Prediction of crop phenology – a component of parallel agriculture management

HUA Jing[1,2], WANG Xiujuan[1,2], WANG Haoyu[1,2], FAN Xingrong[1,3], KANG Mengzhen[1,2,*]

[1]State Key Laboratory of Management and Control for Complex Systems, LIAMA, CASIA, Beijing, 100190 China
[2]Qingdao Academy of Intelligent Sciences, Qingdao, China
[3]Electronic Information and Networking Research Institute, Chongqing University of Posts and Telecommunications, Chongqing, 400065, China
[*]mengzhen.kang@ia.ac.cn

***Abstract*—Accurate prediction of crop development stages (phenolgoy) plays an important role in crop production management. Many biochemical actions of plants change with their development stage, therefore, fertilizing the plants at plants' need can increase the fertilizer use efficiency and decrease the waste. For this purpose, this paper presents a prediction model of crop phenology, describing development stages of crops as a function of the most important environmental factors, temperature and photoperiod. Machine learning is used to calibrate the crop phenology according to the predicted temperature one month later. The results indicated that our method can improve the prediction precision. The resulting model will be incorporated in a crop production management system platform developed for agriculture.**

Keywords—Crop phenology model; Prediction; Machine learning; Agriculture production management; PDT

I. INTRODUCTION

Agriculture covers a wide range of regions, involving a wide range of fields and influencing factors, complex data collection and difficult decision-making management. Therefore, precision agricultural production management is a challenging and long-term task. Precision agriculture comprises a set of technologies that combines sensors, information systems, enhanced machinery, and informed management to optimize production by accounting for variability and uncertainties within agricultural systems [1].

Precision management [2] has been proposed for agriculture in order to achieve benefits in profitability, productivity, sustainability, crop quality, food safety, environmental protection, on-farm quality of life, and rural economic development. The management of complex agricultural production system needs to integrate multi-domain knowledge, including eco-physiology, soil science, climatology, computer science, automation, etc.

In recent years, unreasonable fertilizer application leads to low nutrient utilization, waste of resources, increased costs of agricultural inputs and quality reduction in agricultural product, also cause the deterioration of soil properties, eco-environmental pollution and other issues [3, 4].

Timing of many management practices like weed and insect control, applications of fertilizer are usually determined by crop development stages (crop phenology). The stages and the duration of the stages of crops are determined by genetic factors, soil fertility and cultural practices [5]. In addition, it is also influenced by the environmental conditions of cultivation area (such as temperature and photoperiod). As autumn and winter crops grow slowly with a long growth period due to low temperature in winter; spring and summer sowing crops grow fast with a short growth period due to high temperatures. Moreover, the same species planted in different latitudes has different growth periods due to the differences of temperature and light. Phonological development of crops is an important aspect of yield formation process [6]. Hence, its better understanding requires quantitative insights in the phonological development of crop in relation to environmental factors. Furthermore, the understanding of crop phenology and its influencing factors will be helpful for cultivation managements (fertilization and irrigation) and cultivar selection.

Several numbering systems or scales have been developed for naming and describing crop stages [7, 8]. As these scales became more widely used and degree day-based growth stage (phenology) models were developed, it was apparent that combining crop staging information could forecast management practices [9]. Plant development depends on temperature. Plants require a specific amount of heat to develop from one point in their lifecycle to another. People often use a calendar to predict plant development for management decisions [10, 11]. However, calendar days can be misleading, especially for early crop growth stages. Research has shown that measuring the heat accumulated over time provides a more accurate physiological estimate than counting calendar days [9]. Each developmental stage of an organism has its own total heat requirement. Development can be estimated by accumulating degree days between the high and low temperature thresholds throughout the season [9]. The mothods mentioned above can well compute the developmental stages for some crops under some circumstances, but they are based on the history environmental data, which will be not accurate to be used to other sites and future situation. As the developmental stages vary with the temperature and photoperiod, the key issue of prediction for phenology is how to compute the phenological development according to the real-time environmental data. In recent years, many studies have been done to predict the indoor temperature using statistical analysis [12, 13], CFD model [14] or neural network [15, 16].

The objective of this study described in this paper was to develop a prediction model of crop phenology based on the prediction of greenhouse temperature with neural network, describing development stages of crops as a function of the most important environmental factors including temperature and photoperiod. The development stages and the three critical temperatures of tomato are used as an example. The resulting model will be incorporated in a crop production management system platform developed for agriculture.

