C~12°C이며 서버 풍량은 랙 전후면의 온도차가 11°C가 되도록 제어된다. CRAH는 랙에서 회수되는 총 풍량과 같은 풍량을 급기한다. CRAH 내부의 냉각 코일에는 R134a 냉매가 순환한다. 해당 냉매는 판형 열교환기로 인입되어 냉수와 열교환 하며 2N 구조로써 열교환기 1개가 컨테이너의 모든 냉방부하를 처리할 수 있도록 설계되었다. 각 열교환기는 냉동기에 연결되고 최종적인 냉방부하는 냉각탑에서 처리된다. 각 열매체들 간의 열교환은 열역학에 기반하여 (Eq. 1)에 의해 정의된다.

$$Q = m_a c_a (T_{aRA} - T_{aSA}) = m_b (h_{RA} - h_{SA}) \quad (\text{Eq. 1})$$

여기서,  $Q$  [kcal/hr]는 냉방부하,  $m$  [kg/hr]는 열매체의 질량 유량,  $c$  [kcal/kg°C]는 열매체의 비열,  $T_{RA}$  및  $T_{SA}$  [°C]는 열매체의 회수, 공급온도,  $h_{RA}$  및  $h_{SA}$  [kcal/kg]는 열매체의 회수, 공급 엔탈피를 의미한다.

일반적으로 데이터센터 권장 온도는 랙의 인입 온도로 측정되나, 본 연구에서는 냉복도가 차폐되어 있으며 스탠드형 CRAH의 특성을 반영하여 CRAH 공급 공기 온도와 랙 인입 공기의 온도 차  $\Delta T$ 를 0으로 가정한다. 공급 공기 온도는 냉매와 열교환을 일으키지만, 냉매는 공기와 냉수의 단순 열전달 매체로써 온도 제어에 영향이 없다. 냉수는 7°C로 공급되어 12°C로 회수되는 것으로 설정되며, 냉수 유량은 CRAH 공급 공기 온도 제어의 주체가 된다. 냉각수는 32°C로 공급되고 5°C의 레인지를 갖도록 설정하였다.

### 4. 열환경 예측모델 개발

#### 4.1. 데이터세트 구축

본 연구에서 활용되는 4가지의 기계학습은 모두 지도학습 알고리즘으로써 데이터세트 구축이 필수적이다. 예측모델의 실시간 최적 제어알고리즘 적용을 위해서는 빠른 응답시간 및 연산속도, 현장 취득이 용이한 데이터가 활용되어야 한다. 예측모델의 개발 목적은 추

후 개발 예정인 CPS (Cyber Physical System) 및 적응형 최적 제어 알고리즘에 적용되어 제어 변수 조절로 인한 실내 열환경 및 에너지 소비량을 예측하기 위함이다. 즉, CPS 및 최적 제어알고리즘은 MPC (Model Predictive Control)의 개념을 기반으로 하며, 예측모델은 Predictor의 역할을 한다. 따라서 본 연구에서 개발되는 예측모델의 출력변수는 CRAH 공급 공기 온도로 선정하였다.

입력변수는 (Eq. 1.)에서 변수 간의 관계에 기인하여 선정하였다. 냉동기와 냉각탑은 시스템 설정온도에 따라 별도로 자동제어되며, 서버의 안정성을 보장하기 위해 충분히 큰 용량으로 설계되기 때문에 냉수 및 냉각수 온도는 설정값으로 고정된다. 이에 지속적으로 변동하는 변수인 CRAH 회수 공기 온도, IT 부하, 냉수 유량으로 입력변수를 최소화하였다.

온도 범위는 ASHRAE (American Society of Heating, Refrigerating and Air-Conditioning Engineers)에서 권장하는 A1 등급 데이터센터 온도 기준인 18~27°C를 기준으로 하였다[22]. 이에 29~38°C를 CRAC 회수 공기 온도 범위로 설정하였으며, IT 부하는 최대 부하의 약 30%인 30kW부터 5kW 간격으로 상승시켰다. 냉수 유량은 4,000kg/hr를 기준으로 800kg/hr 씩 상승시키며 총 2,254개의 데이터셋트를 구축하였다. 학습 및 성능평가를 위하여 해당 데이터셋트는 MATLAB의 divideind 함수를 통해 Train (60%), Validation (20%), Test Data (20%)로 분류하였다. Train Data는 학습에 사용되고, Validation Data는 학습되지는 않지만, 과적합 방지, 성능향상을 위해 학습 과정에 활용된다. 최종 성능평가는 Test Data에 의해 진행되며 모든 데이터는 정규화(Normalization) 혹은 표준화(Standardization) 처리 후 학습 및 성능평가에 사용된다.

