Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 60 (2015) 860 – 869

Procedia Computer Science

19th International Conference on Knowledge Based and Intelligent Information and Engineering Systems

Greenhouse environmental control system based on SW-SVR

Yukimasa Kaneda[a], Hirofumi Ibayashi[b], Naoki Oishi[c], Hiroshi Mineno[a]

[a]Graduate School of Integrated Science and Technology, Shizuoka University, 3-5-1 Johoku, Naka-ku, Hamamatsu, Shizuoka 432-8011, Japan
[b]Graduate School of Informatics, Shizuoka University, 3-5-1 Johoku, Naka-ku, Hamamatsu, Shizuoka 432-8011, Japan
[c]Shizuoka Prefectural Research Institute of Agriculture and Forestry, 678-1 Toioka, Iwata, Shizuoka 438-0803, Japan

Abstract

Greenhouse environmental control systems using sensor networks are becoming more widespread and sophisticated. To match the produce of expert farmers, these systems collect data about cultivation environment and growth situation, and aim to control the environment for cultivating high quality crops. However, with no agriculture experience, it is difficult for system users to set control parameters of several devices properly. In order to reproduce prediction control performed by expert farmers' cultivation without human intervention, the authors propose a smart greenhouse environmental control system based on sliding window-based support vector regression (SW-SVR). The proposed system performs prediction control based on accurate predictions in real time. SW-SVR is a new machine learning algorithm for time series data prediction. The prediction model automatically adjusts to the current environment periodically, predicts time series data with high accuracy and low computational complexity. The proposed system using SW-SVR enables system users to optimize controls for crops. Meanwhile, since plant growth is related to the photosynthesis and transpiration of leaves, the authors developed wireless scattered light sensors which measure leaf area size indirectly so as to estimate plant growth. Our experimental results, using data of scattered light sensors on-site, outside weather data, and forecast data as independent variables of SW-SVR for hydroponic culture of tomatoes, show the proposed system reduced prediction error of nitrogen absorption amount by 59.44% as Mean Absolute Error (MAE) and 52.89% as Root Mean Squared Error (RMSE) compared with SVR, and reduced training data by 43.07% on average. Furthermore, the sugar content of tomatoes cultivated by the prototype system increased 1.54 times compared with usual tomatoes.

Peer-review under responsibility of KES International

Keywords: Support vector regression (SVR); time series data prediction; greenhouse environmental control system; scattered light sensors

1. Introduction

Greenhouse environmental control systems using sensor networks have been studied and developed increasingly [1-3]. To match the produce of expert farmers, these systems collect data about cultivation environment and growth situation, and aim to control the environment for cultivating high quality crops. However, with no agriculture

Peer-review under responsibility of KES International
doi:10.1016/j.procs.2015.08.249

experience, it is difficult for system users to set control parameters of several devices properly depending on environmental factors such as seasons and weathers. Many conventional systems use control devices based on comparison between sensor data and setting values that individual system users have set. These settings depend on the ability of the system users.

In actual fact, prediction control is done by expert farmers depending on their experiences for cultivating high quality crops, and optimizes control for crops. For example, in melon cultivation, they predict the water absorption amount based on the growth level of melons and micrometeorological data, such as air temperature, moisture, CO2 concentration, and soil moisture so as to decide appropriate amount of watering every morning. However, it is hard to reproduce expert farmers' implicit prediction from agricultural data which properties change with the lapse of time. For that reason, although many researchers have investigated fundamental researches of prediction methods which can apply to agricultural data, it is hard to apply these conventional prediction methods to environments required for accurate predictions in real time in order to perform necessary control with appropriate timing.

In this paper, we propose a smart greenhouse environmental control system based on SW-SVR [4]. The proposed system performs prediction control based on accurate predictions in real time. The purpose of this study is to reproduce prediction control performed by expert farmers' cultivation without human intervention. SW-SVR is a new machine learning algorithm for time series data prediction. The prediction model automatically adjusts to the current environment periodically, predicts time series data with high accuracy and low computational complexity.