II. MATERIALS AND METHODS

A. Existed crop phenology models

Accurate simulation of plant development is an important component in crop simulation model since many biochemical actions change with plant development stage [17]. Generally, crop phenology can be computed based on three methods: growing degree-day (GDD), physiological development time (PDT) and exponential sine equation. The GDD is the most traditional method to compute the development stages of crops, whose advantage is simple equation with the temperature as the variable of model input. Mohammad et al. used the GDD method to compute the planting dates for vegetables [18]. Perry et al. predicted the harvest date of cucumber using a heat unit model [19]. However, the GDD method does not consider the effect of day length on development rate. Besides, the GDD method assumes that the development rate and temperature follow the same linear relationship between the upper and lower temperature, ignoring the effect of high temperature for the development. Thus, the prediction error is big when the model is used to the other field and cultivars [7]. The physiological development time (PDT) is the duration time needed for a certain developmental stage under the optimal temperature and light conditions. For one genotype, the PDT keeps constant for some specific developmental stage [20]. Therefore, the PDT was used to simulate growth and development processes of crops and it unified the physiological time scale for different genotype, such as the simulations of development stages in tomato development [21] and cucumber [7]. Exponential sine equation is used to compute the flowering period and the effect of climate change on the flowering period [22]. Li et al. simulated the phenology of cucumber [23].

B. Model Description

1) Development stages of crops

The development of tomato can be divided into five stages: germination, seedling, flowering, fruit setting and harvest stages (Table 1). During the experiment, the development situation will be observed and the beginning date of each development stage will be recorded.

TABLE I. THE MORPHOLOGIC STANDARD OF TOMATO AT DIFFERENT DEVELOPMENT STAGES

Development stages	The standard of tomato shape
Germination	From fifty percent seed germination to first leaf appearance
Seedling stage	From first leaf appearance to first flower flowering
Flowering stage	From first flowering to first fruit setting
Fruit setting stage	From first fruit setting to first fruit harvest
Harvest stage	From first fruit harvest to ending date

2) Modelling of crop phenolgoy (PDT)

Crop development is influenced by the genetic factors and the environmental conditions (temperature and photoperiod). For the specific species, the temperature and photoperiod are the main influencing factors. The PDT is the needed time of crop growth under the optimal temperature and photoperiod transmitted from that under the real conditions. The ratio of one-day growth between the optimal and real light condition is defined as the relative photoperiod effectiveness (*RPE*). The ratio of one-day growth between the suitable and real temperature and light condition is defined as the relative physiological development effectiveness (RPDE). The PDT can be computed as the sum of RPDE for certain developmental stage. The PDT per day can be computed according to three critical points of temperature for crop development, critical photoperiod and optimum photoperiod. The ratio of one-day growth between the optimum and real temperature condition is defined as the relative thermal effectiveness (*RTE*), which is described in (1):

$$\mathrm{RTE} = \begin{cases} 0 & (\mathrm{T} \le \mathrm{T_b}) \\ \dfrac{\mathrm{T} - \mathrm{T_b}}{\mathrm{T_{ob}} - \mathrm{T_b}} & (\mathrm{T_b} < T < \mathrm{T_{ob}}) \\ 1 & (\mathrm{T_{ob}} \le \mathrm{T} \le \mathrm{T_{ou}}) \\ \dfrac{\mathrm{T_m} - \mathrm{T}}{\mathrm{T_m} - \mathrm{T_{ou}}} & (\mathrm{T_{ou}} \le \mathrm{T} \le \mathrm{T_m}) \\ 0 & \mathrm{T} > \mathrm{T_m} \end{cases} \tag{1}$$

Where T_{ob} is the lower optimum temperature; T_{ou} is the upper optimum temperature limit; T_b is the minimum temperature; T_m is the maximum temperature; T is the daily temperature defined as the average day temperature of 24 hours.

The three critical points of temperature for tomato are given in table 2 referred to [21]. The RPE can be calculated according to the day length, the critical and optimum day lengths. The RPE can be described using (2).

$$\mathrm{RPE} = \begin{cases} 0 & \mathrm{DL} \ge \mathrm{DLc} \\ \dfrac{\mathrm{DLc} - \mathrm{DL}}{\mathrm{DLc} - \mathrm{DLo}} & \mathrm{DLo} < DL < DLc \\ 1 & \mathrm{DL} \le \mathrm{DLo} \end{cases} \tag{2}$$

Where *DLc* and *DLo* are the critical and optimum day length of development. Generally, the values of *DLo* and *DLc* can be obtained by experience [24].

