Research on Intelligent Decision-Making Irrigation Model of Water and Fertilizer Based on Multi-source Data Input

Shanshan Li[1,2,3], Yisheng Miao[1,2,3], Xiao Han[1,2,3], and Wei Guo[1,2,3](✉)

[1] National Engineering Research Center for Information Technology in Agriculture, Beijing 100097, China
guowei@nercita.org.cn
[2] Research Center of Information Technology, Beijing Academy of Agriculture and Forestry Sciences, Beijing 100097, China
[3] Key Laboratory of Digital Village Technology, Ministry of Agriculture and Rural Affairs, People's Republic of China, Beijing 100097, China

Abstract. At present, the integrated irrigation management and control system of water and fertilizer has met the requirements of automatic control of farmland water and fertilizer, gradually transforming the traditional manual operation into facility industrialization. However, this method has a weak use of data, and there is still a large gap between the calculation method and intelligent management and control. Taking greenhouse cabbage as the main research object, based on the cultivation environmental parameters, growth morphological parameters, water and fertilizer irrigation requirements during the growth period of cabbage, and using the efficient allocation ability of attention mechanism to data feature weights, this paper proposes the establishment of water and fertilizer intelligent decision-making management and control model integrating multi-source data input. The results showed that the prediction error of the intelligent decision-making irrigation model for water and fertilizer for greenhouse cabbage was relatively small, RMSE was 0.002447 m^3/Day, MAE is 0.001779 m^3/Day, and the coupling relationship between multi-source data is comprehensively analyzed, and the overall performance of model decision-making is improved through multi-feature extraction.

Keywords: Water and fertilizer · Intelligence · Multi-source data · Irrigation decision

1 Introduction

Agricultural production is the pillar of a country's economy toward self-sustainable development [1]. Over the years, water and fertilizer resource management has spanned different stages and methods in different geographical locations according to availability [2–5]. Increasing crop yield and limiting adverse environmental impacts, is a widely stated goal in agricultural science [6]. More accurate application of water and fertilizer

L. Fang et al. (Eds.): CICAI 2022, LNAI 13605, pp. 206–217, 2022.
https://doi.org/10.1007/978-3-031-20500-2_17

irrigation is regarded as one of the main means of this goal. Relative to crop demand, excessive application of nutrients will not only reduce farm profits but also degrade the environment [7]. Excessive irrigation will unsustainably deplete groundwater and streams, while excessive fertilization will rapidly reduce soil fertility, and a large number of chemical residues will endanger human health. Therefore, a large number of studies have proved that it is very important to develop a scientific decision-making and control method for water and fertilizer irrigation to better meet the needs of crops for water and nutrients at different times [8–10]. Scientific irrigation can not only effectively regulate the absorption and transfer of nutrients, but also improve the photosynthesis and transpiration of leaves, the division, and growth of plant tissues and cells, as well as the synthesis and transformation of organic substances. It can significantly promote the growth of cabbage leaf bulbs and coordinate the physiological activities of cabbage crops. To promote the healthy development of the soil micro-ecosystem, ensure the water and nutrient balance of greenhouse vegetables, and achieve high-quality production, it is very important to seek scientific irrigation and fertilization schedules.

The management and control of water and fertilizer nutrient solution have been systematically studied at home and abroad [11, 12]. At present, the regulation of nutrient solution is mainly based on the EC Value and pH value of nutrient solution at home and abroad [13], at the same time, this method is most widely used in actual production. Most studies on nutrient solution management and control system are also based on this control mode [14–18]. Some scholars use the cyclic neural network to predict the EC nutrient in the root zone of pepper in closed-loop soilless cultivation management, and have achieved good results [19]. Others use the mixed signal processing method based on the neural network to predict and analyze the ion distribution in the hydroponic solution [20]. However, the way of using multi-source data based on deep learning method to make the management decision of greenhouse cabbage nutrient solution still needs to be studied, Therefore, we propose an improved convolution neural network CNN and two-way long-term and short-term memory network AT-CNN-BiLSTM neural network model based on attention mechanism to solve this problem. The purpose is to deeply mine the existing data and accurately determine the amount of nutrient solution for water and fertilizer irrigation in real time.

