

스마트 수직농장의 에너지 효율 향상 및 생산비 절감에 관한 연구

(A Study on Improving Energy Efficiency and Reducing Operating Costs in Smart Vertical Farms)

정광훈*, 이명훈**

(Gwang Hoon Jung, Meong-Hun Lee)

요약

스마트 수직농장은 도시 공간을 효율적으로 활용하여 연중 안정적인 농산물 생산이 가능하지만 높은 에너지 소비와 운영비로 인해 경제성 확보에 한계가 있다. 본 연구의 목적은 환경·운영 데이터를 통합적으로 분석하여 수직농장의 에너지 효율을 향상시키고 생산비를 절감할 수 있는 머신러닝 기반 운영 최적화 방안을 제시하는 것이다. 이를 위해 국내 실제 수직농장에서 2024년 3월부터 12월까지 약 9개월간 수집된 환경 데이터(온도, 습도, 토양 수분, 풍향, 풍속 등)와 운영 데이터(전력 사용량, 인력 투입 시간, 물 사용량, 생산비 등)를 하나의 시계열 데이터셋으로 통합하였다. 결측치 및 이상치 처리, 범주형 인코딩, 스케일링, 시계열 분할 등의 전처리 과정을 거쳐 Random Forest, XGBoost, LSTM 모델을 구축하고 예측 성능을 비교하였다. 실험 결과, XGBoost 모델이 평균적으로 가장 낮은 예측 오차를 보였으며, 전통적인 규칙 기반 예측 대비 약 17~18%, 이동평균 기반 예측 대비 약 12~13%의 오차 개선 효과를 나타냈다. 연구 대상 수직농장의 일별 누적 전력 사용량이 평균 약 3,400~4,000 kWh/day 수준임을 고려할 때, RMSE 3.37은 평균 사용량 대비 약 0.1% 미만의 상대 오차에 해당하며, 이는 실제 수직농장 운영 의사결정에 충분히 활용 가능한 정확도로 해석된다. 이러한 예측 정확도 향상은 불필요한 조명 및 냉난방 설비의 과잉 운영을 줄이고 에너지 사용의 불확실성을 완화함으로써, 에너지 비용 및 생산비 절감에 기여할 수 있는 잠재력을 가진다. 본 연구는 환경 변수뿐 아니라 전력, 인력, 자원 투입을 동시에 고려하는 통합 운영 관점의 데이터 기반 최적화 모델을 제안함으로써, 단순 제어 규칙을 넘어 생산성과 경제성을 함께 고려할 수 있는 스마트 수직농장 운영 전략을 제시한다.

■ 중심어 : 스마트 수직 농장 ; 에너지 효율성 ; 운영 비용 ; 머신 러닝 ; 탄소절감

Abstract

While smart vertical farms enable stable year-round crop production by efficiently utilizing urban spaces, they face economic limitations due to high energy consumption and operational costs. The purpose of this study is to propose a machine learning-based operational optimization strategy to improve energy efficiency and reduce production costs in vertical farms through an integrated analysis of environmental and operational data. To this end, an integrated time-series dataset was constructed using environmental data (temperature, humidity, soil moisture, wind direction, wind speed, etc.) and operational data (power consumption, labor hours, water usage, production costs, etc.) collected from a real-world vertical farm in Korea over a nine-month period from March to December 2024. Following preprocessing steps—including missing value and outlier treatment, categorical encoding, scaling, and time-series splitting—predictive models such as Random Forest, XGBoost, and LSTM were developed and compared for performance. Experimental results indicated that the XGBoost model achieved the lowest average prediction error, demonstrating an improvement of approximately 17–18% over traditional rule-based predictions and 12–13% over moving average-based predictions. Given that the daily cumulative power consumption of the subject farm averages between 3,400 and 4,000 kWh/day, the observed RMSE of 3.37 represents a relative error of less than 0.1% signifying a level of accuracy highly applicable to real-world operational decision-making. Such improvements in predictive accuracy have the potential to reduce energy and production costs by minimizing the excessive operation of lighting and HVAC systems and mitigating uncertainty in energy demand. This study suggests a data-driven optimization model from an integrated operational perspective—considering power, labor, and resource inputs alongside environmental variables—thereby providing a strategic framework for smart vertical farm operations that enhances both productivity and economic viability beyond simple control rules.

■ keywords : Smart vertical farm ; Energy efficiency ; Operating costs ; Machine learning ; Low carbon

I. 서론

세계 인구의 도시 집중화와 농업 환경의 급격

* 정회원, 국립순천대학교 스마트농업전공 박사과정

** 종신회원, 국립순천대학교 융합바이오시스템기계공학과 교수

"이 논문은 2025학년도 국립순천대학교 글로벌대학 사업비의 지원을 받아 연구되었음"

접수일자 : 2025년 12월 18일

수정일자 : 2026년 01월 13일

게재확정일 : 2026년 02월 09일

교신저자 : 이명훈 e-mail : leemh5544@gmail.com

한 변화로 인해, 식량안보와 에너지 문제는 지구적 차원의 주요 관심사가 되었다[1]. 특히 도심 지역에서는 경작지가 제한적이고 인건비나 운영비가 높아, 안정적으로 농산물을 생산하기가 쉽지 않다[2]. 이러한 배경에서 등장한 스마트 수직농장(Smart Vertical Farm)은 건물 내부나 옥상, 혹은 고층 빌딩과 같은 도심 공간을 활용하여 고밀도로 작물을 재배함으로써, 도시 내 농업 생산성을 높일 수 있는 새로운 대안으로 주목받고 있다[3]. 수직농장은 외부 기후 조건에 구애받지 않고 연중 생산이 가능하나, 에너지 사용량이 과도하고, 기계·설비 운영비용이 높다는 문제점도 함께 지적되고 있다[4].