TABLE II. MINIMUM, OPTIMUM AND MAXIMUM TEMPERATURE OF TOMATO AT DIFFERENT DEVELOPMENT STAGE.

Development stages	Tb (ºC)	Tob~Tou (ºC)	Tm (ºC)
Germination	15	25~30	60
Seedling stage	10	25~30	35
Flowering stage	15	25~30	35
Fruit setting stage	15	20~30	35
Harvest stage	15	25~30	35

[a] Tb, Tob, Tou and Tm are the minimum, the lower optimum, the upper optimum and the maximum temperature, respectively.

$$\mathrm{DL} = 12 \cdot \left[1 + \left({}^{2}/_{\pi} \cdot \mathrm{asin}\left({}^{a}/_{b}\right)\right)\right] \qquad (3)$$

Where

$$a = \sin\lambda \cdot \sin\delta, b = \cos\lambda \cdot \cos\delta$$

$$\sin\delta = -\sin\left(\pi \cdot \frac{23.45}{180}\right) \cdot \cos\left[2\pi \cdot \frac{\mathrm{DAY} + 10}{365}\right]$$

$$\cos\delta = \sqrt{1 - \sin\delta \cdot \sin\delta}$$

Where DAY is the sequence of days (1 is 1^{st}, Jan and 365 is 31^{st}, Dec); λ is the geographical latitude; δ is the solar declination.

Generally, the germination, flowering, fruit-setting and harvest date are not influenced by the photoperiod, thus, the developmental rate is decided by the RTE for those stages. The seeding is influenced by temperature and photoperiod, the development rate is decided by the interaction between *RTE* and *RPE*. Therefore, the *RPDE* can be calculated by (4).

$$\mathrm{RPDE} = \begin{cases} \mathrm{RTE} & \mathrm{PDT} \leq \mathrm{GER} \\ \mathrm{RTE} \cdot \mathrm{RPE} & \mathrm{GER} < PDT < FLO \\ \mathrm{RTE} & \mathrm{PDT} \geq \mathrm{FLO} \end{cases} \qquad (4)$$

Where *GER* is the PDT of germination period; *FLO* is the PDT from seeding to flowering period. The values of *GER* and *FLO* are two parameters needed to be adjusted. The PDT is the sum of RPDE, which can be obtained with (5).

$$\mathrm{PDT} = \mathrm{SUM(RPDE)} \qquad (5)$$

C. Temperature Prediction

To predict crop growth period, we first need to predict the daily average temperature in the greenhouse over the next period of time. In China, most greenhouses are solar greenhouses and lack artificial regulation measures, and their environment is basically determined by outdoor weather conditions. Therefore, we can use outdoor weather data to predict future indoor temperature by machine learning algorithm.

1) Learning Alrorithm

Neural network algorithm is used in this paper. The inputs of the network are outdoor weather data, totally six dimensions, including daytime weather condition (sunny, cloudy, rain, fog, snow or cloud), night weather condition, maximum temperature, minimum temperature, maximum daytime wind speed and maximum night wind speed. The output of the network is the average temperature in the greenhouse. 10-fold cross validation was used by dividing the training data into 10 groups. Mean and variance of the prediction errors were calculated.

2) Data Source

The indoor temperature data used in this paper comes from an experiment in a solar greenhouse in Chengyang District, Qingdao. The temperature sensor was set in the greenhouse, and the temperature was collected every five minutes and uploaded to the server. The daily average of the temperature is used as the output of the neural network. The experiment lasted nearly 10 months, from November 4, 2016 to August 25, 2017. The outdoor data used in this paper is from the public data of the China Meteorological Administration. We extracted the weather data from the Qingdao area as the input of the neural network.

D. Schema of Prediction Method

To estimate the crop phenology, we need to calculate and accumulate the daily RPDE until the day at which the cumulative value reaches the PDT, which is the end of this development stage. At each day, we have the known temperatures recorded by the sensor before this day, and predicted temperatures calculated using neural network algorithm from weather forecasts after this day, thus we can use these two parts of data to predict the crop phenology. As new observations are made every day, we can update the data sets and the prediction results every day. As the known data grows more and more, the results will be more and more accurate.

III. RESULTS

A. Prediction of the Indoor Temperature

Data from March 15, 2017 to May 15, 2017 are used as testing sets. The result of the calculation is shown in the Fig.1, in which the red dotted line represents the predicted daily average temperatures and the blue solid line represents the temperatures recorded by the sensor.