#### 4.2. 초기 열환경 예측모델 개발

기계학습 알고리즘은 회귀에 주로 사용되는 Gaussian Process Regression (GPR), Support Vector Machine (SVM), Random Forest (RF), Artificial Neural Network (ANN) 4가지 모델을 채택하였다. 각 모델의 성능평가는 적합도 판단의 척도인 결정 계수  $R^2$ 와 ASHRAE에서 제시하는 Cv(RMSE)를 통해 정답 CRAH 공급

공기 온도와 예측 온도를 비교하여 실시하였다. ASHRAE Guideline 14 기준  $R^2 > 0.8$ , Cv(RMSE) < 30%를 만족시켜야 예측모델로서의 타당성이 확보된다[23]. 각 학습 알고리즘에 대한 초기 하이퍼파라미터는 MATLAB에서 제공하는 Default 값으로 설정하였으며 추후 최적화 과정을 거친다. 각 학습 알고리즘에 대한 초기값 세팅 및 성능평가 결과는 Table 1.과 같다.

성능평가 결과  $R^2$ 의 경우 MinMax 정규화를 사용한 ANN 모델이 0.89로 가장 높았으며 SVM는 MinMax, 표준화 두 경우 모두 기준치인 0.8 초과를 만족시키지 못하였다. Cv(RMSE)는 MinMax 정규화가 처리된 RF모델이 17.85%로 가장 정확도가 높았다. GPR, RF, ANN 모델은 30% 미만으로 기준치를 만족하였으나 SVM는 39.1%로 정확도가 낮은 것으로 분석되었다.

### 5. 최적화 및 성능평가

#### 5.1. 예측모델 최적화

초기 예측모델은 기본값을 기반으로 개발되었으나, 학습 알고리즘은 데이터 특성, 알고리즘의 메커니즘에 따라 필요로 하는 변수가 다르며 하이퍼파라미터의 설정에 따라 성능이 달라진다. 따라서 최적 성능을 만족시키는 하이퍼파라미터의 선정은 필수적이며 최적화 결과에 따라 초기 예측모델의 성능이 개선될 수 있다. 베이지안 최적화(Bayesian Optimization)는 미지의 목적함수를 최대 혹은 최소로 만들어주는 해를 찾는 기법으로 반복 탐색을 통해 최적의 하이퍼파라미터를 산출한다[24]. 따라서 본 연구에서는 각 예측모델에 대해 베이지안 최적화를 적용하였으며 결과는 Table 2.와 같다.

$R^2$ 는 ANN, RF, GPR, SVM이 각각 0.99, 0.82, 0.77, 0.51로 ANN이 가장 높은 수치를 나타내었으며 ANN, RF, GPR은 기준치를 만족하였다. Cv(RMSE)는 ANN, GPR, SVM, RF가 각각 0.94%, 4.08%, 5.31%, 15.53%로 ANN이 가장 예측모델 정확도가 높았으며 4가지 모델 모두 기준치를 만족하였다. ANN은 은닉층 2개, 은닉 뉴런 14개로 비교적 간단한 구조로 결정됨과 동시에 연산 속도 또한 빠른 것으로 나타났다. 따라서 ANN 모델을 컨테이너형 데이터센터 모델에 적용시킨 후 성능평가를 실시하였다.

Table 1. Initial Structure of Prediction Models

Learning Algorithms	Hyperparameters	Values	$R^2$						Cv(RMSE) (%)					
			Train		Validation		Test		Train		Validation		Test	
			Min Max	Std*	Min Max	Std	Min Max	Std	Min Max	Std	Min Max	Std	Min Max	Std
GPR	Sigma	6.7743	0.89	0.90	0.84	0.84	0.84	0.85	18.42	18.25	23.26	23.11	23.50	23.12
	Basis Function	pureQuadratic												
	Kernel Function	Squared Exponential												
SVM	Epsilon	1.1739	0.56	0.77	0.55	0.58	0.55	0.08	36.01	26.98	37.82	42.01	39.10	52.03
	Box Constraint	11.7388												
	Kernel Scale	1												
RF	Method	LSBoost	0.95	0.95	0.85	0.85	0.83	0.83	13.03	18.25	18.75	23.11	17.85	23.12
	Learning Rate	1												
	NumCycling	100												
ANN	Hidden Layers	1	0.93	0.92	0.90	0.88	0.89	0.88	14.78	14.92	19.35	21.12	18.80	20.87
	Hidden Neurons	10												
	Mu	0.001												