최근 다양한 기술적 시도가 이뤄지고 있지만, 수직농장의 에너지 효율 향상과 생산비 절감은 여전히 쉽지 않은 과제로 남아 있다[5]. 설비·장비 자체의 성능 향상만으로는 한계가 분명하며, 복잡한 환경 인자(광량, 온도, 습도, CO₂ 농도, 양액 공급 등) 간 상호작용을 체계적으로 이해하고 최적화해야 하기 때문이다[6]. 이에 따라, 대규모로 축적되는 농장 운영 데이터를 바탕으로 머신러닝(Machine Learning) 기법을 적용하여 운영 효율을 전반적으로 개선하고자 하는 연구가 활발히 시도되고 있다[7-9].

본 연구는 이러한 문제의식에서 출발하여, 스마트 수직농장에서 축적되는 운영 데이터를 머신러닝 기법으로 분석함으로써 에너지 효율을 높이고 운영비를 절감할 수 있는 방안을 제시하고자 한다. 구체적으로, 다차원적 특성을 갖는 수직농장 데이터를 다룰 때 효과가 높은 Random Forest, XGBoost, LSTM 등 대표적인 머신러닝·딥러닝 모델을 선정하여 에너지 소비량 및 생산비를 예측·최적화한다. 이를 통해 세 모델 간 예측 성능을 비교·평가하고, 어떤 모델이 수직농장 운영 개선에 가장 유리한 결과를 도출하는지를 정량적으로 살펴보는 것을 핵심 목표로 삼는다.

본 논문의 구성은 다음과 같다. 2장에서는 수직농장의 개념과 머신러닝 기반 최적화 연구에 대한 관련연구를 기술한다. 3장에서는 연구 데이터를 포함한 본문으로, 모델 구축 과정과 주요 변

수 설정, 훈련 및 성능 평가 방법을 기술한다. 마지막으로 4장에서는 연구 결과를 요약하고 추후 연구방안에 대해 기술한다.

스마트 수직농장이 미래 농업의 핵심 대안으로 거론되는 만큼, 본 연구를 통해 환경·자원 요소를 통합적으로 다루는 머신러닝 접근이 수직농장 운영의 실효적 해법이 될 수 있음을 보여주고자 한다.

II. 관련 연구

1. 국내외 스마트 수직농장 사례

가. 해외 사례

(1) AeroFarms (미국 뉴저지)

세계 최대 규모의 실내 수직농장 중 하나로 꼽히며, 빅데이터와 IoT 센서를 활용해 재배 환경(광량, 온도, 습도 등)을 실시간 모니터링하고 자동 제어하며, 작물 생육 정보를 클라우드로 전송해, 머신러닝 알고리즘을 통해 생육 속도, 작물 품질 및 병해 위험 등을 예측한다[10]. 이를 바탕으로 광 세기, 양액 조성, 공조 시스템 운용 등 세부 조건을 자동으로 조정해 생산성을 높이고 에너지 사용을 줄이는 전략을 구사한다[11].

(2) Bowery Farming (미국 뉴욕)

‘BoweryOS’라는 자체 개발 소프트웨어 플랫폼에 기반하여, 농장 운영에 필요한 모든 데이터를 센서와 카메라 등을 통해 수집하며, 데이터를 머신러닝으로 분석해 작물별 최적 생육 조건을 찾아내며, 독자적인 AI 모델을 통해 조명 제어와 양액 주기를 자동화하여 인력 부담을 줄이고 전력 소비를 최소화한다[12]. 회사 발표에 따르면 전통 농업 대비 약 95% 적은 물을 사용하면서도 균일한 생산 품질을 유지할 수 있다고 알려져 있다[13].

(3) Spread ‘TechnoFarm’ (일본 교토)

일본에서 대규모 상업용 스마트 수직농장을 운영하는 ‘Spread’사는 2018년부터 TechnoFarm 프로젝트를 가동했으며, 재배 환경 정보를 자동 수집하고 로봇을 통한 일부 공정(파종, 수확 등)을 도입해 인건비를 낮추고 있으며, 빅데이터를

활용해 에너지 소비 패턴을 분석한다[14]. 이 과정에서 머신러닝 기반의 예측 모델을 구축하여, LED 조명 조절 및 냉난방 제어를 고도화한다[15].

(4) Infarm (유럽)

슈퍼마켓이나 레스토랑 내부에 설치되는 분산형(mini) 수직농장으로 유명하며, 각 유닛(소규모 재배 모듈)에서 수집되는 생육·환경 데이터를 본사 클라우드에 축적하고, 머신러닝을 이용해 작물별 재배 프로토콜을 최적화한다. 해당 모델을 기반으로 모든 유닛에 원격 제어 명령을 내려, 작물 품질과 에너지 효율을 표준화한다[16].