2 Materials and Methods

2.1 Data Acquisition

From September 2020 to March 2021, the water and fertilizer irrigation experiment of greenhouse cabbage was carried out in the No. 14 Solar Greenhouse of Beijing National precision agriculture research demonstration base (Fig. 1). Its geographical location is 116.26 °E, 40.10 °N, with an altitude of about 50 m. The planting area belongs to a temperate continental monsoon climate, and the temperature is appropriate. The test greenhouse is 50 m long and 7.5 m wide. It is composed of an arched steel frame structure. The maximum height in the middle is about 3 m. Crops are planted in the east-west direction, and north-south direction, and covered with polyethylene trickle anti-aging film on the sunny side. A large number of intelligent control devices are set in the greenhouse. In combination with the current software and hardware configuration

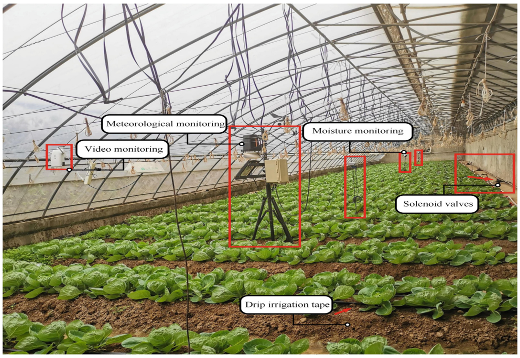

Fig. 1. Location and overview of test site.

of the greenhouse, 26 ridges were built in the test area, with an average ridge spacing of 1.5 m and a ridge top width of 0.8 m. Two rows of cabbage were planted on each ridge, with a plant spacing of 0.3 m and a row spacing of 0.5 m. The planting density was 2402 plants/667 m^2.

In the experiment, according to the different water and nutrient requirements of cabbage in different growth cycles, 12 cabbage plants with basically the same growth potential were selected for observation in each plot at the cabbage seedling stage, and the data of plant height, stem diameter (rhizome), number of leaves and width of their aboveground parts were collected day by day (among them, the stem diameter was measured by electronic vernier caliper, and the plant height and width were measured by a ruler with an accuracy of 0.01 mm). 1440 groups of cabbage growth potential data were obtained. At the same time, the WS-1800 remote meteorological monitoring station is used to monitor the greenhouse micrometeorological data. The measurement parameters include air temperature, relative humidity, wind speed, and solar radiation intensity. The collection frequency is preset to 4 pieces/minute. After the meteorological station obtains the data, the data is transmitted back through the RS485 bus, and 691200 pieces of meteorological data are selected as the basis for the construction of the greenhouse meteorological data set. The current scientific daily irrigation method of a small number of times is adopted for water and fertilizer irrigation, and the daily water and fertilizer consumption is recorded through data monitoring equipment.

2.2 Identification of Influencing Factors and Coupling Verification

This study also considers the data parameters related to internal and external factors of cabbage that are constantly changing in the process of water and fertilizer irrigation in the greenhouse. It is expected to further determine the rationality of the multi-source data set by analyzing the interaction between internal and external factors and the correlation between this factor and water and fertilizer irrigation decisions and eliminate invalid data to avoid interference with the prediction effect of the model, Improve the prediction accuracy of water and fertilizer irrigation amount of greenhouse cabbage.

External Factors. The external factors affecting the irrigation amount of water and fertilizer for greenhouse cabbage mainly include air temperature, relative humidity, solar radiation intensity, EC value, pH value and temperature of nutrient solution, and potential evapotranspiration of greenhouse cabbage. Because there are two opposite physiological processes of assimilation and dissimilation in the process of crop growth, in this process, crops continuously accumulate organic matter. When the organic matter produced by assimilation is greater than dissimilation, crops show a growth trend, and temperature has an important impact on this process. Dissolving fertilizer in water to form organic nutrient solution to irrigate cabbage. At this time, the fertility of nutrient solution is the main concern of irrigation. Appropriate EC value, pH value and irrigation temperature of nutrient solution can promote the absorption and growth of cabbage nutrients, while excessive or small amount will have a negative impact on the growth of cabbage. Generally, when the fertility is too large, we need to carry out crop leaching to avoid fertilizer damage.