The remainder of this paper is organized as follows. Section 2 shows related work in time series data prediction. The basic idea of the proposed system is described in section 3. Prototype implementation and experimental results are shown in section 4 and section 5 respectively. Conclusion is finally made in section 6.

2. Related work

Researchers have investigated prediction performance of time series data using machine learning [5-7]. One machine learning algorithms, Support Vector Regression (SVR) [8] which is an extension of Support Vector Machine (SVM), enables the building of a high accuracy prediction model for time series data. SVR can build a model having high generalization ability by projecting training data on higher dimensional space via kernel functions to enable linear separation and defining a discriminative plane based on margin maximization. Although several papers suggest that SVR performs well in many predictions of time series data such as the natural environmental data, continuing use of the prediction model based on training data from certain periods over a long period increases prediction error. That is because properties of time series data change with the lapse of time. For that reason, in order to continue high accuracy prediction for time series data, by applying online learning or rebuilding of the prediction model, the prediction model always has to continually learn the latest environment.

Several researches suggest that the online learning performs well in many time series data prediction. For example, Incremental and decremental support vector machine learning which applies SVR to online learning is mentioned [9]. In addition, Passive Aggressive (PA) [10] as used by Jubatus [11] attracted attention as distributed processing foundation of machine learning. In time series data predictions, training data continues to be generated with the lapse of time. For that reason, since online learning, which updates the prediction model with low computational complexity even if training data increases, always updates the latest environment rapidly, high accuracy predictions in real time is promising by using online learning. Meanwhile, conventional learning methods including online learning do not take account of the necessity of learning for each training data, learn all accumulated training data. In the natural environmental prediction, by extracting proper training data depending on properties of prediction target data, it is revealed that prediction accuracy improves [12].

In light of the aforementioned research, for high accuracy prediction of the natural environmental data such as agricultural data, we considered that it is important to continually learn the latest environment and extract minimum training data required for prediction depending on properties of prediction target data. In this paper, we propose a new machine learning algorithm which extracts data similar to prediction target data based on properties of data and automatically rebuilds itself periodically whenever environmental properties change. Extracting proper training data enables not only improvement of prediction accuracy but also reduction computational complexity. The algorithm can be applied to prediction control in an environment that requires accurate predictions in real time.

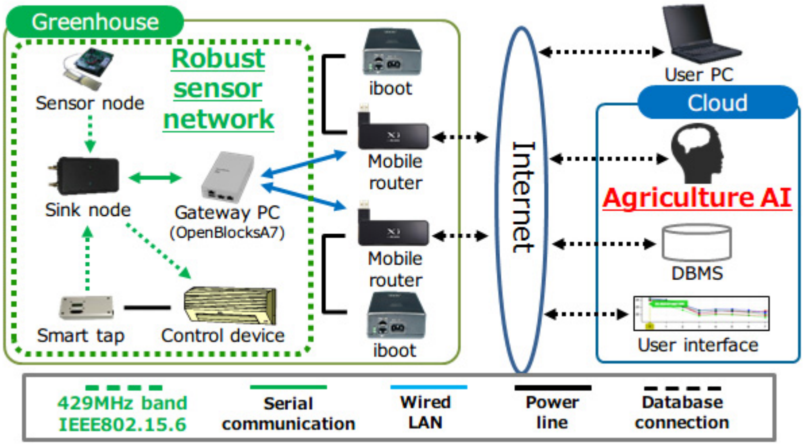

Fig. 1. Architecture of greenhouse environmental control system based on SW-SVR.

3. Greenhouse environmental control system based on SW-SVR

3.1. Overview

This paper proposes a smart greenhouse environmental control system based on SW-SVR. The proposed system performs prediction control based on accurate predictions in real time. The purpose of this study is to reproduce prediction control performed by expert farmers' cultivation without human intervention. SW-SVR is a new machine learning algorithm for time series data prediction. The prediction model automatically adjusts to the current environment periodically, predicts time series data with high accuracy and low computational complexity.