\*Std : Standardization

### 5.2. 예측모델 성능평가

컨테이너 및 시스템 모델에서 예측모델은 피드백에 기반하여 다음 Time Step의 CRAH 공급 공기 온도를 예측한다. n번째 시뮬레이션이 진행된 후 CRAH 회수 공기 온도( $T_{RA(n)}$ ), IT 부하( $Q_{IT(n)}$ ), 냉수 유량( $m_{w(n)}$ ) 데이터는 예측모델에 입력변수로 입력된다. 예측 모델에서 출력된 n+1번째 Time Step의 CRAH 공급 공기 온도( $T_{SA(n+1)}$ )는 설정온도와 비교된다. 이 과정에서 냉수 유량은 4,000 ~ 40,000kg/hr 사이의 값을 800kg/hr 단위로 변경되어 입력되며 이에 따라 예측모델에서 출력된  $T_{SA(n+1)}$ 는 45가지의 값을 출력한다. 설정온도와 예측값의 차이는 *error*로 정의되며  $|error| < 0.1$ 를 만족시키는 최대 냉수 유량으로 n+1번째 Time Step에 대한 시뮬레이션이 진행된다. IT 부하는 실제 데이터센터 부하 발생 추이를 반영하여 약 50%의 부분부하로 난수를 생성하여 적용하였다. 설정온도는 18°C로 고정하였으며 시뮬레이션 결과는 Fig. 2와 같다.

설정온도와 예측값의 평균 오차는 0.09°C이며 최대 오차는 0.10°C로 산출되었다. 이에 따른 RMSE (Root Mean Square Error)는 0.09로 예측 정확도는 매우 높았다. IT 부하가 지속적으로 변동함에도 예측모델은 이상 값을 예측하지 않았으며 안정적으로 설정온도에 수렴하였다.

Fig. 3.은 IT 부하 변동에 따라 온도 예측값을 산출하는데 결정된 냉수 유량이다. 냉수 유량은 IT 부하와 비례하는 것으로 분석되었으며 이는 예측모델에 입력변수와 출력변수의 열역학적 관계가 잘 반영되었음을 시사한다.

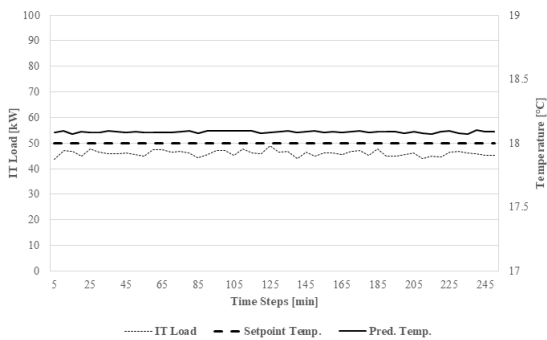


Fig. 2. Simulation Result of ANN Model

### 6. 결론

최근 급증하는 데이터 기반 산업에 대응하기 위하여 데이터센터의 수요가 급증하고 있다. 데이터센터 전력 소비량의 약 50.5%는 냉방 에너지에서 사용되고 있어 해당 부분의 에너지 절감은 필수적이다. 선행연구 분석 결과 에너지 절감을 위해 소프트웨어적인 접근이 이루어지고 있으나 국내의 경우 관련 연구 및 개발이 미비한 실정이다. 따라서 본 연구에서는 MPC 기반의 최적 열환경 제어알고리즘 구축을 위한 기초 작업으로써 열환경 예측모델을 개발하였다.

