Prediction of Greenhouse Air Vapor Pressure Deficit Based on Machine Learning

KyeongMin Jang
Information and Communication Engineering
P.O. Box 57922
Korea
jangkm0625@scnu.ac.kr

HanByeol Oh
Information and Communication Engineering
P.O. Box 57922
Korea
ohb8095@scnu.ac.kr

MyeongBae Lee
Information and Communication Engineering
P.O. Box 57922
Korea
imb@scnu.ac.kr

ChangSun Shin
Information and Communication Engineering
P.O. Box 57922
Korea
csshin@scnu.ac.kr

JangWoo Park
Information and Communication Engineering
P.O. Box 57922
Korea
jwpark@scnu.ac.kr

ABSTRACT

Vapor Pressure Deficit (VPD) in greenhouses is a factor that affects not only plant growth but also pore function and photosynthesis due to nutrient absorption of plants. In this study, the optimal model for predicting VPD was found. The correlation between VPD and environmental factors inside and outside the greenhouse was confirmed. We also check how the conditions of the data (the spacing and model of the data) affect the prediction accuracy by varying applications, and find an optimal model for predicting VPD after 20-minute. The prediction results of the various models learned showed accuracy of 0.011 for MAE and 0.008 for RMSE when 1 day data with 20-minute intervals were applied to the LGBM model. It was confirmed that the factor that had the greatest influence on VPD prediction after 20-minute was VPD (VPD_Y_71) from 20 minutes ago rather than environmental factors. Using the results of this study, it can be used not only for environmental data of greenhouses in the future, but also for various fields such as production prediction and smart farm control model implementation.

KEYWORDS

VPD:Vapor Pressure Deficit, Prediction model, Smart Farm, LGBM Ridge, RandomForest

1 INTRODUCTION

In the case of facility cultivation, it is possible to reduce the influence of external weather such as temperature, humidity, and wind, and control environmental factors necessary for crop growth under optimal conditions, so that crop yield and quality can be improved [1]. Due to these advantages, the number and area of facility cultivation facilities are steadily increasing (53,274.2ha in 2018, about 135.0×103ha in 2018), while the actual farmland area in Korea continues to decrease (2,138.0×103ha in 1988 → 1,595.4% in 2018) [2] Among the factors involved in the growth of crops and controllable in the greenhouse, Vapor Pressure Deficit (VPD) is a value calculated through temperature and relative humidity collected in the greenhouse [3]. VPD is a major driver of water transport inside plants, and a factor that affects the absorption of nutrients in plants, which affects not only plant growth but also pore function and photosynthesis due to the absorption of nutrients in plants. In addition, it is possible to determine the threat of crop diseases, cold damage, and moisture supply by using VPD values. This is because when VPD levels are high (above 2.2 kPa), photosynthesis in plants decreases, and the flow of production and water is limited, which tends to reduce nutrient absorption and accumulation [4-7]. In recent decades, temperatures have risen worldwide, increasing the average VPD exponentially. In the case of outdoor cultivation, the temperature and humidity cannot be controlled, so the optimal VPD for plant growth cannot be maintained, but the internal environmental factors for facility cultivation can be adjusted to maintain the optimal VPD of 0.4 kPa to 0.8 kPa.[8, 9] In particular, the temperature of the paprika greenhouse should be managed at 22-25°C during the day, 18-20°C at night, and 70-75% humidity, and the VPD suitable for these environmental conditions is 0.66 kPa to 0.95 kPa during the day and 0.52 kPa to 0.70 kPa at night.[10, 11]. The VPD data is also collected by the automatic greenhouse management system of the Netherlands' leading smart farm solution company, Priva, and the U.S. commercial greenhouse company Argus Control Systems operates a humidification system inside the greenhouse to control VPD [12].

In this study, we checked the correlation of VPD with various elements of the greenhouse environment, applied various analytical data conditions (data amount, spacing, and model) to determine how it affects prediction accuracy according to data conditions, and found the optimal model for predicting VPD after 20-minute.

2 DATA COLLECTION AND PREPROCESSIN

2.1 Data Collection

Greenhouse environmental data were collected from the Venlo type greenhouse located in Goheung, Jeollanam-do. Benro-type greenhouses are a type of greenhouse that is widely used due to the increased preference of farmers because they are resistant to meteorological disasters [13]. The types of crops being cultivated are paprika, consisting of internal environmental data (in_temp (temperature), in_RH (humidity), VPD (saturated water vapor), CO2 (out_temp (temperature), out_RH (humidity), wind_d1 (wind direction), wind_s (wind speed), RI (d), and a total of 11 trillion (I) per day. The data were collected at 20-minute intervals and analyzed using data collected for about 14 months (2020-12-30-2022-02-22).