Internal Factors. The internal factors that affect the irrigation amount of water and fertilizer in greenhouse are the growth morphological parameters of cabbage, such as plant height, stem diameter, leaf number and plant width. Among them, cabbage plant height refers to the straight-line distance from the root of the plant to the highest part of the plant. Excessive irrigation will lead to overgrowth of the plant and unable to support the growth of leaf bulbs. Appropriate irrigation plays a positive role in plant height growth. Stem diameter refers to the diameter of the root and stems at 1cm from the ground surface, the number of leaves refers to the number of fresh leaves surviving in the growth process of cabbage, and the width refers to the span length parallel to the ridge width of cabbage.

Coupling Relationship Verification. Collect and sort out the irrigation related data in the whole growth cycle of greenhouse cabbage, and get the initial data of different indicators (Table 1). The importance of each indicator is determined by the entropy method, and the grey correlation analysis method is used to measure the correlation between the factors that change with time.

Table 1. Data indicators.

Influence factor	Input	Evaluating indicator
External factors	x_1	Air Temperature/°C
	x_2	Elative Humidity/%
	x_3	Solar Radiation Intensity/MJ (m^2·d)
	x_4	Nutrient Solution Conductivity/mS·cm^{-1}

(*continued*)

Table 1. (*continued*)

Influence factor	Input	Evaluating indicator
	x_5	pH Value of Nutrient Solution
	x_6	Nutrient Solution Temperature/°C
	x_7	Potential Evapotranspiration/mm
Internal factors	x_8	Plant Height/mm
	x_9	Stem Diameter/mm
	x_{10}	Number of Blades/leaf
	x_{11}	Plant Width/mm

During the training process of neural network model, the potential features of data are continuously extracted through the optimization and adjustment of parameters, and this process is very time-consuming. In order to avoid the negative impact of data on the model training and prediction results, the entropy method (Table 2) and grey correlation analysis method (Table 3) were used to verify the coupling relationship between the model input data, and the input variables were analyzed through entropy calculation and correlation. Entropy method is a method to reflect the disorder degree of information in information theory. The smaller the value, the lower the disorder degree and the greater the weight; On the contrary, the larger the value, the higher the degree of disorder and the smaller the weight. Grey correlation analysis judges the close relationship between data by determining the geometric similarity between the target data column and other comparison data columns. It is usually used to reflect the influence degree between data curves, and can also be used to solve the systematic analysis of comprehensive evaluation problems. Its core idea is to determine the parent sequence according to specific rules, take the object to be evaluated as a sub sequence, and solve the correlation degree between each sub sequence and the parent sequence, And draw a conclusion.

Table 2. Calculation results of entropy method.

Index	Entropy	Order
x_1	0.059456	11
x_2	0.095377	5
x_3	0.103611	3
x_4	0.081602	8

(*continued*)

Table 2. (*continued*)

Index	Entropy	Order
x_5	0.070074	10
x_6	0.075437	9
x_7	0.082517	6
x_8	0.101482	4
x_9	0.139337	1
x_{10}	0.081913	7
x_{11}	0.109193	2

When analyzing the data coupling relationship with the grey correlation degree, if the change trends of the two elements are consistent, it means that the correlation between them is very high; On the contrary, the correlation is low. When the correlation degree ranges from 0.00 to 0.35, it means that the correlation is low and the coupling effect is very weak; When the correlation degree ranges from 0.35 to 0.45, it means low correlation and weak coupling; When the correlation degree ranges from 0.45 to 0.65, it indicates medium correlation and medium coupling; When the correlation degree ranges from 0.65 to 0.85, it indicates high correlation and strong coupling; When the correlation degree ranges from 0.85 to 1.00, it means extremely high correlation and strong coupling.

Table 3. Results of grey relational analysis.