나. 국내 사례

(1) Farm8 ‘Planty Cube’

국내 기업 Farm8이 운영하는 ‘Planty Cube’는 주로 백화점이나 대형마트 등에 설치되는 소규모 수직농장 형태로, 작물의 성장 주기를 데이터로 축적해 AI 분석에 활용하며, 환경 변화에 따른 작물 반응을 예측하고, 소비지 근거리 설치를 통해 물류비를 절감하는 동시에 수확 시점을 정밀하게 관리한다. 일부 매장에서 에너지 소비량을 실시간 모니터링하여, 일정 수준을 넘으면 조명 강도를 자동으로 조정하는 방식을 시도 중이다[17].

(2) LG CNS ‘스마트 팜 팩토리’

LG CNS가 SI(System Integration) 기술을 활용해 구축한 스마트팜 시설로, 온실 환경과 식물공장 시스템을 통합 관리하며, IoT 센서와 엣지 컴퓨팅을 통한 데이터 수집 체계를 갖추고 있으며, 머신러닝 알고리즘을 적용해 병해충 예찰 및 조명-냉난방 최적화 등을 구현한다. 아직 대규모 상용화 단계는 아니지만, 에너지 효율과 인력 절감 효과를 극대화하기 위한 테스트 베드로 운용되고 있다[18].

2. 머신러닝 기반 최적화 적용 사례

가. AeroFarms의 AI-Driven Operation

AeroFarms는 수직농장 운영 시스템에 딥러닝을 접목해, LED 조명 스펙트럼과 공조 시스템 운용 스케줄을 자동으로 제어하고 있으며, 특정 작물의 성장 단계와 에너지 사용량 간의 상관관계를 머신러닝 모델로 학습해, 필요한 시점에만 냉난방 또는

조명을 가동함으로써 전력 소비를 효율화했다. 회사측 발표에 따르면, 머신러닝 도입 이후 동일 생산량 기준으로 15% 이상의 에너지 절감을 달성했고, 수확 주기를 10%가량 단축했다고 보고했다[19].

나. Bowery Farming의 예측 모델 기반 운영

Bowery Farming은 자사의 ‘BoweryOS’에서 머신러닝 회귀 모델을 활용해 시간대별 전력 부하, 온도 상승 패턴, 생육 특성 등을 예측하며, 이를 통해 일별·주별 조명 가동 시간을 자동으로 최적화하며, 소비 전력을 평균 20%가량 절감하는 효과를 얻었다고 밝히고 있다. 병해충 위험 관리에도 머신비전과 AI를 접목, 잎에 나타나는 미세 점박이나 색상 변화를 실시간 감지·학습해 초기 단계에서 방제함으로써 병해 비용을 줄이고 작물 품질을 안정화했다[20].

III. 연구 설계 및 데이터 전처리

본 논문은 스마트 수직농장의 다양한 환경·운영 데이터를 수집·분석하여, 에너지 소비량 및 운영비를 예측하고 최적화하기 위한 머신러닝 모델을 구축했다. 구체적으로, Random Forest (RF), XGBoost, LSTM 등 세 가지 알고리즘을 활용하여 모델별 예측 성능을 비교하고, 각 모델이 에너지 효율 및 비용 절감에 어떻게 기여할 수 있는지 평가하며 이를 위해 실제 운영 중인 국내 수직농장에서 2024년 03월 01일부터 2024년 12월 31일까지의 약 9개월간 수집된 환경 데이터(온도, 습도, 토양수분, 풍속, 풍향 등)와 운영 데이터(전력 사용량, 인력 투입, 물 사용량, 생산비 등)를 활용하였다.

표 1. 환경 데이터 특성

데이터 특성	
temp	온도 (°C)
humid	습도 (%)
soil_moisture	토양 수분 (%)
wind_direction	풍향 (°)
wind_speed	풍속 (m/s)
weather	날씨 (맑음, 흐림 등)
point	관측 지점 (위치 정보)

<표 1>은 재배 환경을 정의하는 핵심 요소들로 구성된다. temp(온도)와 humi(습도)는 작물 생육 조건을 결정짓는 대표적인 기후 변수로, 작물의 성장 속도와 병해충 발생 가능성에 직결된다. soil_moisture(토양 수분)는 작물 뿌리 부근의 습윤 상태를 나타내며, 물 사용량 최적화와 작물 스트레스 관리를 위해 중요한 지표다. 외부 기상 영향과 실내 환기의 효율을 판단하기 위해 wind_direction(풍향), wind_speed(풍속), weather(날씨)를 함께 측정하였으며, 센서가 설치된 point(관측 지점) 정보를 추가해 장소별·구역별 변화를 비교할 수 있도록 했다.