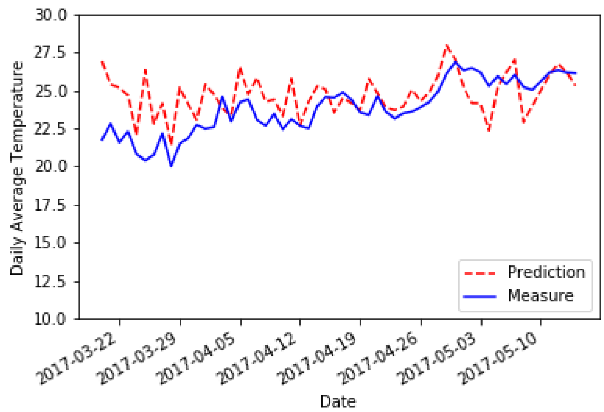

Fig. 1. Predicted and measured temperature

In the process of cross validation, we calculated the mean and standard deviation of prediction errors, which were 1.82 and 1.54, respectively. The result indicates that it is feasible to use neural network algorithm to predict the daily average temperature in the greenhouse.

B. Prediction of Crop Phenolgoy

In the prediction of the crop phenology, we also used the data from March 15, 2017 to May 15, 2017. The objective of the simulation experiment was to calculate the plant development processes of tomato, take the prediction of flowering stage as an example. The related parameters are listed in the Table 2. The results of the calculation are shown in the Fig. 2. The x-axis represents the number of days of plant growth, and the y-axis represents the cumulative RPDE value. When the cumulative value reaches the PDT, it is the beginning of flowering stage.

We can calculate a cumulative RPDE value for each day. The blue symbols are the observed data while the red ones are the predicted data. The predicted PDTs for the beginning of flowering stage were 43, 45 and 47 according to the predicted temperatures of 5 days, 15 days and 35 days, respectively.

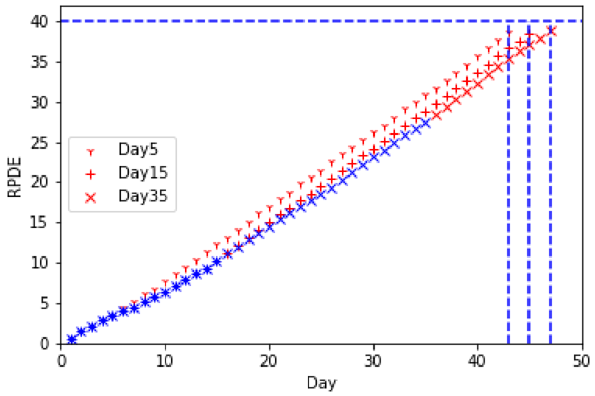

Fig. 2. Predicted PDTs of flowering time in tomato according to the predicted temperatures of 5 days, 15 days and 35 days.

C. Error analysis of crop phenology

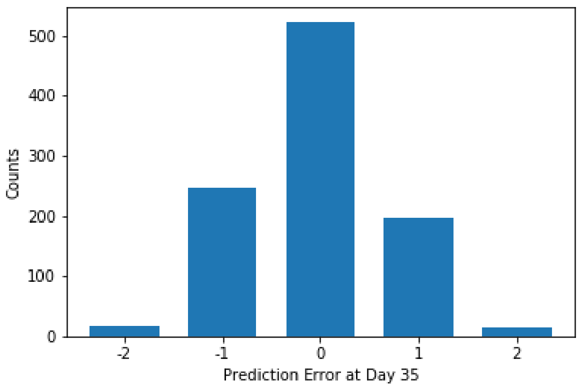

Fig. 3. Prediction Errors in simulation exprements

The crop phenology has an important role in the timing of fertilization. Therefore, it is very meaningful to make an error analysis of the calculation results in the previous section. Since the experimental period is very long, it is very difficult to analyze the errors by accumulating the data through many experiments. Instead, a simulation experiment has been made with random sampling based on the measured data of this experiment. Each measured temperature was added by a random number, which is in the range of -2.5 to 2.5 and meet the uniform distribution, and all data generated in this way were used in a simulation experiment. The error between prediction result on the 35th day and the actual PTD was calculated. A total of 1000 sets of simulation were carried out, and the histogram of the errors is shown in Fig.3. The maximum error is 2 days, and the predicated flowering day is exactly the same as the actual day in more than 50% cases. The results are quite convincing.