Fig. 1 shows the architecture of greenhouse environmental control system based on SW-SVR. The proposed system applies the Agricultural AI (Fig. 1 right) which performs prediction control by SW-SVR to the control part of robust sensor network system (Fig. 1 left) which aims at data arrival rate 99.99% [13]. Utilizing the cloud for the Agricultural AI reduces the load on devices to be installed in environments of high temperatures and humidity. In this way, integrated prediction analysis for cloud type service is enabled.

The operation procedure of the proposed system is shown below. Firstly, the proposed system collects environmental data definitely by robust sensor network system, then accumulates the collected data in the DBMS in a cloud. Next, the Agricultural AI predicts future environment based on accumulated data using SW-SVR, and it generates control signals based on the prediction result. Finally, the control signals generated by the Agricultural AI are delivered through robust sensor network system to on-site control devices.

3.2. Agricultural AI

3.2.1. Overview

Agricultural AI generates control signals based on prediction results of future environment and sends the control signals to control devices. By patterning predictions of expert farmers using machine learning, the proposed system can perform expert farmers' control which could not be done in conventional device control systems without human intervention.

Applying SW-SVR to prediction algorithm reproduces predictions of expert farmers properly. Since SW-SVR focuses on properties of time series data, and builds the prediction model specialized to the current environment periodically, enables a high accuracy prediction with low computational complexity. The algorithm is suitable for applications such as device control systems treating time series data.

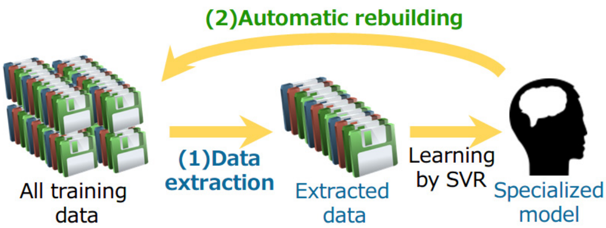

Fig. 2. Overview of SW-SVR.

3.2.2. SW-SVR

SW-SVR is an improved algorithm of SVR that specializes to time series data prediction. SVR is machine learning which applies SVM to regression algorithm, and is one of the algorithms having highest generalization ability in machine learning. Meanwhile, since the natural environment changes properties with the lapse of time, it is necessary to rebuild prediction model regularly so as to predict with high accuracy in the natural environment using SVR. However, computational complexity of SVR is from $O(n^2)$ to $O(n^3)$. For that reason, unless SVR rebuilds with less computational complexity, it is difficult to apply SVR to device control systems required for real time. SW-SVR extracts data similar to prediction target data based on properties of data, and rebuilds itself periodically whenever environmental properties change. Since extracting proper training data enables not only reduction of computational complexity in rebuilding but also improvement of prediction accuracy by removing data which becomes noise factor, SW-SVR may predict with higher accuracy than conventional prediction methods that lack extracting training data.

Fig. 2 depicts the overview of SW-SVR processing. SW-SVR extracts appropriate training data from all training data by data extraction (Fig. 2(1)), and rebuilds the model with automatic rebuilding function (Fig. 2 (2)) when environmental properties change with the lapse of time. In data extraction, SW-SVR focuses on the regularity of time series data so as to extract appropriate data. Meanwhile, because the prediction model built from extracted training data is specialized to current environment, prediction accuracy of the model deteriorates with changing environmental properties with the lapse of time. Thus, in automatic rebuilding, SW-SVR calculates prediction accuracy of the model periodically. When prediction accuracy is less than a threshold, SW-SVR considers that environmental properties changed with the lapse of time and rebuilds the prediction model. Although most prediction accuracy cannot be calculated, that of time series prediction is calculated by comparing prediction value with measured values sequentially. How to extract data and define the threshold is described later.