표 2. 운영 데이터 특성

데이터 특성	
power_consumption	전력 사용량 (kWh)
manpower_hours	인력 투입 시간 (인-시, Man-hour)
water_usage	물 사용량 (L)
nutrient_solution	양액 사용량 (L)
harvest_weight	작물 수확량 (kg)
production_cost	생산비 (원, 재료비 + 인건비 + 전력비 등 합산)
selling_price	판매 단가 (원/kg)

<표 2>는 농장의 경제성과 효율성을 가늠하는 운영데이터 항목을 담고 있고 power_consumption(전력 사용량)은 조명, 냉난방, 양액 공급 등에 소요되는 전체 에너지를 나타내며, 스마트 수직농장 비용 구조에서 가장 큰 비중을 차지하기에 핵심 분석 대상이다. manpower_hours(인력 투입 시간)는 작업(파종, 수확, 유지보수 등)에 필요한 노동량을 정량화한 것으로, 자동화 수준 및 운영 효율 평가에 활용된다. 이와 함께 water_usage(물 사용량)와 nutrient_solution(양액 사용량)을 측정하여, 재배 과정에서 소요되는 자원 투입 규모를 파악했다. 한편, harvest_weight(수확량)은 기간별 생산성을 정량적으로 확인할 수 있는 지표이며, production_cost(생산비)와 selling_price(판매 단가)를 연동하면 실제 손익 구조를 구체적으로 산출할 수 있다. 본 연구에서는 1분 단위를 기준으로 약 22만 건의 환경

데이터(온도, 습도, 풍속, 토양 수분 등)와 운영 데이터(전력 사용량, 인력 투입, 물 사용량, 생산비 등)를 9개월 동안 수집하고, 이를 하나의 데이터프레임(df)에 통합한 상태로 머신러닝 모델 학습을 준비하였다. 데이터 전처리는 크게 결측치 처리, 이상치 제거, 범주형 인코딩, 스케일링, 시계열 분할로 구분하였으며, 다음과 같은 과정을 거쳤다.

```
df['temp'] = df['temp'].interpolate(method='linear', limit_direction='forward')
df['humi'] = df['humi'].interpolate(method='linear', limit_direction='forward')
df['wind_speed'] = df['wind_speed'].interpolate(method='linear', limit_direction='forward')
df['soil_moist'] = df['soil_moist'].interpolate(method='linear', limit_direction='forward')

df['power_consumption'] = df['power_consumption'].fillna(df['power_consumption'].mean())
df['manpower_hours'] = df['manpower_hours'].fillna(df['manpower_hours'].mean())
df['water_usage'] = df['water_usage'].fillna(df['water_usage'].mean())
df['production_cost'] = df['production_cost'].fillna(df['production_cost'].mean())
df['nutrient_solution'] = df['nutrient_solution'].fillna(df['nutrient_solution'].mean())
df['harvest_weight'] = df['harvest_weight'].fillna(df['harvest_weight'].mean())

df.dropna(inplace=True)
```

그림 1. 데이터 결측치 처리

수직농장에서는 센서 장애나 로그 누락으로 인해 온도, 습도, 풍속, 토양 수분 등 환경 변수나 전력 사용량, 인력 투입, 물 사용량, 생산비 등 운영 변수에 결측치가 발생할 수 있다.

<그림 1>은 데이터 결측치 처리 과정으로 시계열 간격이 짧은 환경 변수는 선형 보간, 운영 변수는 평균값 대체를 우선 사용하고, 처리 불가능한 경우는 해당 샘플을 제거하여 모델 성능 저하를 방지했다.

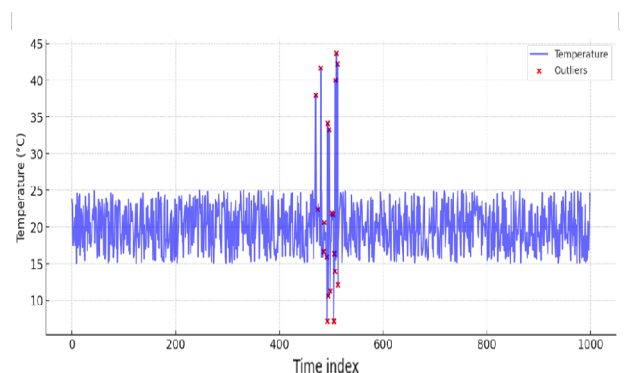


그림 2. 이상치 데이터 확인

<그림 2>는 시계열 순서에 따라 정렬된 데이터 샘플 인덱스를 x축으로 하고, 각 환경·운영

변수의 측정값을 y축으로 나타낸 이상치 시각화 결과이다. x축은 시간 흐름에 따른 데이터 발생 순서를 의미하며, 특정 시점에서 비정상적으로 큰 값이나 음수 값이 나타나는 경우 센서 오류 또는 로그 기록 이상으로 판단하였다. 데이터의 신뢰성을 높이기 위해 이상치를 제거하거나 적절한 값으로 대체하는 전처리 과정을 수행하였으며 또한, 이상치 제거 전·후의 데이터를 비교 분석하여 모델 성능에 미치는 영향을 평가하였다.

```
cond_temp_low = (df['temp'] < 15)
cond_temp_high = (df['temp'] > 30)
cond_humi_low = (df['humi'] < 0)
cond_humi_high = (df['humi'] > 100)
cond_wind_high = (df['wind_speed'] > 50)
cond_moist_low = (df['soil_moist'] < 0)
cond_moist_high = (df['soil_moist'] > 100)

cond_pwr_neg = (df['power_consumption'] < 0)
cond_pwr_extreme = (df['power_consumption'] > 10000)
cond_man_neg = (df['manpower_hours'] < 0)
cond_man_extreme = (df['manpower_hours'] > 24)
cond_water_neg = (df['water_usage'] < 0)
cond_water_ext = (df['water_usage'] > 100000)
cond_prod_neg = (df['production_cost'] < 0)
cond_prod_ext = (df['production_cost'] > 1e7)
cond_nutr_neg = (df['nutrient_solution'] < 0)
cond_nutr_ext = (df['nutrient_solution'] > 1000)
cond_harv_neg = (df['harvest_weight'] < 0)
cond_harv_ext = (df['harvest_weight'] > 1000)
```

