Journal of Agricultural Engineering

https://www.agroengineering.org/

Forecasting the vapor pressure deficit in vertical farming facilities aiming to provide optimal indoor conditions

Carlos Alejandro Perez Garcia,[1] Dafni Despoina Avgoustaki,[2] Enrica Santolini,[1] Daniele Torreggiani,[1] Thomas Bartzanas,[2] Patrizia Tassinari,[1] Marco Bovo[1]

[1]Department of Agricultural and Food Sciences, University of Bologna, Italy

[2]Department of Natural Resources Management and Agricultural Engineering, Agricultural University of Athens, Greece

Corresponding author: Marco Bovo, Department of Agricultural and Food Sciences, University of Bologna, Viale G. Fanin 40-50, 40127 Bologna Italy. E-mail: marco.bovo@unibo.it

Please cite this article as doi: 10.4081/jae.2025.1793

Submitted: 2 April 2025
Accepted: 1 July 2025

Forecasting the vapor pressure deficit in vertical farming facilities aiming to provide optimal indoor conditions

Carlos Alejandro Perez Garcia,[1] Dafni Despoina Avgoustaki,[2] Enrica Santolini,[1] Daniele Torreggiani,[1] Thomas Bartzanas,[2] Patrizia Tassinari,[1] Marco Bovo,[1]

[1]Department of Agricultural and Food Sciences, University of Bologna, Italy

[2]Department of Natural Resources Management and Agricultural Engineering, Agricultural University of Athens, Greece

Corresponding author: Marco Bovo, Department of Agricultural and Food Sciences, University of Bologna, Viale G. Fanin 40-50, 40127 Bologna Italy. E-mail: marco.bovo@unibo.it

Abstract

Vertical farming is a sustainable solution for urban agriculture by optimizing space and resources. However, this requires ideal indoor climatic conditions to achieve maximum crop yield and quality. This research develops and validates a prediction model based on NeuralProphet algorithm to assess the vapor pressure deficit in a vertical farming facility. The model uses environmental data such as temperature, relative humidity, and solar radiation to predict vapor pressure deficit (VPD), a key indicator of vegetation health and crop growth status. The model shows high accuracy and reliability with a root mean squared error (RMSE) of 34.80 and a mean absolute error (MAE) of 25.28. The model, demonstrating satisfactory performance in predicting VPD, enables optimization of indoor growth conditions, thereby improving resources use efficiency and minimizing operational costs. Finally, it indicates a promising application of advanced artificial intelligence tools in vertical farming management to establish a sustainable and economically feasible agricultural practice since the model can help to produce high quality crops through a precise control of environmental parameters.

Key words: Microclimate forecasting; indoor agriculture; NeuralProphet; vertical farming; vapor prssure deficit.

Introduction

Context of the work

Vertical farming (VF) is an indoor crop-growing practice that stacks crop vertically in a protected environment. Advantages of VF include efficiency in the utilization of space, low water consumption, cultivation in the off-season, and reduced or even elimination use of pesticides and herbicides

(Eldridge *et al*., 2020; Oh and Lu, 2023; Shamshiri *et al*., 2018). The VF structure is fully insulated from the outside environment, where farmers having the capability to manipulate the internal environment in regard to temperature, humidity, CO_2, among others (Avgoustaki and Xydis, 2020). One of the key aspects of VF operations is climate control due to its economic and environmental significance. Effective systems for climate control are therefore crucial, whereby the systems will create an optimum growing indoor environment that suits high-quality and high-yielding crop production and, at the same time, conserves energy. For ideal conditions, cooling and heating systems are usually required; in fact, such systems account for 65-85% of the energy demand in these buildings (Morales-García *et al*., 2023). This indicates that the amount saved is significant since heaters, humidifiers, curtain actuators, among others, can be switched on only when there is a need for them and turned off immediately after attaining the desired climate conditions. What this, therefore, guarantees are precision in temperature control done through the used sensors and actuators. In the past, attempts were made to optimize climate control in VFs. For instance, Revathi S *et al*., (2017) proposed a classical control system with a PID controller. Nevertheless, the principal drawback of PID-related systems is that they target the present condition of a sole environmental factor; hence, actions are made a posteriori.

The widespread deployment of Internet of Things (IoT) infrastructures for climate control in VF has enabled the acquisition of relevant measurements about the VF's operations with high temporal resolution and accuracy. This has enabled a paradigm shift in terms of control strategy, specifically towards Advanced Control strategies combined with Artificial Intelligence (AI) tools. In particular, the use of time series predictive model enables the development of advanced control strategies that use anticipation of disturbance variables to feed the environmental control system. (Castañeda-Miranda and Castaño-Meneses, 2020; Guillén-Navarro *et al*., 2021; Tzounis *et al*., 2017). In traditional feedforward control based on statistical models, is remarkable their inability to capture the complex patterns and relationships among variables (Vanegas-Ayala *et al*., 2022). In contrast, AI time-series models are promising for climate forecasting because they can analyze sequential data and make accurate predictions. By using historical climate data, AI models can identify subtle patterns and trends that may not be detected by traditional ones, resulting in improved predictive capabilities (Eraliev and Lee, 2023).

Vapor pressure deficit (VPD) is a pivotal metric in agricultural science, quantifying the discrepancy between the actual vapor pressure of the air and the saturation vapor pressure at a specific temperature (Oke, 2002). VPD can be defined as an indicator of the capacity of the air to retain moisture. A high VPD indicates a greater potential for water evaporation from surfaces, including plant leaves, due to a larger disparity between the actual moisture level in the air and its saturation point. Conversely, a low VPD value indicates that the air is closer to its saturation point, which in turn leads to reduced rates of water evaporation. The primary determinants of VPD are temperature and relative humidity where warmer temperatures and lower humidity levels tend to result in higher values. Previous research, such as that of Zhang *et al*. (2016), has demonstrated the strong relationship between VPD values and the main environmental parameters optimal for the crop growth.