Index	Rosette stage		Heading stage		Mature period	
	Relevancy	Coupling strength	Relevancy	Coupling strength	Relevancy	Coupling strength
x_1	0.7738	Strong	0.6741	Strong	0.7482	Strong
x_2	0.6451	Secondary	0.7358	Strong	0.7838	Strong
x_3	0.8323	Strong	0.6087	Secondary	0.6847	Strong
x_4	0.5589	Secondary	0.7899	Strong	0.7983	Strong
x_5	0.6845	Strong	0.6246	Secondary	0.7341	Strong
x_6	0.7264	Strong	0.7288	Strong	0.7699	Strong
x_7	0.5811	Secondary	0.5032	Secondary	0.6320	Secondary
x_8	0.6857	Strong	0.8015	Strong	0.7859	Strong
x_9	0.6763	Strong	0.7975	Strong	0.8236	Strong
x_{10}	0.7452	Strong	0.7140	Strong	0.7664	Strong
x_{11}	0.6842	Strong	0.7923	Strong	0.7911	Strong

3 Intelligent Decision-Making Irrigation Model of Water and Fertilizer

In recent years, the use of the attention mechanism has made remarkable achievements in image recognition and document classification. It can select the information that is the most critical to the current task target from a large amount of information, capture the sequence and mark the remote dependency between the context information. For traditional convolutional neural networks and cyclic neural network, the contribution value of each eigenvector is the same, and the difference between them is ignored. In this paper, the attention mechanism is used to predict the irrigation amount of vegetable water and fertilizer. It selectively pays more attention to some important information, and assigns the corresponding weight to the output characteristics of BiLSTM neural network, to promote the model to get better prediction results. In this paper, the CNN BiLSTM greenhouse cabbage water and fertilizer intelligent decision irrigation model (AT-CNN-BiLSTM model for short) based on attention mechanism is improved. The model mainly adds an attention mechanism layer based on CNN BiLSTM model and enriches the preprocessing analysis of multi-source data and the adjustment of parameters at each layer. Its network structure is shown in Fig. 2.

After preprocessing the greenhouse multi-source irrigation data, Xi represents the ith parameter of the input sample, and T represents the time length of the sample. The processed data set is transmitted to the CNN layer, and the convolution layer is also used for convolution operation. Here, the new features generated by the convolution layer can be expressed as:

$$C_i^k = f(x_i \otimes W_k + b_k) \tag{1}$$

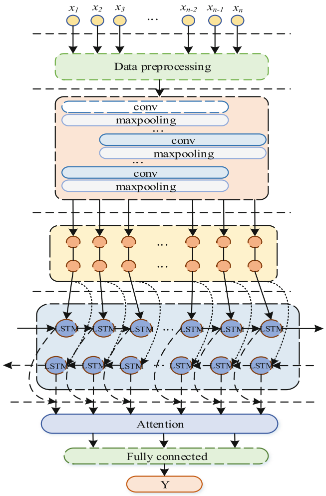

Fig. 2. AT-CNN-BiLSTM neural network structure.

Where represents the convolution operator; W is the weight vector of the convolution kernel; B is the offset term; $f(\cdot)$ represents the nonlinear excitation function. The activation function of the hidden layer adopts the relu function, which can avoid abnormal problems such as slow convergence and local maximum caused by the disappearance of the gradient. k different convolution kernels are set to promote the comprehensiveness of data feature extraction. After the convolution operation, the maximum pooling function is used for the pooling operation. The specific process can be expressed as follows:

$$C_m = \left[C_1, C_2, \ldots, C_{n-k+1}\right] \tag{2}$$

$$P_m = \max(C_m) \tag{3}$$

Taking advantage of the bi-directional feature extraction of BiLSTM, considering the interaction between each data in the greenhouse cabbage sequential growth environment data and its forward and backward data, we can use the data features at this time point to obtain the backward unit through the forward LSTM unit. In the experiment, in order to capture the long-distance dependence feature, P_m is input into the BiLSTM model, which is connected by LSTM modules in two directions and has multiple shared weights. At each time step *t,* each gate is represented by the output of the previous module and the input P_t at the current time. The three gates work together to complete the selection of attribute information, forgetting and updating of cell status.