그림 3. 이상치 제거 조건 설정

<그림 3>은 제시된 조건 범위를 벗어나는 데이터를 이상치로 판단하여 제거하였다. 예컨대 온도(temp)가 15도 미만이거나 30도를 초과하면 비정상 상태로 간주하고, 습도(humi)가 0% 미만 또는 100%를 넘어서면 센서 장애로 판단하였다. 또한 전력 사용량(power_consumption)이 음수로 기록되거나 매우 큰 값(10,000 kWh 이상)이면 실제 운영 현실에 부합하지 않으므로 제거 대상으로 삼았다.

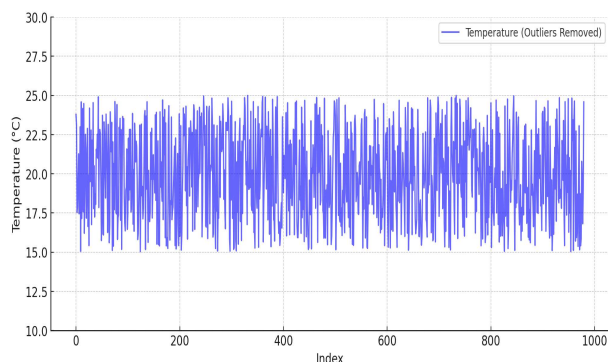


그림 4. 이상치 데이터 제거 후 시각화

<그림 4>은 이상치 데이터를 삭제하여 시각화한 그림이다. 이상치를 제거한 결과 데이터가 일정 범위 안에 분포함을 확인할 수 있다. 이를 통해 전처리를 마친 데이터가 물리적·현실적으로 타당한 범위에 집중되었으며, 이로써 모델 학습 시 이상치에 의한 성능 저하나 왜곡 가능성을 줄일 수 있게 되었다.

숫자형태로 저장되는 일반적인 환경변수와 달리 날씨(맑음·흐림·비·눈)나 재배 구역(A존·B존·C존)처럼 문자열 형태로 저장된 범주형 변수는 머신러닝 모델에서 직접 해석하기 어렵다. 이를 해결하기 위해, 본 연구에서는 원-핫 인코딩(One-Hot Encoding) 방식을 사용하여 각 카테고리를 0·1 값으로 변환했다. 예를 들어, weather 열에 “맑음”이 기록되면 weather_맑음=1, 나머지(weather_흐림, weather_비, weather_눈)는 0으로 처리한다. 이렇게 하면 모델이 범주별 차이를 정량적으로 학습할 수 있어, 날씨나 구역 변화가 에너지 사용량이나 생산비에 미치는 영향을 쉽게 파악할 수 있다.

```
from sklearn.preprocessing import MinMaxScaler

scaler = MinMaxScaler()
num_cols = [
    'temp', 'humi', 'wind_speed', 'soil_moist', 'power_consumption',
    'manpower_hours', 'water_usage', 'production_cost',
    'nutrient_solution', 'harvest_weight'
]
df[num_cols] = scaler.fit_transform(df[num_cols])
```

그림 5. Min-Max scaling 진행

<그림 5>는 스케일링을 진행하는 코드이며, 온도(°C), 습도(%), 전력(kWh), 생산비(원) 등 서로 다른 물리 단위를 사용하는 변수가 공존할 경우, 특정 변수가 값의 절대 크기 때문에 모델 학습 과정에서 과도하게 지배적인 역할을 할 우려가 있다. 이를 방지하기 위해 본 연구에서는 모든 수치형 변수를 0~1 범위로 정규화하는 Min-Max Scaling 방식을 적용하였다.

이로써 변수 간 균형을 맞추고, 모델이 다양한 입력 변수를 고르게 학습할 수 있도록 했다. 또한, 데이터 분포 특성이나 모델 목적에 따라 표준화

(Z-score) 기법을 적용하는 방안도 검토했으나, 실제 데이터의 분포 및 범주형 변환 등을 종합 고려해 Min-Max 방식을 최종 선택하였다. 이를 통해 온도와 전력 등 상이한 스케일의 변수가 동일한 기준 위에서 학습될 수 있도록 하였다.