On the other hand, current literature identifies three kinds of approaches to developing predictive models: i) white-box models which utilize first principles to build a physics-based model that requires large inputs and expert knowledge for setting up; ii) black-box models generated using input-output data in purely data-driven methods without considering physical relations, and iii) grey-box models which combine structure established from physical laws with parameters identified from input-output data (Hauge Broholt *et al*., 2022). Black-box modeling might be adequate as a first approach for several reasons, given the nature of a facility for vertical farming. First, there is no prior knowledge about the system; it helps when preliminary understanding of these complex interrelations between environmental factors is really low and secondly, when there is limited data availability from the installed supervisory system.

Neural networks represent a significant branch of AI, reflecting an effort to emulate intricate, highly integrated structures and learn from vast quantities of data. A neural network comprises numerous layers of interconnected nodes, or neurons, wherein each node performs a relatively straightforward mathematical operation. By arranging these layers in a hierarchical structure, neural networks facilitate automated feature extraction, enabling the identification of patterns within data and the relationships between them. In the context of VF, where optimizing indoor environmental conditions such as VPD is crucial for maximizing crop yield and resources efficiency, the selection of the proper model for setting up a forecast tool is a strategic decision. The careful selection of the most suitable model to use for the aim of the present work, led the authors to adopt NeuralProphet model (Triebe *et al*., 2021). In fact, NeuralProphet extends the principles of neural networks specifically to time series forecasting tasks, which are essential for accurately predicting VPD fluctuations over time. This tool is adept at handling the inherent complexities of time series data, including seasonality, trends, and irregular patterns, without requiring extensive manual feature engineering (Triebe *et al*., 2021).

Literature review of previous works

Several researchers worked on predicting and controlling environmental variables useful in agricultural facilities (Morales-García *et al*., 2023). For instance, in Singh *et al*. (2018), a mathematical model was developed to predict VPD in a greenhouse independently using internal and external climate parameters as inputs. For VPD, the mean standard deviation for the observed and predicted data was calculated to be 2.16 and 2.41 kPa, respectively, while the mean RMSE and R^2 were 0.56 and 0.94 kPa, respectively. Despite the accuracy obtained in the predictions, given the short time period of the productive cycles in structures such as the FVs, it will not provide an adequate model fit for the application context of the present investigation.

In Frausto *et al*. (2003) a study aimed to predict the indoor temperature of a greenhouse employs linear autoregressive models with external input and autoregressive moving average models with external input. Input variables such as outdoor air temperature, relative humidity, global solar radiation, and sky cloudiness are utilized for these models. While these models effectively capture the evolution of greenhouse temperature under normal conditions, they exhibit limitations during ventilation phases due to the nonlinear nature of ventilation strategies. Consequently, the accuracy of these models is diminished when the greenhouse undergoes ventilation processes.

Another study (Patil *et al*., 2008) investigates the prediction of greenhouse indoor temperature in Thailand using autoregressive models during ventilation processes. In addition to these models, the study incorporates a neural network for forecasting. Results demonstrate that the combination of neural networks with autoregressive models yields significantly improved accuracy in predicting this variable compared to using autoregressive models alone. In Castañeda-Miranda and Castaño (2017), a similar work to the previous study is conducted in Mexico, with the objective of forecasting the indoor temperature of a greenhouse. The authors propose a hybrid approach combining an autoregressive model with a neural network trained using the Levenberg-Marquardt backpropagation algorithm. The integrated model employs external climatic variables to enhance predictive accuracy. The results showed a 95% confidence temperature prediction, with a coefficient of determination of 0.9549 and 0.9590, for summer and winter, respectively.

Moreover, a study addressing greenhouse temperature prediction is discussed by Li *et al*. (2020) where the historical values of humidity, temperature, and light intensity are employs as input parameters for forecasting. The approach employs a Neural Network augmented with a K-Nearest Neighbor algorithm, representing a hybrid method aimed at improving prediction performance through the synergy of different machine learning techniques.

Aim of the present work

The main aim of this paper is to develop a forecast model using NeuralProphet for predicting Vapor Pressure Deficit in a VF facility. The research involves several stages, starting with comprehensive data analysis and preprocessing to ensure the quality and integrity of the dataset. This step requires cleaning the data and transforming variables as needed to prepare them for modeling. The training phase involved the NeuralProphet framework to train a predictive model on the preprocessed data.

The model architecture integrates historical VPD values up to 1 hour before, as well as exogenous variables representing outdoor environmental conditions such as temperature, relative humidity, and solar radiation. During this stage, hyperparameter tuning is performed to optimize the model's performance and ensure its effectiveness in capturing the complex relationships inherent in VPD dynamics within VF. After model training, evaluation procedures are carried out to assess the model's predictive accuracy and generalization capabilities. This involves analyzing different performance metrics, such as mean absolute error and root mean square error, to assess the model's effectiveness in predicting VPD values. Furthermore, the findings are carefully examined to reveal insights into the temporal patterns and causes of VPD fluctuations within the VF environment. This paper aims to deliver a robust and reliable forecasting model for VPD in VF settings by systematically traversing through research stages.