The computational complexity (CC) of SW-SVR is as follows: CC of SW-SVR for one model building is determined by both CC of the training data extraction algorithm and CC of SVR. As mentioned later, CC of training data extraction is $O(n)$. Meanwhile, CC of SVR is from $O(n^2)$ to $O(n^3)$. For that reason, when n' cases of training data was extracted from n cases of all training data, if the extraction rate is very low ($n' \ll n$), CC of SW-SVR is $O(n)$. Meanwhile, in the case of other values ($n' < n$), it is from $O(n^2)$ to $O(n^3)$. Except for the case with the very high extraction rate ($n' \approx n$), SW-SVR reduces the computational complexity compared with SVR.

3.2.3. Data extraction

Fig. 3 depicts the overview of Short distance Data Collection (SDC). SDC collects the training data within the range of radius r centered on the latest data G. Then, SDC can extract the data similar to the latest data required for high accuracy prediction. In time series data, the data similar to the latest data is also similar to changing environmental properties with the lapse of time. Thus, by using the data similar to the latest data only, we consider that the model specialized to current environment is built.

Formula (1) is the extraction condition of SDC, where $d(A, B)$ is the distance between data A and data B, G is the latest data, and P is the training data. The formula describes that SDC extracts only training data P_n where the distance with the latest data G is less than radius r. Distance index d can be defined variously as 'euclidean distance', 'mahalanobis distance', 'manhattan distance', etc, while the data index should be chosen depending on prediction

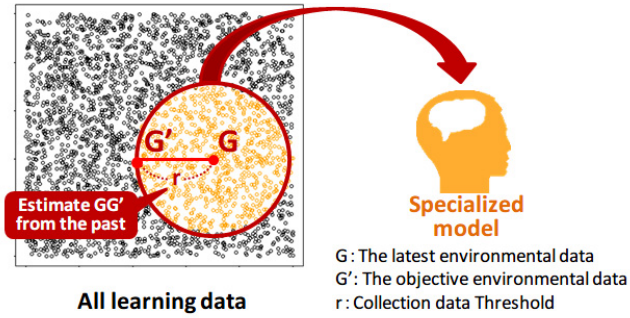

Fig. 3. Overview of SDC.

objects. Meanwhile, regardless of the distance index, SDC standardizes training data to remove the differences of the scale of each independent variable.

$$r > d(G, P_n)\ . \tag{1}$$

The determination of radius *r* with SDC is described below. SDC defines radius *r* as distance between the latest data *G* and *G'* after prediction time from *G*. By defining radius *r* as *d*(*G, G'*), when the change of environmental properties is gradual, SDC extracts less training data. On the other hand, when the change of environmental properties is rapid or prediction time is long, SDC extracts more training data. Thus, SDC extracts training data properly depending on the difficulty of the prediction. However, *d*(*G, G'*) cannot be calculated directly because data *G'* is not observed at prediction point in time. For that reason, when the movement of data *X* is defined as distance between data *X* and data *X'* after prediction time from *X*, SDC estimates *d*(*G, G'*) (the movement of *G*) from movements of training data near *G*. Since training data near *G* has similar properties, the movement of *G* shows also similar values to the movement of close data. So that SDC defines radius *r* as a weighted mean of the movements of each training data *P*, when weights are applied to a reciprocal of *d*(*P, G*). By applying the reciprocal of the distance to weights, movements of training data similar to the latest data greatly affects the value of *r*.