열환경 예측모델은 4가지 기계학습 알고리즘에 의해 개발되었으며 각 모델에 대한 최적화 및 성능평가를 실시하였다. 초기 모델의 정확도는 R<sup>2</sup>와 Cv(RMSE)로 평가하였으며 ANN의 R<sup>2</sup>가 0.89, RF의 Cv(RMSE)가 17.85%로 ASHRAE 기준을 상회하는 결과를 나타내었다. 각 모델에 베이지안 최적화를 실시한 후 결과는 ANN의 R<sup>2</sup>가 0.99, Cv(RMSE)가 0.94%로 가장 높은 성능을 나타내었다. 해당 최적화 모델은 컨테이너형 데이터센터 모델에 적용하여 시뮬레이션을 실시하였으며 IT 부하가 지속적으로 변동됨에도 불구하고 RMSE 0.09로 매우 높은 예측 정확도를 나타내는 것으로 분석되었다. 따라서 예측모델은 적은 수의 입력변수로도 안정적인 예측 운영이 가능하며 실시간 제어에 활용될 수 있을 것으로 판단된다.

본 연구에서는 단일 시나리오에 국한되어 성능평가가 이루어졌으나 실제 데이터센터는 외기, 냉방 시스템의 종류, IT 장비 배치 여부 등의 다양한 변수가 존재한다. 따라서 실시간 학습을 통해 최적화

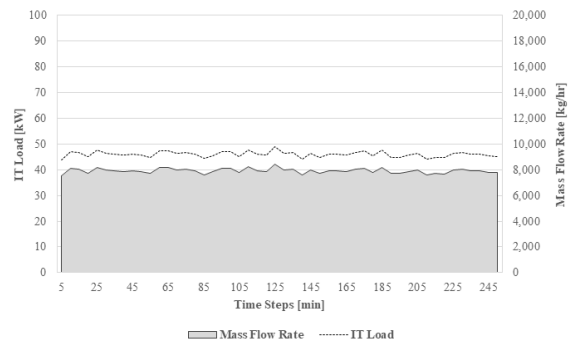


Fig. 3. IT Load & Calculated Mass Flow Rate

Table 2. Optimized Structure of Prediction Model

Learning Algorithms	Hyperparameters	Values	Preprocessing	R <sup>2</sup>			Cv(RMSE) (%)		
				Train	Validation	Test	Train	Validation	Test
GPR	Sigma	28.4274	Std	0.99	0.79	0.77	6.41	9.83	4.08
	Basis Function	pureQuadratic							
	Kernel Function	Ard Squared Exponential							
SVM	Epsilon	7.7074	MinMax	0.53	0.51	0.51	13.33	8.19	5.31
	Box Constraint	399.0687							
	Kernel Scale	13.6945							
RF	Method	LSBoost	MinMax	0.97	0.86	0.82	10.47	17.33	15.53
	Learning Rate	0.9629							
	NumCycling	336							
ANN	Hidden Layers	2	MinMax	0.99	3.65	0.99	1.00	0.95	0.94
	Hidden Neurons	14							
	Mu	5.1143e-04							

및 적용이 가능한 최적 열환경 제어알고리즘의 개발은 필수적이며 추후 연구로 진행될 예정이다.

## Acknowledgement

본 연구는 산업통상자원부(MOTIE)와 한국에너지기술평가원(KETEP)의 지원을 받아 수행한 연구 과제입니다. (No. 20182010600010)