Materials and Methods

Theoretical background

Oke (2002) defines the vapor pressure deficit as the amount of vapor needed to reach atmospheric saturation. To understand this concept, it is important to know two key parameters: saturation vapor pressure "e_0" (in Pascal) and vapor pressure "e" (in Pascal). Saturation vapor pressure is the air pressure at the point where it can no longer hold additional water molecules, reaching its maximum capacity for vapor retention. On the other hand, e is the partial pressure exerted by water vapor molecules in the air. In Allen *et al*. (1998), is presented the relationship between e_0 (Pa) and e (Pa) as follow:

$$e = e_0 \cdot RH/100 \text{ (Pa)} \qquad \text{(Eq. 1)}$$

Where: RH (%) is the relative humidity of the environment, while e_0 (Pa) can be calculated from equation proposed by Yount (2017):

$$\ln e_0 = C_1/T + C_2 + C_3T + C_4T^2 + C_5T^3 + C_6 \ln T \qquad \text{(Eq. 2)}$$

Where: T is the absolute temperature in Kelvin; e_0 (Pa) and C_1(Pa ·°C), C_2(Pa), C_3 (Pa ·°C^{-1}), C_4(Pa ·°C^{-2}), C_5 (Pa ·°C^{-3}) and C_6 (Pa ·°C^{-1}) are the coefficients given for a temperature range from 0°C to 200°C. Following the equations described above, a Python script was developed to perform the conversion of internal temperature and humidity variables into VPD.

Description of the case study

The experiment was conducted within the pilot vertical farm located at the Laboratory of Farm Structures, established for research under the VF2FARM project at the Agricultural University of Athens, Greece (coordinates: N 37° 58.947120 E 23° 42.266640). The climate variables of the zone are mild, wet winters and hot, dry summers with average temperatures will be from 14°C to 20°C, reducing gradually as November advances. Relative humidity also augments, keeping a variation between 60% and 80%.

The pilot vertical farm was housed within a container (Figure 1) located at the Agricultural University of Athens - Laboratory of Farm Structures, providing approximately 30 m^2 of internal space. The external dimensions of the cube are 12.0 m in length, 2.4 m in width, and 3.0 m in height. The indoor space was equipped with four growing towers designated as T1, T2, T3, and T4. Each tower was configured with three growing layers labeled L1, L2, and L3. These growing towers and layers were integral to the experimental setup, facilitating controlled cultivation environments for various crops (Avgoustaki *et al*., 2024).

Figure 1. Shipping container housing the vertical farming facility (left) and controlled indoor environment equipped with growing towers and environmental sensors (rigth).

The indoor climate of the structure is controlled by the operation of two air conditioning units, each with a cooling capacity of 2.70 kW (i.e., 9212 Btu/h). These air conditioning units were essential in maintaining a stable and optimal growing environment by regulating the temperature inside the vertical farm. To ensure continuous air exchange with the outdoor environment, two exhaust fans located on the roof of the structure operated continuously, removing the inside air and replacing it with fresh outside air. To provide effective thermal insulation for the interior of the structure against external temperature fluctuations, the container envelope was clad with plates of polystyrene.

Monitoring and control variables

During the experimental test the cooling system was activated when the indoor air temperature reached the 25°C during the light period and 22°C during the dark period. The same temperature setpoint values were also adopted to turn off the cooling system. The atmospheric CO_2 level was ensured, internally the container, by continuous air exchange with the outdoor environment. To monitor environmental variables, data was recorded at a 10-minute interval using a GP1 Data Logger (Delta-T Devices, Cambridge, UK) and a thermo-hygrometric sensor TBSHT06 (TekBox Digital Solutions, Singapore) measuring air relative humidity and air temperature (for details see Figure 2). The GP1 can log up to two differential analog voltages, two temperature channels and two pulse counters. This is complemented by a high-precision sensor PR2 SDI-12 (Delta-T Devices, Cambridge, UK) measuring the substrate moisture content. The digital temperature sensor provided an accuracy of ±0.2°C over the range from 0°C to 90°C, while the humidity sensor had an accuracy of ±2% over the range from 0% to 100%. This robust monitoring system ensured that internal climate conditions were consistently recorded and maintained within the desired parameters to support optimal plant growth. The accuracy of the sensor measuring the substrate moisture content was ±4%.

Figure 2. GP1 Data Logger (left) and TBSHT06 air relative humidity and air temperature sensor (right).

Data processing

The dataset used for this study span from October 25th, 2023 to November 20th, 2023 (Figure 3). The indoor temperature ranged from approximately 18°C to 30°C, which can be attributed to the scheduled operation of the ventilation system designed to maintain the climate conditions within the desired parameters. The relative humidity measurements exhibited an inverse relationship with temperature, as expected. Initially, the relative humidity showed a wider range of values, but from November 4th onwards, it stabilized, ranging between 50% and 80%.