In order to obtain more accurate prediction accuracy, the output results of BiLSTM are input to the attention mechanism layer. The attention mechanism can highlight the key features that affect the irrigation volume, reduce the impact of non key features on irrigation, help the BiLSTM layer make predictions, and will not increase the calculation and storage costs of the model. The weight calculation formula of the attention mechanism can be expressed as:

$$e_t = u_a \tanh(w_a h_t + b_a) \tag{4}$$

$$a_t = \frac{\exp(e_t)}{\sum_{j=1}^{t} \exp(e_j)} \tag{5}$$

where, h_t represents the hidden layer state vector of BiLSTM neural network at time t; e_t represents the attention probability distribution value; a_t indicates attention score; u_a and w_a are attention weight vectors; b_a is the attention bias vector.

4 Results and Discussion

The establishment of a model to predict the next day irrigation amount of greenhouse cabbage belongs to the category of regression problem. Therefore, root mean square error (RMSE) and mean absolute error (MAE) are selected as the model evaluation indicators, and the formulas are shown in Eqs. (6) and (7).

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n} (P_i - Q_i)^2}{n}} \tag{6}$$

$$MAE = \frac{1}{n} \sum_{i=1}^{n} |P_i - Q_i| \tag{7}$$

4.1 Comparison of Prediction Performance of Different Models

It can be seen from Table 4 that when different variable factors are used as model inputs, the AT-CNN-BiLSTM model proposed in this paper has the best effect, and its RMSE and Mae values are the lowest among the three neural network models, of which the RMSE variation range is 0.0024 m^3/day–0.0030 m^3/The MAE range is 0.0017–0.0025 m^3/Between days, it shows that the prediction accuracy of the improved model is effectively improved by using the attention mechanism method, which confirms the necessity of introducing the attention mechanism method.

Table 4. Comparison of prediction performance of the model under different input factors.

Model	Internal factor input		External factor input		All factor input	
	RMSE	MAE	RMSE	MAE	RMSE	MAE
BiLSTM	0.003020	0.002521	0.003432	0.002756	0.002731	0.002312
CNN-BiLSTM	0.003014	0.002583	0.003704	0.002938	0.002985	0.002269
AT-CNN-BiLSTM	0.002771	0.002405	0.002960	0.002474	0.002447	0.001779

4.2 Analysis of Factors Affecting the Prediction Performance of the Model

As shown in Fig. 3, three different time-series neural network models, BiLSTM, CNN BiLSTM and AT-CNN-BiLSTM, have shown good prediction effect on the irrigation amount per plant of greenhouse cabbage in the day after the prediction of the previous day's greenhouse cabbage growth environment data, and the irrigation trend is basically the same as the standard value. From the perspective of the factors affecting the prediction performance of the model, the multi-source data has a greater impact on the final water and fertilizer irrigation amount of cabbage. The model takes a variety of data variables as inputs, and the prediction effect is better than other inputs. The prediction result of internal factors is the second, and the effect of external factors is the worst. Due to the complexity of cabbage growth environment, the influencing factors on irrigation amount are complex and diverse. Under the condition of frequent farming operations, there are still defects in using external influencing factors alone to predict irrigation amount of water and fertilizer. However, prediction based on multi-source data can consider the coupling relationship between variables more, increase the amount of information and the effective basis for decision-making.

4.3 Prediction Error Analysis of Water and Fertilizer Irrigation Quantity

In the process of greenhouse irrigation decision-making and control, the size of irrigation error plays a vital role in cultivation management. The smaller the calculation error is, the more balanced the supply and demand is. Only in this way can crop growth be guaranteed, and water and fertilizer disasters are not easy to occur, which will negatively affect crop