Input: Dataset D with a date column
Output: Training set, Validation set, Test set

1. Sort dataset D in ascending order based on the date column
2. Let n be the total number of samples in D
3. Define training size as 80 percent of n
4. Define validation size as 10 percent of n
5. Assign the first training size samples to the training set
6. Assign the next validation size samples to the validation set
7. Assign the remaining samples to the test set
8. Return the training, validation, and test sets

그림 6. 시계열 분할

<그림 6>은 시계열 분할에 관한 그림으로 수직농장의 환경·운영 데이터는 시간 순서에 따라 누적되므로, 과거 정보를 학습하여 미래 시점을 예측하는 형태의 분석이 일반적이다. 이를 구현하기 위해, 전체 데이터를 시간 순서로 정렬한 뒤 훈련(train), 검증(validation), 테스트(test) 세 구간으로 나누었다. 예를 들어, 80%는 훈련, 10%는 검증, 10%는 테스트 세트로 할당한다. 이렇게 하면 모델이 과거 데이터(훈련)를 학습하고, 검증 세트로 하이퍼파라미터 등을 조정한 후, 실제 미래 상황과 유사한 테스트 세트로 최종 성능을 평가할 수 있다. 이 과정을 통해 데이터 누수(Data Leakage) 없이 모델을 개발·검증할 수 있으며, 미래 시점 에너지 소비 예측이나 생산비 관리 등에 대한 정확도를 높일 수 있다.

IV. 모델 구성 및 성능 평가

스마트 수직농장의 에너지 효율 향상 및 생산비 절감을 확인하기 위해 시간 순서대로 축적된 환경·운영 데이터를 통해 미래 시점의 전력 사용량(또는 생산비 등)을 예측한다. 이를 위해 Random Forest, XGBoost, LSTM 세 가지 모델을 비교·분석하였다. 각 모델의 하이퍼 파라미터

는 검증세트를 활용해 다음과 같이 최적화 하였다.

표 3. 모델별 하이퍼파라미터 탐색 결과

모델	하이퍼 파라미터	후보 값	검증세트 RMSE	최종 선택 값
Random Forest	n_estimators	100, 200 300, 400	100→3.58, 200→3.49, 300→3.47, 400→3.48	200
	max_depth	5, 7, 9	5→3.58, 7→3.49, 9→3.47	7
	min_samples_split	2, 5, 10	2→3.50, 5→3.48, 10→3.47	10
XGBoost	n_estimators	150, 200, 250, 300	150→3.45, 200→3.38, 250→3.37, 300→3.39	150
	learning_rate	0.01, 0.05, 0.1	0.01→3.47, , 0.05→3.37, , 0.1→3.38	0.05
	max_depth	3, 4, 5	3→3.42, 4→3.37, 5→3.40	4
	reg_lambda	0.0, 0.5, 1.0	0.0→3.39, 0.5→3.37, 1.0→3.38	0.5
LSTM	hidden_units	32, 64, 128	1→3.52, 2→3.44, 3→3.47	64
	num_layers	1, 2, 3	1→3.52, 2→3.44, 3→3.47	2
	dropout	0.1, 0.2, 0.3	0.1→3.46, 0.2→3.44, 0.3→3.45	0.2
	epoch	50, 100, 150	50→3.49, 100→3.46, 150→3.44	150

모델별 하이퍼 파라미터 탐색 결과는 다음 <표 3>과 같으며 Random Forest의 경우 트리 개수가 증가할수록 전반적으로 RMSE가 조금씩 낮아졌으며, 300개 부근에서 최적화되었으며, 트리 깊이는 9에서 최저 오차를 달성했다. min_samples_split을 10으로

했을 때 과적합 없이 안정적인 성능을 보였으며, RMSE 3.47을 기록했다. XGBoost의 경우 `n_estimators`는 250에서 RMSE 3.37를 기록하였으며, `learning_rate`를 0.05로 낮추면 수렴 속도는 느려지지만(훈련 시간이 약간 늘어남), 과적합 위험을 줄여 검증 오차가 가장 낮았다. `max_depth=4`, `reg_lambda=0.5`로 정규화를 적용해 모델이 지나치게 복잡해지는 것을 억제했다. LSTM의 경우 은닉 유닛이 64, 층 수가 2, `dropout=0.2`, `epoch=150`일 때 RMSE가 3.44로 가장 낮았다. 시계열 특성을 반영해 급작스러운 변동 구간을 잘 캐치하지만, 학습에 많은 시간이 소요되었고 하이퍼파라미터 튜닝 범위에 따라 소폭의 오차 변동이 있었다.

세 모델 중 XGBoost가 평균적으로 가장 낮은 오차를 보였으며 전력 사용량 예측에 있어서 우수한 성능을 달성했다. 다만 시계열 특화 모델인 LSTM이 변동성이 큰 일부 구간에 대해서는 더 세밀한 추론 능력을 보였으므로, 목적에 따라 모델 선택이 달라질 수 있다.

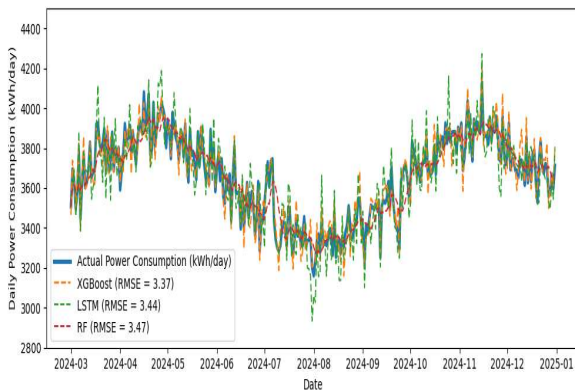


그림 7. 일별 전력 사용량에 대한 세 모델(RF, LSTM, XGBoost)의 시계열 예측 성능 비교

<그림 7>은 일별 전력 사용량에 대한 RF, LSTM, XGBoost의 시계열 예측 성능 비교를 나타낸 그림이다.