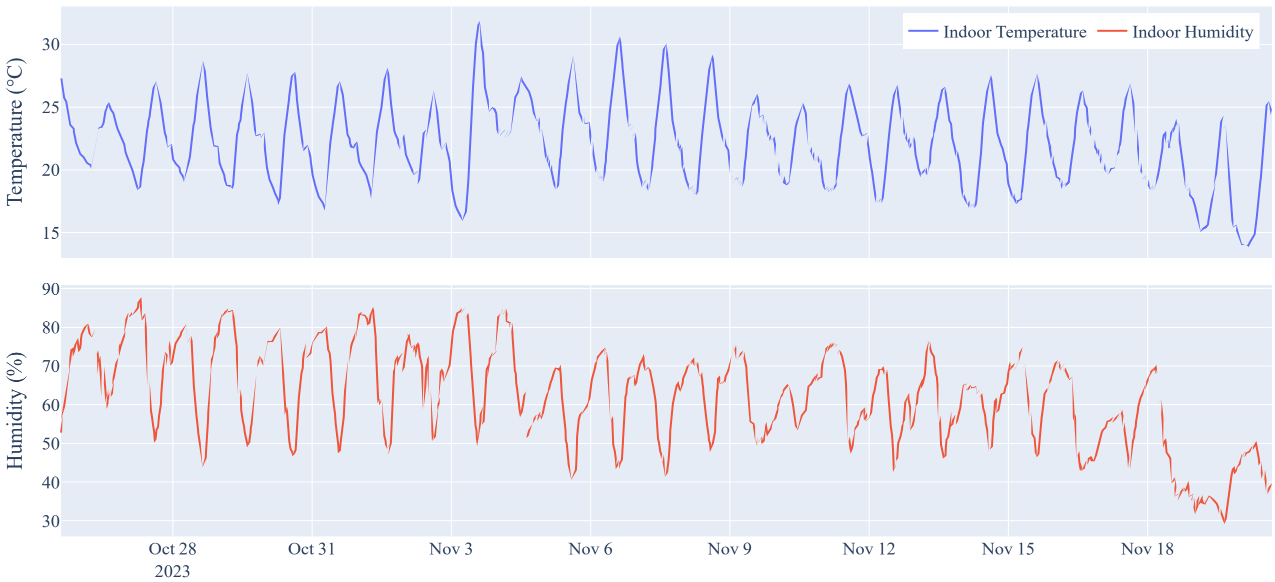

Figure 3. Temperature and relative humidity measurements.

During the initial data exploration, no outliers or missing values were observed upon visual inspection. Therefore, it was possible to proceed with the data transformation stage. In this stage, the internal temperature and relative humidity variables were used to calculate the VPD from the (Eq. 1 and **Errore. L'origine riferimento non è stata trovata.**. The absence of outliers and missing values ensured a smooth transition to the data transformation phase, ensuring the integrity and reliability of subsequent analyses.

In the training phase, the input variables for the model are indoor VPD, outdoor temperature, outdoor relative humidity and outdoor solar radiation (see as an example the Figure 4). These variables were chosen based on insights from a comprehensive literature review, which identified them as commonly used measurements in forecasting climatic variables. On the other hand, when the model has been trained and the model is used for the assessment of the indoor VPD (target variable or output variable), the necessary input variables remain outdoor temperature, outdoor relative humidity and outdoor solar radiation. Numerous studies have emphasized the significance of internal temperature, relative humidity, and external solar radiation from Open-Meteo (https://open-meteo.com/) in shaping environmental conditions within agricultural settings, particularly Vertical Farming (VF) facilities. Therefore, it is best practice to include these factors in climatic modeling and forecasting. Utilizing these variables not only guarantees access to dependable data but also improves the predictive capabilities of the model, thus enabling informed decision-making in precision agriculture and environmental management.

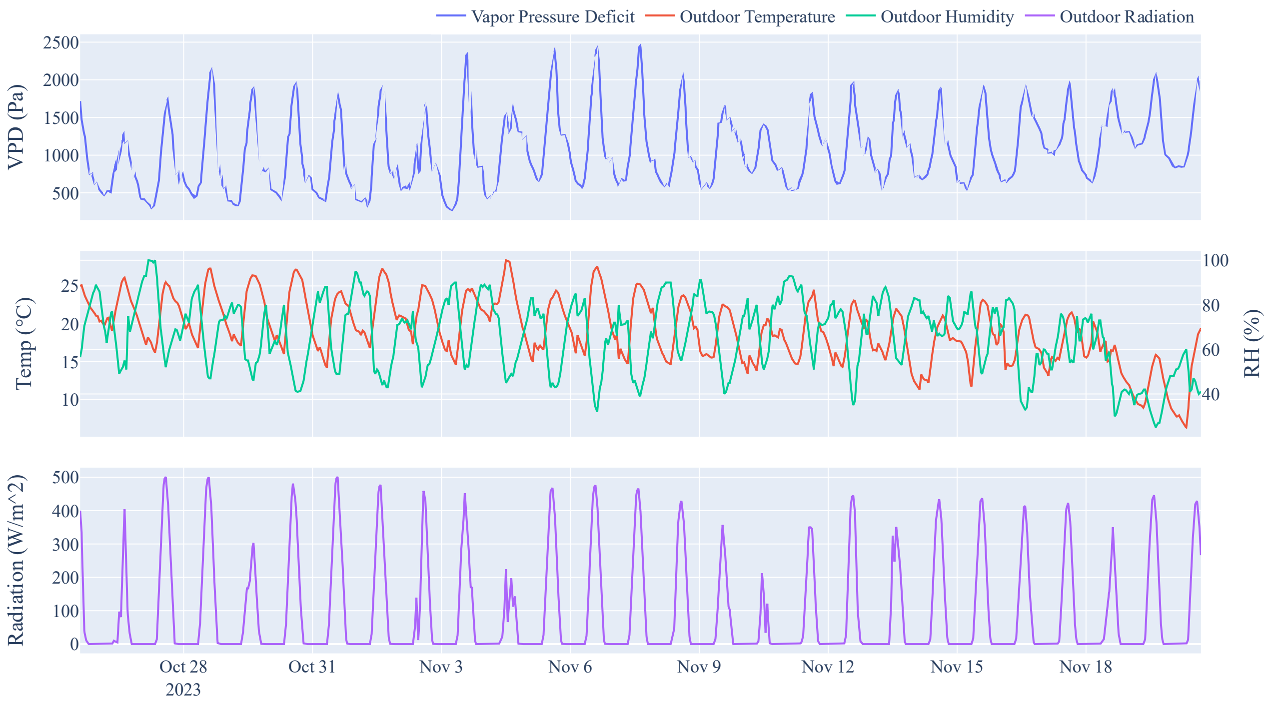

Figure 4. Input variables for the model.