IV. Discussion and Conclusions

Most of existed prediction models for crop development stages are based on the GDD [10, 11, 25] or PDT methods [7, 21, 26]. They adopted the piecewise linear functions to compute the thermal time according to the temperature and photoperiod, and then compare the results with the measured values. Although this kind of models has been claimed as predict model, in fact, they are only the calibration for the model using the known environmental data. Generally, they are based on offline and static information, and even if being model based, the model is not suitable for the other situation. In this paper, first we compute the development stages based on the PDT method using the known temperature; second, the development stages were recomputed according to the real-time temperature (several days or weeks later), thus, we can predict the development stages. The method of prediction in this paper is only a simulation of crop phenology, which needs to be verified with the measurement data in the greenhouse.

Recent advances in technologies and science bring new chances for the development of smarter expert system [27-29]. Applications of IoT (internet of things) facilities are bringing much more data than ever about agricultural environment, such as the air temperature, light intensity, etc. [30]. Widely-used mobile phones are not only the tools of viewing the data but also giving the user behavior through agricultural information management system [31]. Besides, there are enormous open data source that can be achieved through Web crawler [32]. However, such data need to be deeply exploited to realize their value.

The model developed in this study can predict the development stages of crop using machine learning theory. This method should compute crop phenology more precise. Moreover, the prediction model of crop development stages according to predicted environment data can serve for cultural practices (such as irrigation and fertilization) by combined with other models, such as plant growth model, fertilization model.

Acknowledgment

This work was supported by the National High Technology Research and Development Program (863 program) of China (2012AA101906-2), and also by the National Natural Science Foundation of China (31400623, 71232006, 61233001).

References

[1] R. Gebbers and V. I. Adamchuk, "Precision Agriculture and Food Security," *Science,* vol. 327, p. 828, 2010.

[2] P. C. Robert, "Precision agriculture: a challenge for crop nutrition management," *Plant and Soil,* vol. 247, pp. 143-149, 2002.

[3] J. Zhang, Z. Li, Z. Qiao, Y. Nie, L. Di, N. Yang, and N. Yu, "The application status of farmland fertilizer in China," *Journal of Northern Agriculture,* vol. 44, pp. 118-119, 2016.

[4] Z. Zhang and W. Zhang, "The situation and trend of fertilizer application in China," *Phosphate & Compound Fertilizer,* vol. 23, pp. 9-12, 2008.

[5] Z. Starck, E. Stahl, and B. Witek-Czupryniska, "Responsiveness of Tomato Plants to Growth Regulators Depends on Light and Temperature Conditions," *Journal of Plant Physiology,* vol. 128, pp. 121-131, 1987.

[6] B. Habekotte, "A model of the phenological development of winter oilseed rape (Brassica napus L.)," *Field Crops Research,* vol. 54, pp. 127-136, 1997.

[7] Y. Li, W. Luo, J. Ni, Y. Chen, G. Xu, L. Han, and C. Chen, "Simulation of Development Stages of Greenhouse Cucumber," *Journal of Nanjing Institute of Meteorology,* vol. 31, pp. 257-263, 2008.

[8] P. G. Acosta-Quezada, T. Riofrío-Cuenca, J. Rojas, S. Vilanova, M. Plazas, and J. Prohens, "Phenological growth stages of tree tomato (Solanum betaceum Cav.), an emerging fruit crop, according to the basic and extended BBCH scales," *Scientia Horticulturae,* vol. 199, pp. 216-223, 2016.

[9] P. L. Miller, W. and S. Brandt, "Using Growing Degree Days to Predict Plant Stages," *Montana State University Extension Service,* vol. 9, p. MT00103 AG 7/2001, 2001.

[10] J. Amir and T. R. Sinclair, "A model of the temperature and solar-radiation effects on spring wheat growth and yield," *Field Crops Research,* vol. 28, pp. 47-58, 1991.

[11] Z. Xie, C. Chen, and J. Zhao, "Study on cucumber growth and development in automatic control glasshouse and its development model based on effective accumulated temperature," *Acta Agriculturae Shanghai,* vol. 23, pp. 46-49, 2007.

[12] G. Zhang, Y. Zou, and J. Wang, "Temperature Variation Rules and Prediction Model of Solar Greenhouse in West Liaoning," *Temperature Variation Rules and Prediction Model of Solar Greenhouse in West Liaoning,* vol. 29, pp. 117-122, 2013.

[13] Z. Zuo, H. Mao, X. Zhang, J. Hu, l. Han, and J. Ni, "Forecast Model of Greenhouse Temperature Based on Time Series Method," *Transactions of the Chinese Society for Agricultural Machinery,* vol. 41, pp. 173-177, 2010.