Formula (2) shows how to determine the radius *r* in SDC. This formula is a weighted mean, and weight *W* is defined by **formula (3)**, where *N* is the amount of training data, *P'* is training data after prediction time from *P*, and *p* is to decide strength of the weight. However, when the amount of training data extracted by formula (2) is extremely little, prediction model cannot be built from the training data. In that case, radius *r* is calculated by **formula (4)** that traditional determine method of radius *r* [14].

$$r = \frac{\sum_{k=1}^{N} W_k d(P_k, P_k')}{\sum_{k=1}^{N} W_k}\ . \tag{2}$$

$$W_x = \frac{1}{d(P_x, G)^p}\ . \tag{3}$$

$$r = \sqrt{\frac{\sum_{k=1}^{N}\{d(P_k, G)\}^2}{N}}\ . \tag{4}$$

SDC evaluates radius *r* from 3 phases. In first phase, SDC decides the individual weight variables W_x for each training data. A weight variable W_x becomes higher as $d(G, P_x)$ become lower. In second phase, SDC finds the movement of each training data P_x. For example, if a prediction model predicts after one hour, SDC evaluates how each training data moves in 1 hour ($d(P_k, P'_k)$). In third phase, SDC determines the radius *r* using weight variables and movements of each training data.

3.2.4. Automatic rebuilding

SW-SVR rebuilds the model as properties changes when prediction accuracy is less than a threshold. It is necessary to consider the threshold for model rebuilding because determination of the threshold is connected directly with performances of SW-SVR. When environmental properties change with the lapse of time, a prediction target *G'* will go outside of the circle in Fig. 3. Then SW-SVR should rebuild a prediction model. If the prediction target goes outside of the circle, we consider that the prediction error for the prediction target exceeds that for all training data in the circle. Thus, SW-SVR predicts all extracted training data used for the model building, and the prediction error of the prediction is used as the threshold. This value means the prediction accuracy of data within a circle constructed by SDC. By applying the result of the prediction to the threshold, SW-SVR enables to determine whether the point *G'* is outside the circle.

3.3. Robust sensor network system

Robust sensor network system which is the base of the proposed system surely collects environmental data required for predictions by wireless communication [13]. Generally, although wireless sensors are used to reduce wiring costs with device control systems based on sensor data, reliability of wireless communication may fail due to environmental conditions. By the way, because prediction is based on collected data, it is important to collect data certainly to control using prediction. The system achieves the robust wireless communication at greenhouse environment with many obstacles by using a low frequency band wireless communication having high wavelength and long diffractive.

As an evaluation of the system, the researchers compared the 2.4GHz band wireless communication which was frequently used in the past with the 429MHz band wireless communication adopted by robust sensor network system. As a result, it was revealed that the packet arrival rate declines with growth of plants in a conventional wireless communication standard using 2.4GHz band. Meanwhile, wireless communication standard using 429MHz band adopted robust sensor network system maintains high packet arrival rate regardless of growth of plants. Thus, the system with 429MHz band wireless communication can collect data required for predictions certainly.

4. 1. Prototype implementation

4.1. Nitrogen absorption amount prediction control system using the scattered light sensors

A Prototype of greenhouse environmental control system based on SW-SVR is implemented. By focusing on nitrogen absorption which is a function of plants, the prototype collects environmental information at greenhouse by using the scattered light sensors that get strength of light scattered in the air and some sensors that get the data required for nitrogen absorption amount. The prototype supplies nitrogen depending on a prediction result of the nitrogen amount that tomatoes absorb on the control day. When nitrogen is supplied every day, by determining the nitrogen amount which should be supplied by Agricultural AI and maintaining appropriate supply of nitrogen, high quality crops are cultivated.

Summary data for scattered light sensors implemented for prototype is described below. Scattered light sensors collect strength of light scattered in the air. The sensors measure leaf area size indirectly to estimate plant growth, being installed on top and bottom of plants. Since plant growth is related to photosynthesis and transpiration of leaves, scattered light sensors are useful for prediction of plant growth such as the amount of photosynthesis and transpiration. In fact, amount of transpiration is estimated at error less than 3% by using data that scattered light sensors collects [15].