Model development

Predictive models can be categorized into three main paradigms: White Box, Black Box, and Gray Box models (De Coninck *et al*., 2016; Ferracuti *et al*., 2017). Although the boundaries between these categories are blurred and often overlap, this paradigm is useful for understanding the modeling process. When the building parameters are available and the development of the dynamic model is not possible, the black box model assumes a remarkable role; therefore, this approach is chosen for the present work. So, in the current configuration the model appears as a black box model and physical and thermal properties of the walls of the container are not explicitly introduced.

The NeuralProph*et al*gorithm is well-suited for predicting significant environmental variables in indoor control conditions and optimizing plant growth. It excels in handling time series data, allowing for the incorporation of temporal patterns and dependencies inherent in VPD dynamics. This capability enables the model to capture short-term fluctuations and long-term trends, which are essential for accurate forecasting. NeuralProphet offers flexibility in modeling seasonal patterns, allowing for exploration of daily, weekly, or yearly seasonality components. This feature is particularly useful in VF environments where periodic fluctuations in environmental conditions may influence VPD dynamics. Additionally, the framework employs an adaptive learning approach, enabling the model to continuously update its parameters based on new data inputs. This adaptability ensures that the model remains responsive to changes in VF conditions, enhancing its robustness and reliability in real-world applications.

Model training

The model training process involved using the NeuralProphet framework with custom parameters obtained through a hyperparameter optimization process. These parameters were chosen to optimize the performance of the model in predicting VPD. In this study, *n_lags* was set to 6, corresponding to the previous six 10-minute intervals (representing the past hour) for each exogenous variable. This configuration enables the model to learn short-term temporal patterns relevant to VPD dynamics within the vertical farming environment. The absence of *yearly_seasonality* and *weekly_seasonality* indicates that the model does not account for yearly or weekly seasonal patterns in the data. Instead, *daily_seasonality* is set to 0.25, indicating a weak daily seasonal pattern with a period of about 4 days. This parameter choice is based on the understanding that VPD dynamics may exhibit short-term daily fluctuations. *epochs* specify the number of iterations over the entire dataset during the training process, while *learning_rate* controls the step size at each iteration, which affects the rate of model convergence. A value of 50 epochs and a learning rate of 0.025 were chosen to balance model training efficiency and convergence speed. The parameters *n_changepoints* and *changepoints_range* relate to the detection and incorporation of changepoints in the data, allowing the model to adapt to abrupt changes in VPD dynamics. A total of 20 changepoints have been considered, covering 95% of the data range. "Normalize" is set to "minmax", indicating that the input features were scaled to a range of [0, 1] to ensure uniformity and aid model convergence. *impute_missing* is enabled to handle any missing values in the dataset by imputing interpolated values. Finally, *collect_metrics* specifies the evaluation metrics to monitor during training, including mean squared error (MSE), mean absolute error (MAE), and root mean squared error (RMSE), which provide insight into the performance of the model (Torto *et al*., 2024). A summary of the main model parameters is collected in Table 1.

Table 1. NeuralProphet model parameters.

Parameter	Value	Description
n_lags	6	Number of past time steps (six 10-minute intervals) used as input features.
yearly_seasonality	False	Yearly seasonality not included.
weekly_seasonality	False	Weekly seasonality not included.
daily_seasonality	0.25	Weak daily seasonality (period of ~4 days).
epochs	50	Number of full passes over the training dataset.
learning_rate	0.025	Step size used during model optimization.
n_changepoints	20	Number of changepoints to detect abrupt shifts in the time series.
changepoints_range	0.95	Proportion of history where changepoints are allowed.
normalize	"minmax"	Input features scaled to [0, 1] range.
impute_missing	True	Enables interpolation to fill missing values.
collect_metrics	MSE, MAE, RMSE	Metrics used to evaluate model performance.

Evaluation of the model accuracy

After the training session is completed, the performance of the trained model needs to be evaluated on unseen data to assess its generalization ability. The model has been evaluated using appropriate metrics such as MAE, RMSE and a Loss function depending on the MAE. These indicators have been calculated using the following equations:

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \quad \text{(Eq. 2)}$$

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \quad \text{(Eq. 4)}$$

$$\text{Loss} = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \quad \text{(Eq. 5)}$$

where:

- n is the total number of observations.
- $y_i$is the true value of the target variable for the i^{th} observation.
- $\hat{y}_i$ is the predicted value of the target variable for the i^{th} observation.
- $\bar{y}$ is the mean of the true values of the target variable.

Observing the trajectory of the metrics (Table 2) throughout the training epochs is crucial for evaluating the model's convergence and the effectiveness of the optimization process. The model objective is to minimize the disparity between the predicted and actual values by minimizing the loss function, thereby enhancing its predictive capabilities. The evolution of the model, demonstrated by the decrease in these metrics over subsequent epochs, highlights its ability to continuously improve and enhance predictive accuracy. This thorough evaluation not only clarifies the learning process of the model but also confirms its validity and reliability for real-world use.

Table 2. Metrics of the model.