[14] W. Zhou, Y. Li, and X. Wang, "Model Predictive Control of Air Temperature in Greenhouse Based on CFD Unsteady Model," *Transactions of the Chinese Society for Agricultural Machinery,* vol. 45, pp. 335-340, 2014.

[15] H. Wang, L. Shi, H. Zhao, and Y. Yue, "Temperature Prediction Model for Sunlight Greenhouse Based on Bayesian Regularization BP Neural Network," *Hubei Agricultural Sciences,* vol. 54, pp. 4300-4303, 2015.

[16] Y. Wang, "Greenhouse Temperature Forecast Based on Improved PSO for Optimizing RBF Neural Network," *Computer & Digital Engineering,* vol. 44, pp. 1210-1215, 2016.

[17] Y. Cao, Y. Bi, J. Li, L. Gao, and M. Qu, "Verification and evaluation of development stage model for greenhouse cucumber," *Journal of China Agricultural University,* vol. 19, pp. 145-153, 2014.

[18] M. N. Elnesr, A. A. Alazba, and A. A. Alsadon, "An arithmetic method to determine the most suitable planting dates for vegetables," *Computers and Electronics in Agriculture,* vol. 90, pp. 131-143, 2013.

[19] K. B. Perry and T. C. Wehner, "Prediction of cucumber harvest date using a heat unit model," *Hortscience,* vol. 25, pp. 405-406, 1990.

[20] W. Cao and W. Luo, "Crop System Simulation and Intelligent Management," Beijing: Higher Education Press, 2003, pp. 22-31.

[21] J. Ni, W. Luo, Y. Li, J. Dai, L. Jin, G. Xu, and Y. Chen, "Simulation of the Development of Tomato in Greenhouse," *Scientia Agricultura Sinica,* vol. 38, pp. 1219-1225, 2005.

[22] P. Liancourt, L. A. Spence, B. Boldgiv, A. Lkhagva, B. R. Helliker, B. B. Casper, and P. S. Petraitis, "Vulnerability of the northern Mongolian steppe to climate change: insights from flower production and phenology," *Ecology,* vol. 93, pp. 815-824, 2012.

[23] P. Li, J. Zhou, and J. Wang, "Exponential sine equation for predicting cucumber growth process in greenhouses," *Journal of Jiangsu University,* vol. 32, pp. 262-267, 2005.

[24] J. Goudriaan and H. H. V. Laar, *Modelling Potential Crop Growth Processes* vol. 2: Springer Netherlands, 1994.

[25] Z. Shi, S. Guo, Y. Kang, H. Yin, J. Gao, G. Xu, and C. Zhang, "Simulation of greenhouse muskmelon development based on growing degree days," *Journal of Nanjing Agriculture University,* vol. 28, pp. 129-132, 2005.

[26] P. Zhang, C. He, Z. Zhang, and B. Zhang, "A Simulation Model for Tomato Plant Growth and Development in Solar Greenhouse Based on Physiological Development Time," *Chinese Journal of Agrometeorology,* vol. 27, pp. 314-317, 2005.

[27] F. Y. Wang, "Control 5.0: From Newton to Merton in Popper's Cyber-Social-Physical Spaces," *IEEE/CAA Journal of Automatica Sinica,* vol. 3, pp. 233-234, 2016.

[28] F. Y. Wang, X. Wang, L. Li, and L. Li, "Steps toward Parallel Intelligence," *IEEE/CAA Journal of Automatica Sinica,* vol. 3, pp. 345-348, 2016.

[29] F. Y. Wang, J. Zhang, Q. Wei, X. Zheng, and L. Li, "PDP: parallel dynamic programming," *IEEE/CAA Journal of Automatica Sinica,* vol. 4, pp. 1-5, 2017.

[30] X. Yan, W. Wang, and J. Liang, "Application mode construction of internet of things (IOT) for facility agriculture in Beijing," *Transactions of the Chinese Society of Agricultural Engineering,* vol. 28, pp. 149-154(6), 2012.

[31] M. Kang and F. Wang, "From Parallel Plants to Smart Plants: Intelligent Control and Management for Plant Growth " *Journal of Automatica Sinica,* vol. 4, pp. 161-166, 2017.

[32] J. C. Aker, "Dial “A” for agriculture: a review of information and communication technologies for agricultural extension in developing countries," *Agricultural Economics,* vol. 42, pp. 631–647, 2011.