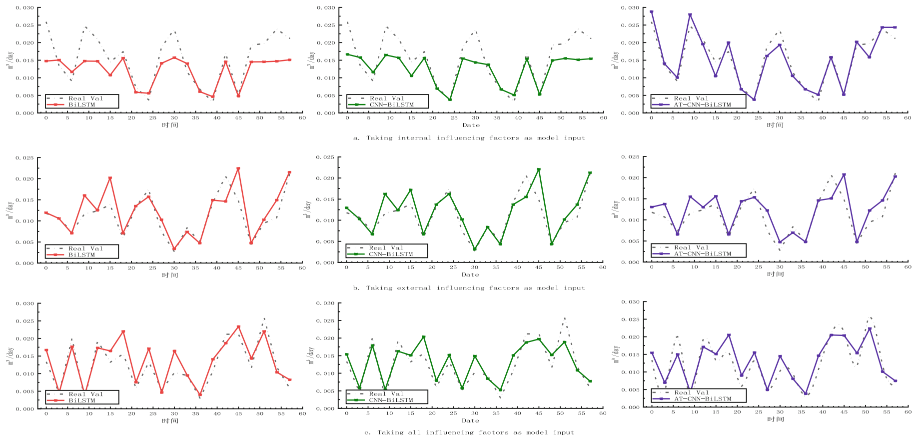

Fig. 3. Comparison between predicted and standard values of irrigation volume of different models.

production. It can be seen from Fig. 4 that the prediction accuracy of three different models with the same prediction trend and real value is more obvious. Among them, the AT-CNN-BiLSTM model proposed in this paper is better, followed by the other two models.

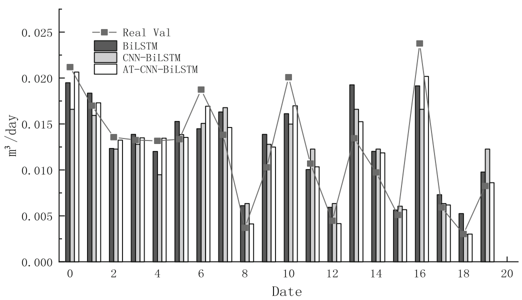

Fig. 4. Comparison of prediction accuracy of different models.

4.4 Model Stability Analysis

The AT-CNN-BiLSTM model has shown good prediction performance in the time series prediction of greenhouse cabbage irrigation. To further confirm the stability of the model and weaken the contingency of high-precision the model, BiLSTM, CNN BiLSTM and AT-CNN-BiLSTM were executed 10 times respectively, and the change range and abnormal performance of RMSE were compared. It can be seen from Fig. 5 that the RMSE variation range of the at-cnn-bilstm model is always 0.0224 ~ 0.0319 m^3/There is no abnormal value between days, indicating that the AT-CNN-BiLSTM model has the best stability. Secondly, the RMSE of BiLSTM model varies from 0.00246 to 0.00409

m^3/Day. After the CNN BiLSTM model was executed 10 times, the RMSE changed greatly and did not show good stability.

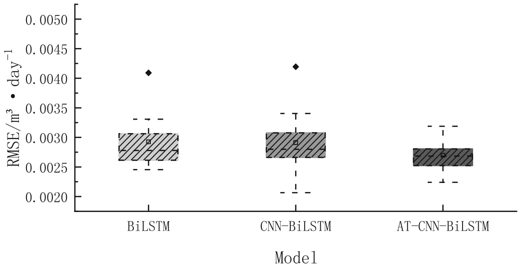

Fig. 5. Comparison of model stability.

5 Conclusion

On the premise of realizing the intelligent decision-making irrigation of water and fertilizer for greenhouse cabbage, this paper uses the front and back characteristics of original time series data and multi-scale high-level front and back characteristics to predict and analyze the irrigation amount of water and fertilizer in the joint architecture of convolution neural network and cyclic neural network, and puts forward an AT-CNN-BiLSTM neural network model with attention mechanism to predict the irrigation amount of the next day, and analyzes the application effect of the model in actual production, The validity and usability of the model are verified by experiments, which provides great help for the intelligent management of vegetable water and fertilizer.

Acknowledgements. This work was supported by Innovation 2030 Major S&T Projects of China (2021ZD0113604) and China Agriculture Research System of MOF and MARA (CARS-23-D07).