Epoch	MAE (Pa)	RMSE (Pa)	Loss (Pa)	Epoch	MAE (Pa)	RMSE (Pa)	Loss (Pa)
0	50.68142	67.88598	0.01722	15	29.62618	40.69444	0.009988
1	43.16586	58.72691	0.014537	16	29.05327	39.91445	0.009787
2	39.26909	53.77192	0.013204	17	28.77737	39.61251	0.009709
3	36.94948	50.4698	0.012429	18	28.43585	39.26314	0.009581
4	35.76921	49.09489	0.01201	19	28.18048	38.72214	0.0095
5	36.01004	48.69224	0.012106	20	27.42133	37.91353	0.009257
6	35.37132	48.09618	0.011894	21	27.21565	37.73999	0.009188
7	34.63443	47.21914	0.011637	22	26.8579	37.16541	0.009064
8	34.27697	46.87327	0.01152	23	26.74077	36.87434	0.009024
9	33.3243	45.42594	0.011203	24	26.17098	36.09068	0.008838
10	32.74167	44.48271	0.011005	25	26.0441	35.98877	0.008802
11	32.12236	43.67097	0.010817	26	25.69304	35.57893	0.008694
12	31.45604	42.90721	0.010597	27	26.01964	35.91535	0.008805
13	30.78403	41.95702	0.010362	28	25.49649	35.11007	0.008639
14	30.18878	41.43113	0.01018	29	25.27834	34.80186	0.008539

Results and Discussion

The initial forecasts generated using the trained NeuralProphet model were evaluated against unseen data, as shown in Figure 5. The model demonstrated high accuracy in both short-term (24-hour) and long-term (48-hour) forecasts, evidenced by the strong correlation between the predicted and actual values of vapor pressure deficit (VPD). The correlation coefficients for these predictions exceeded 0.9, indicating a robust model performance.

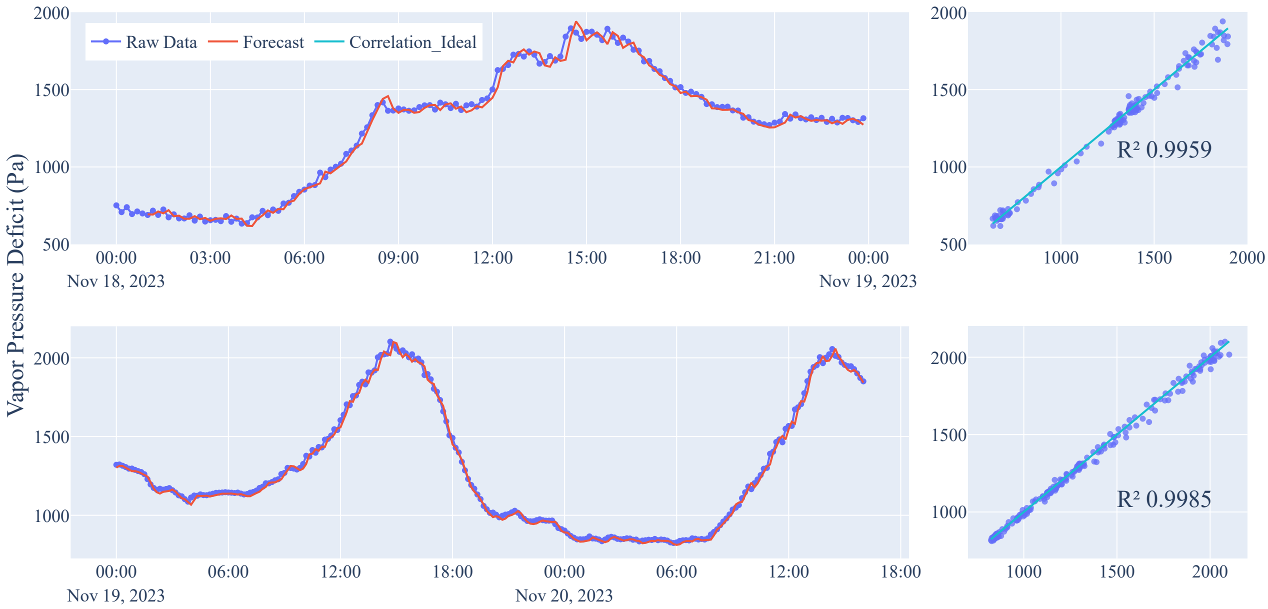

Figure 5. Comparative analysis of 24-hour (top) and 48-hour (bottom). In the right the prediction performance with Forecast *vs* Raw data visualization is reported.

The model's effectiveness was further confirmed by the low values of MAE and RMSE, which steadily decreased over the training epochs (

Table **2**). By the end of the training period, the model achieved an MAE of approximately 25.28 and an RMSE of 34.80, underscoring its capability to minimize prediction errors and improve forecast accuracy over time. While direct comparisons with previous studies must be made with caution due to differences in units, modeling approaches, and system dynamics, it is notable that (Singh *et al.*, 2018) reported an RMSE of 0.56 kPa (560 Pa) for VPD prediction in a greenhouse context. Despite the more complex and dynamic nature of vertical farming environments, our model achieved considerably lower error values using high-frequency (10-minute interval) data and a NeuralProphet framework. This underscores the efficacy of data-driven models in capturing short-term environmental variations and establishes a solid foundation for precision climate control in indoor agricultural systems.

The *plot_parameters* function in NeuralProphet provides valuable insights into the relationships between input variables and the target variable. The autoregression component (Figure 6) revealed a strong dependency on the last 6-time steps, equivalent to one hour, which is consistent with the high temporal resolution of the data. This finding aligns with the expected gradual changes in environmental conditions, where immediate past values significantly influence current conditions. In terms of future regressors, external temperature and solar radiation were identified as the most influential factors affecting VPD. These variables exhibited substantial positive weights, suggesting that increases in outdoor temperature and solar radiation contribute significantly to higher VPD values. Conversely, relative humidity showed less impact, likely due to the controlled indoor environment that mitigates its influence. This differential impact of external variables highlights the importance of considering multiple environmental factors in predictive modeling for vertical farming.