2.2 Preprocessing of data

Vapor Pressure Deficit (VPD) is calculated as the difference between the amount of water (SVD) and the amount of water in the air (AVD) that can be contained when the air is saturated. Each formula is as follows.

$$SVP = 0.6108 \times \exp\left(\frac{17.27\,T}{T+237.3}\right)[kPa] \quad (1)$$

$$AVP = SVP \times \frac{RH}{100} \quad (2)$$

$$VPD = SVP - AVP \quad (3)$$

3 DATA ANALYSIS

3.1 Correlation analysis

The correlation between VPD and environmental characteristics was analyzed. Fig 1 is the result of comparing the correlation with environmental characteristics.

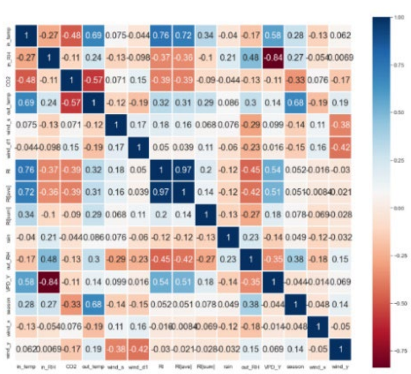

Figure 1: Correlation between VPD and Environmental Characteristics

VPD shows a positive correlation with in_temp (0.58) and RI (0.54), and a negative correlation with in_RH (-0.84). Based on the VPD calculation formula, it is interpreted that the higher the solar radiation, the higher the temperature of the greenhouse, and thus the positive correlation between the temperature and the solar radiation is similar. In the case of internal humidity, since it is inversely proportional to VPD based on the VPD calculation formula, it shows a strong negative correlation.

3.2 Create Dataset

For VPD value prediction, MinMaxScaler was used to convert the values of the data into range values between 0 and 1. Since the meaning of the past order is important because it is time series data, the data were divided into test data and train data without shuffle. The test data was analysed by dividing the total data into about 25% and the train data into 75%.

4 PREDICTION OF DATA

4.1 Changes in the time interval of data

MAE and RMSE were measured by changing the data interval at intervals of 20 minutes and 1 hour in order to compare the accuracy with the change in the data interval. As a result of calculating the MAE and RMSE values, the accuracy was about 78% higher at 20 minutes than when the interval between the data was 1 hour. It also showed about 69% higher accuracy when using the Ridge model than the Lasso model. The temperature and humidity are controlled according to the characteristics of the crops in the greenhouse, and there is a characteristic that the change in weather is not made rapidly. Due to these characteristics, it is judged that the smaller the time interval, the higher the predicted value.

Table 1: Accuracy comparison according to data spacing

Sortation	an hour's interval		20minutes interval	
	RMSE	MAE	RMSE	MAE
Ridge	0.018	0.013	0.012	0.008
Lasso	0.022	0.016	0.020	0.015

4.2 Model change

Various models were applied to the 20-minute interval data to identify models with optimal predictions. All regression models were analyzed by unifying with n_estimators=100 without applying parameters. As a result of accuracy analysis, the Ridge model, LGBM model, and Randomforest model showed high accuracy. Among them, the LGBM model showed the highest accuracy with MAE of 0.011 and RMSE of 0.008.

Table 2: Comparison of accuracy according to model change

Sortation	20minutes interval	
	RMSE	MAE
Ridge	0.012	0.008
Lasso	0.020	0.015
XGBRegressor	0.013	0.009
LGBMRegressor	0.011	0.008
RandomForest Regressor	0.012	0.008

4.3 Results

When using the LGBM models, MAE and RMSE showed the highest accuracy of 0.011 and 0.008, and the difference between the predicted value and the actual value can be confirmed through the graph in Fig. 2.

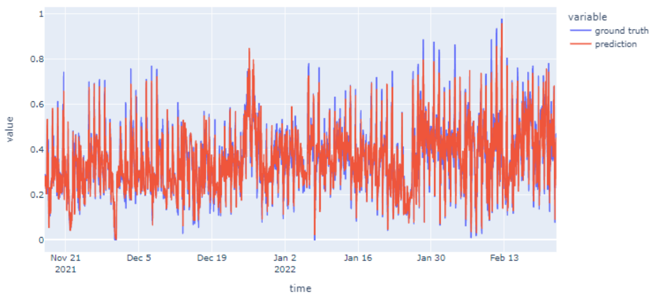

Figure 2: Comparison of predicted and actual values

The importance of the predictor was checked using Random Forest's feature_importances. It was confirmed that the most important variable for predicting the VPD value was VPD_Y_71 (VPD from 20-minute ago) and the importance was about 94%. Next, the important variables were in the order of in_RH_71, RI[ave]_71, past VPD, and in_temp_71, and it was confirmed that the humidity and internal temperature included in the VPD calculation had a slight effect.