본 연구 대상 수직농장의 일별 누적 전력 사용량은 평균적으로 약 3,400 - 4,000 kWh/day 수준으로 나타났으며, 이를 기준으로 할 때 RMSE 3.37은 평균 사용량 대비 약 0.08 - 0.10% 수준의 상대 오차(NRMSE)에 해당한다.

이는 일별 에너지 수요 예측 관점에서 매우 작은 오차 범위로, 실제 운영 의사결정에 충분히 활용 가능한 정확도로 판단된다.

또한 본 연구에서는 단순 통계적 기준 모델과의 정략적 비교를 진행하였다.

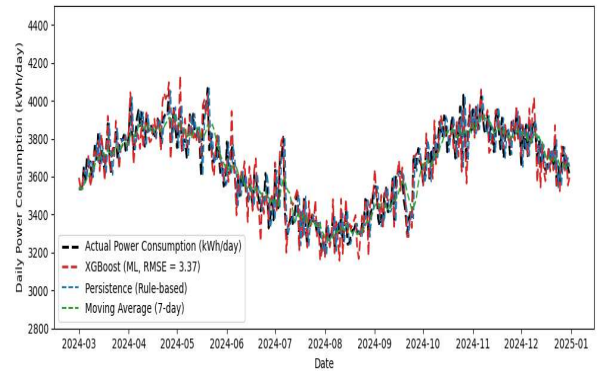


그림 8. 단순 통계적 기준 모델과 XGBoost 기반 일별 전력 사용량 예측 결과의 시계열 비교

<그림 8>은 단순 통계적 기준 모델과 XGBoost 기반 일별 전력 사용량 예측 결과의 시계열 비교를 나타낸 것이며, 비교 결과 XGBoost 모델은 전날 값을 그대로 사용하는 규칙 기반 예측 대비 전력 사용량 예측 오차(RMSE)를 약 17 - 18% 감소시켰으며, 7일 이동평균 기반 예측 대비로도 약 12 - 13%의 오차 개선 효과를 보였다. 이러한 결과는 머신러닝 기반 접근법이 기존의 단순 규칙 또는 통계적 방법에 비해 연간 전기요금 기준 약 5 - 7% 수준의 비용 절감 잠재력을 가지는 것으로 판단된다.

V. 결 론

본 연구를 통해 스마트 수직농장이 단순히 조명 및 냉난방 설비의 효율을 개선하는 방식만으로는 에너지 비용을 획기적으로 절감하기 어렵다는 점을 재확인하였으며, 머신러닝 기반 운영 최적화가 기존 규칙 기반 또는 단순 통계적 운영 방식 대비 실질적인 성능 향상을 제공할 수 있음을 확인하였다. 단순 통계적 기준 모델과의 비교 결과, XGBoost 모델은 전날 값을 그대로 사용하는 규칙 기반 예측 대비 전력 사용량 예측 오차를 약 17.18% 감소시켰으며, 7일 이동평균 기반 예측 대비로도 약 12.13%의 오차 개선 효과를 보였다. 이는 복잡한 전력 부하 패턴과 환경 및 운영 변수 간의 비선형 관계를 고려할 때 머신러닝 기반 접근법의 유효성을 장량적으로 입증하는 결과이다. 본 연구 대상 수직농장의 일별 누적 전력 사용량은 평균적으로 약 3,400 - 4,000 kWh/day 수준으로 나타났으며, 이를 기준으로 할 때 RMSE 3.37은 평균 사용량 대비 약 0.08 - 0.10% 수준의 상대 오차에 해당한다. 이는 실제 수직농장 운영 의사결정에 충분히 활용 가능한 정확도로, 불필요한 조명 및 냉난방 설비의 과잉 운영을 사전에 방지하고 에너지 사용의 불확실성을 감소시키는 데 기여할 수 있다. 다만 본 연구는 약 9개월간 수집된 데이터를 기반으로

분석을 수행하였기 때문에, 작물 종류나 지역 및 계절 조건이 달라질 경우 모델 성능에 변동이 발생할 가능성이 있다. 향후 연구에서는 다양한 규모와 작물을 대상으로 한 장기 관측 데이터를 축적하고, 다중 모델 앙상블이나 실시간 최적화 기법을 적용함으로써 모델의 일반성과 실용성을 더욱 강화할 계획이다. 나아가 에너지 관리뿐만 아니라 작물 선택과 파종 스케줄, 물류까지 포함하는 종합 의사결정 시스템으로 확장함으로써 스마트 수직농장이 미래 도시농업의 핵심 인프라로 자리매김하는 데 기여할 수 있을 것으로 기대한다.