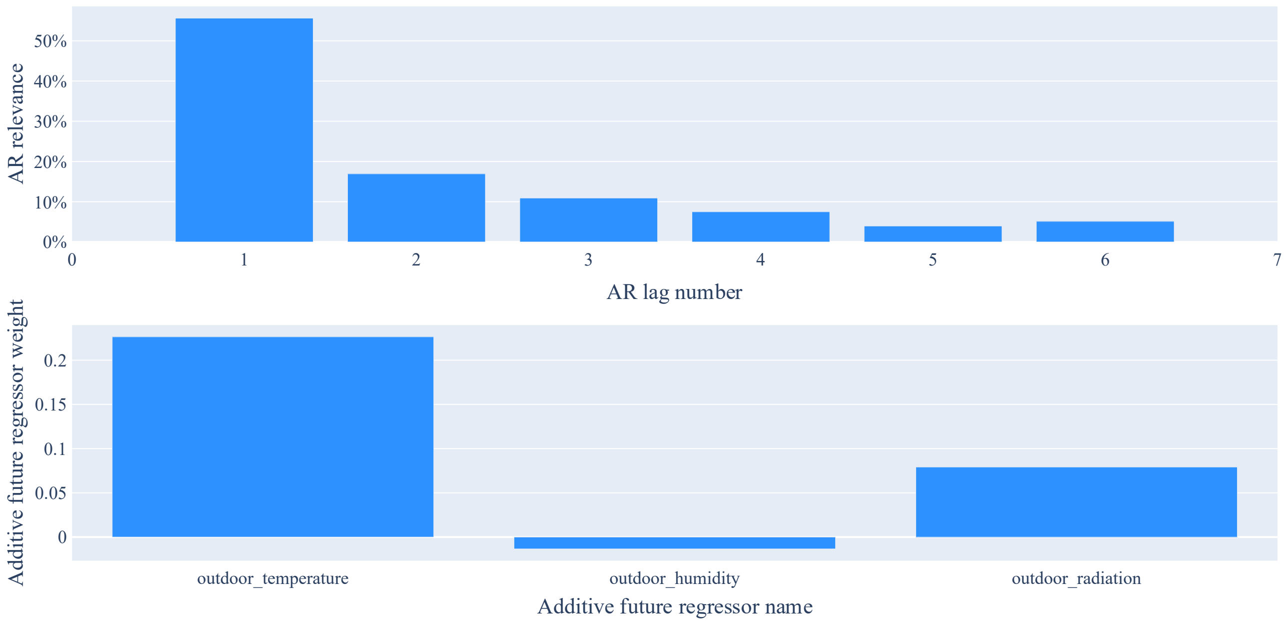

Figure 6. Plot of the parameter importance values.

The insights gained from the forecast model have practical implications for the management of VF facilities. By accurately predicting VPD, the model enables more precise control of the indoor environment, ensuring optimal conditions for plant growth. For instance, understanding the significant influence of outdoor temperature and solar radiation allows for better planning and adjustment of climate control systems, such as air conditioning and ventilation. Moreover, the ability to anticipate VPD fluctuations enhances resource efficiency by reducing the unnecessary operation of heating, cooling, and humidifying systems. This not only conserves energy but also lowers operational costs, contributing to the sustainability and economic viability of VF operations. The predictive capabilities of the NeuralProphet model thus offer a strategic advantage in optimizing environmental conditions, ultimately leading to higher crop yields and improved resources utilization.

Conclusions

The design of a forecast model for the prediction of vapor pressure deficit in vertical farming facilities using the NeuralProphet model has shown substantial progress in environmental control and resources management. High VPD forecasting accuracy, supported by strong performance metrics and insightful analysis of parameters, underlines high potential usefulness for the operation of VFs.

Among the key findings was the high predictive accuracy of the model, as evidenced by the strong correlation coefficient, along with low error metrics (MAE and RMSE) that confirm the model can predict VPD successfully under both short- and mid-term predictions. The analysis goes on to point out the external temperature and solar radiation as the major determinants of VPD and provides critical insights into which environmental factors need to be managed to maintain indoor conditions at an optimum level.

Besides, the predictive capabilities of the model make it easier to run climate control systems, resulting in energy and operational cost savings while ensuring the best-growth conditions. The successful implementation of NeuralProphet in VPD forecasting suggests the necessity of integration with other advanced AI tools for VF management. It provides exact environmental control that, with these features, allows growing high-quality crops with the least possible use of resources. The knowledge drawn from this study will be useful in developing enhanced control strategies that use predictive modeling to anticipate environmental changes and perform preventive actions. Future works may also involve more model tuning by increasing the number of environmental variables and further increasing the prediction period. In fact, a hybrid modeling approach using both neural networks and other. The NeuralProphet forecast model could support in delivering accurate actionable predictions for VPD as one of the leading models in the realization of sustainable and efficient VF practices.

CRediT authorship contribution statement

Carlos Alejandro Perez Garcia: Conceptualization, Methodology, Validation, Formal analysis, Writing – original draft, Visualization. Dafni Despoina Avgoustaki: Conceptualization, Methodology, Validation, Data curation, Writing – review and editing. Enrica Santolini: Conceptualization, Methodology, Validation, Writing – review and editing. Daniele Torreggiani: Conceptualization, Methodology, Writing – review and editing. Thomas Bartzanas: Conceptualization, Methodology, Validation, Writing – review and editing. Patrizia Tassinari: Conceptualization, Methodology, Writing – review and editing, Supervision, Funding acquisition. Marco Bovo: Conceptualization, Methodology, Validation, Data curation, Writing – original draft.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability statement

Dataset available on request from the authors. The raw data used in this study are available on request from the corresponding author due to privacy protection reasons.