Fig. 4 shows the on-site appearance, and **Fig. 5** describes placement of sensors. In the cultivation method of this experiment, a pump connected to a supplying tank scoops solution in the tank, then the solution drains into cultivation beds of plant. The excess solution is recycled in the supplying tank again. The prototype system controls an electromagnetic valve connecting a nitrogen tank and a supplying tank. It decides the nitrogen supplying amount based on a nitrogen absorption amount prediction performed at 6:00 in morning, and it opens the electromagnetic valve only at the time when it is necessary to supply the amount. In this experiment, by focusing on the feature of plants that the time to absorb solution from roots is the daytime, the prototype predicts nitrogen absorption amount of six hours at six o'clock every morning and controls based on the prediction result.

Table 1. Independent variables and dependent variable of training data.

Items	Contents	
Independent variables	On-site	Air temperature, humidity, leaf area index, Nitrogen concentration, sensing time
	Meteorological data	Air temperature, wind speed, rain fall, solar radiation
	Forecast	Air temperature prediction until six hours later, Wind speed prediction until six hours later
Dependent variable	Nitrogen absorption amount of six hours	

Fig. 4. On-site appearance.

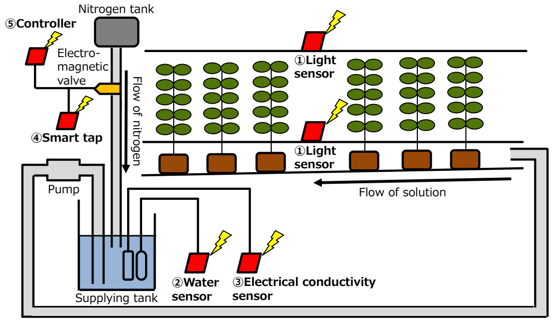

Fig. 5. Placement of sensors.

4.2. Implement of Agricultural AI

Table 1 depicts independent variables and a dependent variable for prediction of nitrogen absorption amount. In this implementation, the proposed system predicts a nitrogen absorption amount of six hours as dependent variable. Independent variables are applied to the items which are generally related to nitrogen absorption and the evaluation criteria that expert farmers use for control. Particularly, since there is connection between transpiration and nitrogen absorption, scattered light sensors which estimate the amount of transpiration with high accuracy have significant effect on prediction of nitrogen absorption amount. Meanwhile, because nitrogen absorption amount and the amount of transpiration are also related to the weather, we consider that wide area environmental data is necessary to express weather six hours later. Therefore meteorological data collected by Japan Meteorological Agency [16] and weather forecasts collected by Japan Weather Association [17] are also used for independent variables.

An operation procedure of a prototype is indicated below. First, it predicts nitrogen absorption amount of six hours at six o'clock and supplies nitrogen of the amount in proportion to prediction result into the tank. Therefore, it determines necessity of rebuilding at twelve o'clock that is six hours later of the prediction. At that time, if it decides that rebuilding is necessary, it rebuilds by applying SDC to the data observed at six o'clock in the next morning.

5. System evaluation

5.1. Experiment contents

The proposed system was evaluated by building prediction models of nitrogen absorption amount for both methods: SVR as a conventional method and SW-SVR as a proposed method. The prediction accuracy and computational complexity of both is compared. Further, the utility of the prediction control with SW-SVR is evaluated by calculating sugar content from fruits harvested after the experiment conclusion.

Table 2. Experiment periods.

	Learning	Evaluation
Period	2014/10/3 ~ 2014/11/25	2014/11/26 ~ 2015/2/1
Data number	4,148	54

Table 3. Parameters of SVR and SW-SVR.

Setting items	SVR	SW-SVR
Method	Epsilon -SVR	Epsilon -SVR
Kernel function	RBF	RBF
Cost parameter	10000	1
Hyperparameter of RBF	0.1	0.0001
Tube parameter	1 (default)	1 (default)
Weight parameter of IDW	—	1

Table 2 shows a learning period and an evaluation period. However, when SW-SVR rebuilds, it also uses training data in the evaluation period generated already as of each rebuilding. Meanwhile, in order to evaluate actual prototype use, the prediction target of the experiment is data observed at six o'clock and twelve o'clock when prediction control is performed. For that reason, the number of evaluation data of the prediction targets is 54 cases.