REFERENCES

- [1] HM Park, DS Ryu, SH Lee, "Development of an IoT-based Fertigation System in Apple Orchards," *ASHS Annual Conference*, pp. 224-228, Aug. 2020.
- [2] GH Jeon, BY Kang, CS Han, "Evaluation of Soil Moisture Sensors for High-Density Cherry Tomato Cultivation," *Journal of Horticultural Science & Biotechnology*, vol. 96, no. 3, pp. 310-315, May 2021.
- [3] AY. Seo, KW. Kim, BR Choi, "Analysis of Soil Nutrient Fluctuation in Pear Orchards under Different Sensor-Driven Fertigation Regimes," *HORTICULTURE ABSTRACTS*, pp. 254-259, Sep. 2022.
- [4] T. Akashi, M Nakayama, "Adoption of Automated Sensing Platforms for Precision Peach Orchard Irrigation," *HortScience*, vol. 57, no. 6, pp. 678-684, Jun. 2022.
- [5] BH Yoon, GW. Sung, "Nutrient Sensor Applications to Improve Pear Orchard Fertilizer Efficiency," *Proceedings of the 29th International Horticulture Congress*, pp. 88-93, Apr. 2022.
- [6] MY. Lee, KW. Jeong, "Comparative Study on Different Soil Moisture Sensors for Citrus Orchard Irrigation Scheduling," *Horticulture, Environment, and Biotechnology*, vol. 63, no. 4, pp. 521-529, Aug. 2022.
- [7] SH Kim, DG Park, ET. Ko, "Design of Sensor-Aided Fertigation for Vineyard Microclimate Control," *HORTICULTURE ABSTRACTS*, pp. 96-100, Sep. 2021.
- [8] Y. Xian, L. Tao, "Evaluating Soil Water Stress Indices with Sensor Arrays in Avocado Orchards," *Acta Horticulturae*, vol. 1330, pp. 77-84, Nov. 2021.
- [9] HQ Choe, H Yoo, and MH Lee, "Machine learning-based smart farm sensor data anomaly detection system research," *Proceedings of Symposium of the Korean Institute of communications and Information Sciences*, pp. 1454, Jun. 2023.
- [10] L. Duque, R. Tamayo, "Real-Time Monitoring of Banana Plantation Soil Moisture for Sustainable Irrigation," *Journal of Horticultural Science & Biotechnology*, vol. 97, no. 2, pp. 215-221, Jan. 2022.
- [11] NJ. O'Connell, MH Kubota, "Framework for Automated Orchard Nutrient Management in Almond Cultivation," *ASHS Annual Conference*, pp. 330-334, May 2023.
- [12] EK Song, JS Kang, DH Hong, "Analyzing Soil Salinity Trends in Strawberry Fields via Multisensor Fusion," *HORTICULTURE ABSTRACTS*, pp. 201-206, Nov. 2022.
- [13] R. Guenero, P. Alvarado, "Multi-Parameter Soil Sensor Calibration for Avocado Orchard Applications," *Horticulture Research*, vol. 8, no. 3, pp. 1-8, Aug. 2021.
- [14] BY. Kim, WS. Rhee, TJ. Gong, "Suitability Assessment of EC and pH Sensors for Organic Apple Orchards," *Symposium on Smart Agriculture*, pp. 144-149, Oct. 2022.
- [15] TD. Clarke, TL. Myra, "Distributed Sensing in a Commercial Pear Orchard: A Case Study," *Journal of Integrative Agriculture*, vol. 21, no. 9, pp. 2437-2445, Sep. 2022.
- [16] HW. Kang, QJ. Lee, "Evaluating Water and Nutrient Management in Hydroponic Strawberry Systems using Sensor Monitoring," *HORTICULTURE ABSTRACTS*, pp. 159-163, Jun. 2022.
- [17] DY. Shin, CW. Lee, "Performance Analysis of Different Soil Sensor Types for Blueberry Orchard Irrigation," *Biosystems Engineering*, vol. 214, pp. 231-239, Dec. 2021.
- [18] LM Song, EK Ahn, GH Wang, "Adaptive Fertilizer Recommendation System for Pear Orchards Based on Sensor Data," *HORTICULTURE ABSTRACTS*, pp. 177-180, Jul. 2022.
- [19] CJ. Tan, YN. Sung, "Machine Learning-Based Nutrient Management using Real-Time Soil Sensors in Grape Vineyards," *Precision Agriculture*, vol. 24, no. 2, pp. 455-469, Mar. 2023.
- [20] TJ. Hong, BD. Lee, MH. Na, "Comparative Analysis of Soil Sensor Accuracy for Organic Apple Orchard Management," *HORTICULTURE ABSTRACTS*, pp. 331-334, Sep. 2021.

저자 소개



정광훈(정회원)

2022년 순천대학교 정보통신공학과
학사 졸업.
2024년 순천대학교 정보통신공학과
석사 졸업.
2025년 순천대학교 정보통신공학과
박사과정 재학

<주관심분야 : 스마트팜, 사물인터넷, 머신러닝, 빅데이터 활용>



이명훈(종신회원)

2004년 국립순천대학교 정보통신공학
학사 졸업
2006년 국립순천대학교 정보통신공학
석사 졸업
2010년 한국전자통신연구원 선임연구원
2011년 국립순천대학교 정보통신공학
박사 졸업
2017년 국립농업과학원 농업연구사
2021년~현재 국립순천대학교 융합마
이오시스템기공학과 교수

<주관심분야 : 스마트농업, 스마트 농식품 표준 >