References

1. Kogo, B.K., Kumar, L., Koech, R.: Climate change and variability in Kenya: a review of impacts on agriculture and food security. Environ. Dev. Sustain. **23**(1), 23–43 (2021)
2. Zhang, Z., Yu, Z., Zhang, Y., et al.: Finding the fertilization optimization to balance grain yield and soil greenhouse gas emissions under water-saving irrigation. Soil Tillage Res. **214**, 105167 (2021).
3. Kuzman, B., Petković, B., Denić, N., et al.: Estimation of optimal fertilizers for optimal crop yield by adaptive neuro fuzzy logic. Rhizosphere **18**, 100358 (2021).
4. Zhai, L., Wang, Z., Zhai, Y, et al.: Partial substitution of chemical fertilizer by organic fertilizer benefits grain yield, water use efficiency, and economic return of summer maize. Soil and Tillage Res. **217**, 105287 (2022).
5. Xiao, H., van Es, H.M., Amsili, J.P., et al.: Lowering soil greenhouse gas emissions without sacrificing yields by increasing crop rotation diversity in the North China Plain. Field Crop. Res.**276**, 108366 (2022).

6. Bwambale, E., Abagale, F.K., Anornu, GK.: Smart irrigation monitoring and control strategies for improving water use efficiency in precision agriculture: a review. Agric. Water Manag. **260**, 107324 (2022).
7. Qin, A., Ning, D., Liu, Z., et al.: Determining threshold values for a crop water stress index-based center pivot irrigation with optimum grain yield. Agriculture **11**(10), 958 (2021)
8. Chen J, Ji X, Wu J, et al.: Review of research on irrigation decision control. In: 2021 the 3rd International Conference on Big Data Engineering and Technology (BDET), pp. 94–98 (2021)
9. Li, D., Wang, X.: Assessing irrigated water utilization to optimize irrigation schedule in the oasis-desert ecotone of Hexi Corridor of China. Agric. Ecosyst. Environ. **322**, 107647 (2021).
10. Liu, H.: Agricultural water management based on the Internet of Things and data analysis. Acta Agric. Scand. Sect. B—Soil & Plant Sci., 1–12 (2021)
11. Gallardo, M., Elia, A., Thompson, R.B.: Decision support systems and models for aiding irrigation and nutrient management of vegetable crops. Agric. Water Manag. **240**, 106209 (2020).
12. Phutthisathian, A., Pantasen, N., Maneerat, N.: Ontology-based nutrient solution control system for hydroponics. In: 2011 First International Conference on Instrumentation, Measurement, Computer, Communication and Control. IEEE, pp. 258–261 (2011)
13. Suhardiyanto, H., Seminar, K.B., Chadirin, Y., et al.: Development of a pH control system for nutrient solution in ebb and flow hydroponic culture based on fuzzy logic. IFAC Proc. Vol. **34**(11), 87–90 (2001)
14. Hyun, S., Yang, S.M., Kim, J., et al.: Development of a mobile computing framework to aid decision-making on organic fertilizer management using a crop growth model. Comput. Electron. Agric. **181**, 105936 (2021).
15. Dursun, M., Özden, S.: Optimization of soil moisture sensor placement for a PV-powered drip irrigation system using a genetic algorithm and artificial neural network. Electr. Eng. **99**(1), 407–419 (2017)
16. González Perea, R., Daccache, A., Rodríguez Díaz, J.A., et al.: Modelling impacts of precision irrigation on crop yield and in-field water management. Precis. Agric. **19**(3), 497–512 (2018)
17. Shan, B., Guo, P., Guo, S., et al.: A price-forecast-based irrigation scheduling optimization model under the response of fruit quality and price to water. Sustainability **11**(7), 2124 (2019).
18. Nguyen, D.C.H., Ascough, J.C., II., Maier, H.R., et al.: Optimization of irrigation scheduling using ant colony algorithms and an advanced cropping system model. Environ. Model. Softw. **97**, 32–45 (2017)
19. Moon, T., Ahn, T.I., Son, J.E.: Forecasting root-zone electrical conductivity of nutrient solutions in closed-loop soilless cultures via a recurrent neural network using environmental and cultivation information. Front. Plant Sci. **9**, 859 (2018).
20. Cho, W.J., Kim, H.J., Jung, D.H., et al.: Hybrid signal-processing method based on neural network for prediction of NO3, K, Ca, and Mg ions in hydroponic solutions using an array of ion-selective electrodes. Sensors **19**(24), 5508 (2019)