5 CONCLUSIONS

When predicting VPD after 20-minute, various conditions (time interval of data, model) were applied to the prediction model, and as a result, the spacing of the data was greatly affected. Since the VPD in the greenhouse does not change rapidly, it is interpreted that the shorter the time interval, the higher the accuracy. In the case of the model, when the LGBM model were used, MAE was 0.011 and RMSE was 0.008, showing the highest accuracy. 20-minute, it was confirmed that the most important factor for VPD prediction was the closest past VPD value rather than the environmental factor of the greenhouse.

Currently, facility horticultural farmers are introducing smart farms, and the demand for farmers who want to help make decisions in the farmhouse system using big data is increasing [14]. Using the predictive results confirmed in this study, it is possible to implement a predictive model such as humidity and CO2 concentration as well as VPD in the greenhouse. Furthermore, it is expected that it will be possible to integrate into various smart farm systems by developing various predictive models that can streamline energy use and increase profits through energy cost reduction.

ACKNOWLEDGMENTS

This research was supported by the MSIT(Ministry of Science and ICT), Korea, under the Grand Information Technology Research Center support program(IITP-2022-2020-0-01489) supervised by the IITP(Institute for Information & communications Technology Planning & Evaluation) and Korea Institute of Planning and Evaluation for Technology in Food, Agriculture and Forestry(IPET) through Smart Farm Innovation Technology Development Program, funded by Ministry of Agriculture, Food and Rural Affairs(MAFRA) and Rural Development Administration(RDA) and Ministry of Science and ICT(MSIT)(421028-3)

REFERENCES

[1] Se-jong Oh, “A Design of Intelligent Information System for Greenhouse Cultivation”, The Society of Digital Policy and Management, Volume 15 Issue 2, Pages.183-190, 2017

[2] Moo-Been SON, Dae-Young Han, “Assessment of Feature Agricultural Land Use and Climate Change Impacts on Irrigation Water Requirement Considering Greenhouse Culitvation,” journal of the Korean Association of Greographic Informatio Studies., 23(4), 2020, pp120-139.

[3] Chiara Amitrano, Carmen Arena, Youssef Rouphael, Stefania De Pascale, Veronica De Micco, ‘Vapour pressure deficit: The hidden driver behind plant morphofunctional traits in controlled environments’, Annals of Applied Biology, Volume175, Issue3, November 2019, Pages 313-325

[4] Juping Ding, Xiaocong Jiao, Ping Bai, Yixin Hu, Jiayu Zhang, Jianming Li, ‘Effect of vapor pressure deficit on the photosynthesis, growth, and nutrient absorption of tomato seedlings’, Scientia Horticulturae, Volume 293, 5 February 2022, 110736

[5] akayasu Inoue, Motoo Sunaga, Mutsuhiro Ito, Qu Yuchen, Yoriko Matsushima, Kazuma Sakoda and Wataru Yamori, ‘Minimizing VPD Fluctuations Maintains Higher Stomatal Conductance and Photosynthesis, Resulting in Improvement of Plant Growth in Lettuce’, Frontiers in Plant Science., 01 April 2021

[6] Prenger, J.J. and P.P. Ling. 2011. Greenhouse condensation control: understanding and using vapor pressure deficit(VPD). The Ohio State University Extension Fact Sheet, AEX-804-01.

[7] Juping Ding, Xiaocong Jiao, Ping Bai, Yixin Hu, Jiayu Zhang, Jianming Li, ‘Effect of vapor pressure deficit on the photosynthesis, growth, and nutrient absorption of tomato seedlings’, Scientia Horticulturae, Volume 293, 5 February 2022, 110736

[8] Charlotte Grossiord, ‘Plant responses to rising vapor pressuredeficit’, The New phytologist v.226 no.6 , 2020, pp.1550 – 1566

[9] Heuvelink., Ep. 2005. Tomatoes. CABI Publishing. Cambridge, MA. USA

[10] Ho Cheol Kim, Sek Gi Jung, Jeong Hyun Lee, Hyang Jong Bae.(2009).Effects of Greenhouse Covering Material on Environment Factors and Fruit Yield in Protected Cultivation of Sweet Pepper.Journal of Bio-Environment Control,18(3),253-257.

[11]https://www.perfectgrower.com/knowledge/knowledge-base/vpd-chart-vapor-pressure-deficit/

[12] Shamshiri, R.; Che Man, A.; Zakaria, A.; Beveren, P.; Wan Ismail, W.; Ahmad, D. Membership function model for defining optimality of vapor pressure deficit in closed-field cultivation of tomato. In Proceedings of the International Conference on Agricultural and Food Engineering, Kuala Lumpur, Malaysia, 21 March 2017; pp. 281–290.

[13] http://www.jjn.co.kr/news/articleView.html?idxno=788176, 29.04.2022.

[14] Seong-Woo Rye, Doo-Yong Park, “Economic Analysis based on Smart Farm Colling Technology”, KIEAE Journal, Vol.21, No.5, pp.55-65, Oct. 2021