## Reference

- [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>, 2020.09.16.
- [2] A. S. G. Andrae, T. Edler, On Global Electricity Usage of Communication Technology: Trends to 2030, challenges, 6, pp.117-157, 2015.
- [3] 냉난방공조 신재생 녹색건축 전문저널 Kham, "데이터센터 냉각 트렌드", <https://www.kham.kr/news/article.html?no=9902>, 2020.09.16. // (Korea Heating Air-conditioning Refrigerations & renewable energy news, "Trends of Data Center Cooling System", <https://www.kham.kr/news/article.html?no=9902>, 2020.09.16.)
- [4] 조진균, 고밀도 데이터센터의 공조효율 및 에너지성능 개선을 위한 공조시스템 계획방법 연구, 연세대학교 박사학위논문, 2008. // (J.K. Cho, A study on the air conditioning system planning method for improving air conditioning efficiency and energy performance in high density data centers, Doctoral dissertation, Yonsei University, 2008.)
- [5] 김민휘, 박준영, 정재원, 데이터 센터 냉방을 위한 제습 증발냉각 시스템의 적용성 분석, 대한건축학회 학술발표대회 논문집, 33(1), pp.283-284, 2013. // (M.H. Kim, J.Y. Park, J.W. Jeong, Applicability of Desiccant and Evaporative Cooling System for Data Center Cooling Energy Savings, Architectural Institute of Korea 2013 Conference, 33(1), pp.283-284, 2013.)
- [6] 김지혜, 임태윤, 정차수, 외기냉방시스템이 적용된 데이터센터 CRAH의 급기온도와 설계 풍량에 따른 에너지성능 분석, 대한건축학회 논문집 - 구조계, 35(11), pp.181-188, 2019.11. // (J.H. Kim, T.Y. Aum, C.S. Jeong, Energy Performance of Air-side Economizer System for Data Center Considering Supply Temperature and Design Airflow Rate of CRAH(Computer Room Air Handler), Journal of the Architectural Institute of Korea Structure & Construction, 35(11), pp.181-188, 2019.11.)
- [7] 송석진, 머신러닝 기반 데이터센터 냉방 장치 자동제어 솔루션 개발 : 한전 사례를 중심으로, 대한전기학회 학술대회 논문집, pp.1793-1794, 2019.7. // (S.J. Song, Development of solution based on machine learning to control CRAC units in data centers : Focusing on KEPCO's case, Journal of The Korean Institute of Electrical Engineers Conference, pp.1793-1794, 2019.7.)
- [8] L. Yao, J.H.Huang, Multi-Objective Optimization of Energy Saving Control for Air Conditioning System in Data Center, Energies, 12(8), pp.1474, 2019.04.18.
- [9] M.H. Beitelmal, C.D. Patel, Model-Based Approach for Optimizing a Data Center Centralized Cooling System, Hewlett-Packard Laboratories White Paper, 2006.
- [10] Y. Chen et al., Integrated management of application performance, power and cooling in data centers, 2010 IEEE Network Operations and Management Symposium - NOMS 2010, pp.19-23, 2010.04.
- [11] L. Wang et al., Task scheduling with ANN-based temperature prediction in a data center: a simulation-based study, Engineering with Computers, 27, pp.381-391, 2011.
- [12] J. Gao, Machine learning applications for data center optimization. Google White Paper, 2014.
- [13] N. Lasic et al., Data center cooling using model-predictive control, Google White Paper, 2018.
- [14] C.E. Bash et al., Dynamic Thermal Management of Air Cooled Data Centers, Thermal and Thermomechanical Proceedings 10th ITERSOCIETY Conference on Phenomena in Electronics Systems, 2006.
- [15] J. Deng et al., Self-tuning PID-type Fuzzy Adaptive Control for CRAC in Datacenters, 7th International Conference on Computer and Computing Technologies in Agriculture (CCTA), pp.215-225, 2013.09.
- [16] D. WEI, et al., Energy-Efficient Mode Transition Strategies of Data Center Cooling Systems, International Journal of Simulation - Systems, Science & Technology, Vol. 17, Issue 37, pp.26.1-26.6, 2016.
- [17] M. Zapater et al., Runtime data center temperature prediction using Grammatical Evolution techniques, Applied Soft Computing, Vol. 49, pp.94-107, 2016.12.
- [18] A. Beghi et al., Modelling and control of a free cooling system for Data Centers, Energy Procedia, Vol.140, pp.447-457, 2017.12.
- [19] J. Ni et al., Simulation of thermal Distribution and Airflow for Efficient Energy Consumption in a Small Data Centers, Sustainability, 9(4), pp.664, 2017.04.
- [20] 김영진, 박철수, 가우시안 프로세스 모델과 냉동기 실시간 최적 제어, 대한건축학회 논문집 - 계획계, 30(7), pp.221-220, 2014.07. // (Y.J. Kim, C.S. Park, Gaussian Process Model for Real-Time Optimal Control of Chiller System, Journal of the Architectural Institute of Korea Planning & Design, 30(7), pp.211-220, 2014.07.)
- [21] J. Niemann, K. Brown, V. Avelar, Impact of Hot and Cold Aisle Containment on Data Center Temperature and Efficiency, Schneider Electric White Paper 135.
- [22] ASHRAE Datacom Series 10, Green Tips for Data Centers, 2011.
- [23] ASHRAE Guideline 14, Measurement of Energy and Demand Savings, 2002.
- [24] 윤상용 외 4인, 베이저안 최적화 기반 능동 학습을 통한 효율적 유기 분자 탐색, 한국정보과학회 학술발표논문집, pp.772-774, 2015.06. // (Y.W. Yoon et al., Efficient Organic Molecule Search via Bayesian Optimization Active Learning, The Korean Institute of Information Scientists and Engineers 2015 Conference, pp. 772-774, 2015.06.)