References

Allen, R.G., Pereira, L.S., Raes, D., Smith, M., 1998. Crop evapotranspiration - Guidelines for computing crop water requirements - FAO Irrigation and drainage paper 56. Available from: https://www.fao.org/4/x0490e/x0490e00.htm

Avgoustaki, D.D., Vatsika, G., Giakoumatos, A., Bartzanas, T., 2024. How different daily light integrals and spectral treatments influence the development of Valerianella locusta plants grown in an indoor vertical farm. Sci. Hortic. 332:113044.

Avgoustaki, D.D., Xydis, G., 2020. Plant factories in the water-food-energy Nexus era: a systematic bibliographical review. Food Secur. 12:253-268.

Castañeda-Miranda, A., Castaño, V.M., 2017. Smart frost control in greenhouses by neural networks models. Comput. Electron. Agr. 137:102-114.

Castañeda-Miranda, A., Castaño-Meneses, V.M., 2020. Internet of things for smart farming and frost intelligent control in greenhouses. Comput. Electron. Agr. 176:105614.

De Coninck, R., Magnusson, F., Åkesson, J., Helsen, L., 2016. Toolbox for development and validation of grey-box building models for forecasting and control. J Build. Perform. Simu. 9:288-303.

Eldridge, B.M., Manzoni, L.R., Graham, C.A., Rodgers, B., Farmer, J.R., Dodd, A.N., 2020. Getting to the roots of aeroponic indoor farming. New Phytol. 228:1183-1192.

Eraliev, O., Lee, C.-H., 2023. Performance analysis of time series deep learning models for climate prediction in indoor hydroponic greenhouses at different time intervals. Plants 12:2316.

Ferracuti, F., Fonti, A., Ciabattoni, L., Pizzuti, S., Arteconi, A., Helsen, L., Comodi, G., 2017. Data-driven models for short-term thermal behavior prediction in real buildings. Appl. Energ. 204:1375-1387.

Frausto, H.U., Pieters, J.G., Deltour, J.M., 2003. Modelling greenhouse temperature by means of auto regressive models. Biosyst. Eng. 84:147-157.

Guillén-Navarro, M.A., Martínez-España, R., López, B., Cecilia, J.M., 2021. A high-performance IoT solution to reduce frost damages in stone fruits. Concurr. Comp.-Pract. E. 33:e5299.

Hauge Broholt, T., Dahl Knudsen, M., Petersen, S., 2022. The robustness of black and grey-box models of thermal building behaviour against weather changes. Energ. Buildings 275:112460.

Li, X., Zhang, X., Wang, Y., Zhang, K., Chen, Y., 2020. Temperature prediction model for solar greenhouse based on improved BP neural network. J. Phys. Conf. Ser. 1639:012036.

Morales-García, J., Bueno-Crespo, A., Martínez-España, R., Cecilia, J.M., 2023. Data-driven evaluation of machine learning models for climate control in operational smart greenhouses. J. Amb. Intell. Smart En. 15:3-17.

Oh, S., Lu, C., 2023. Vertical farming - smart urban agriculture for enhancing resilience and sustainability in food security. J. Hortic. Sci. Biotech. 98:133–140.

Oke, T.R., 2002. Boundary layer climates. Abingdon, Routledge.

Patil, S.L., Tantau, H.J., Salokhe, V.M., 2008. Modelling of tropical greenhouse temperature by auto regressive and neural network models. Biosyst. Eng. 99:423-431.

Revathi, S., Radhakrishnan, T.K., Sivakumaran, N., 2017. Climate control in greenhouse using intelligent control algorithms. Proc. American Control Conference (ACC), Seattle; pp. 887-892.

Shamshiri, R., Kalantari, F., Ting, C.K., Thorp, K.R., Hameed, I.A, Weltzien, C., et al., 2018. Advances in greenhouse automation and controlled environment agriculture: A transition to plant factories and urban agriculture. Int. J. Agric. Biol. Eng. 11:1-22.

Singh, M.C., Singh, J.P., Singh, K.G., 2018. Development of mathematical models for predicting vapour pressure deficit inside a greenhouse from internal and external climate. J. Agrometeorol. 20:238-241.

Torto, S.O G., Pachauri, R.K., Singh, J.G., 2024. Neural Prophet driven day-ahead forecast of global horizontal irradiance for efficient micro-grid management. E-Prime – Adv. Electr. Eng. Electr. En. 10:100817.

Triebe, O., Hewamalage, H., Pilyugina, P., Laptev, N., Bergmeir, C., Rajagopal, R., 2021. NeuralProphet: explainable forecasting at scale. arXiv:2111.15397v1.

Tzounis, A., Katsoulas, N., Bartzanas, T., Kittas, C., 2017. Internet of Things in agriculture, recent advances and future challenges. Biosyst. Eng. 164:31-48.

Vanegas-Ayala, S.C., Barón-Velandia, J., Leal-Lara, D.D., 2022. A systematic review of greenhouse humidity prediction and control models using fuzzy inference systems. Adv. Human-Comp. Interac. 2022: 8483003.

Yount, F.S., 2017. ASHRAE Research: Improving the Quality of Life.

Zhang, B., Xu, D., Liu, Y., Li, F., Cai, J., Du, L., 2016. Multi-scale evapotranspiration of summer maize and the controlling meteorological factors in north China. Agr. Forest Meteorol. 216:1-12.