Table 3 shows parameters used at this evaluation. Parameter tuning for SVR is decided using four-fold cross-validation. On the other hand, tuning for SW-SVR is decided by dividing learning period into two and predicting the newer one after learning the older one, while adjusting various parameters. The reason why we avoid using cross-validation in SW-SVR is because it is improper to use parameters decided by cross-validation for one prediction model, because SW-SVR handles a lot of prediction models by rebuilding.

The indexes of the prediction error is applied to Root Mean Square Error (**Formula (5)**) and Mean Absolute Error (**Formula (6)**), when N is the amount of evaluation data, y is a true value, and $\hat{y}$ is a dependent variable.

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N}(y_i - \hat{y}_i)^2}{N}} \quad . \tag{5}$$

$$MAE = \frac{\sum_{i=1}^{N}\left|y_i - \hat{y}_i\right|}{N} \quad . \tag{6}$$

Target prediction accuracy in the experiment is indicated below. In order to confirm if SW-SVR enables to change expert farmers' implicit knowledge to explicit knowledge, the prediction accuracy of the algorithm for nitrogen absorption amount estimated by expert farmers is set as the target prediction accuracy. The prediction accuracy is 14.4 as RMSE and 10.56 as MAE. Although permitted delay is not apparent when controlling nitrogen for the plants, in order to apply the proposed system to various systems from now on, target prediction time is one minute, which is the measurement cycle of the sensors used by the prototype, so that the proposed system can predict for all measured sensor data.

5.2. Result

Fig. 6 shows true values and predicted values in the evaluation period, and **Fig. 7** describes prediction error for all prediction algorithms for the whole evaluation period. From figures, SW-SVR followed changes of nitrogen absorbed amount more than SVR. It was confirmed that SW-SVR reduced a prediction error by 59.44% as MAE and 52.89% as RMSE in the whole evaluation period. Moreover, by comparing the prediction error of each algorithm, SW-SVR showed the prediction error of SW-SVR very close to that of target compared with SVR. Meanwhile, **Fig. 8** depicts the training data reduction rate. SW-SVR reduced the amount of training data at average 43.07% and most 91.47% more than SVR and contributes to real time improvement.

The reason that SW-SVR reduced the prediction error more than SVR is indicated below. SW-SVR adjusted changes of nitrogen absorption more than SVR. In particular, after Jan/1/2015 when the properties changed substantially, SW-SVR manages changes of nitrogen absorption amount, but SVR cannot. That is because SW-SVR

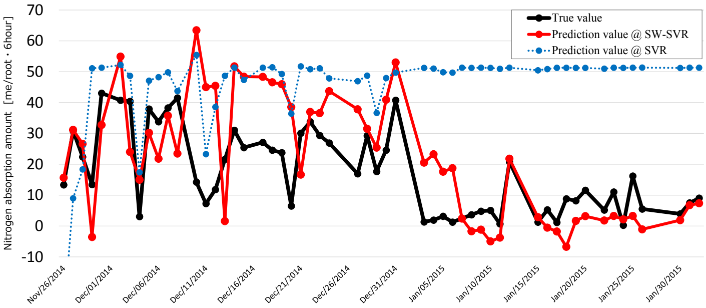

Fig. 6. True value and prediction value.

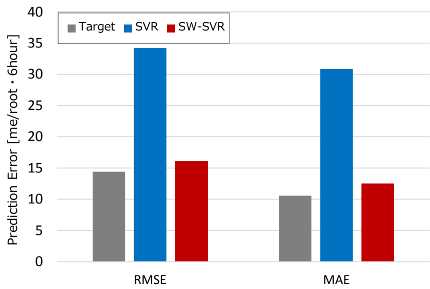

Fig. 7. True value and prediction value.

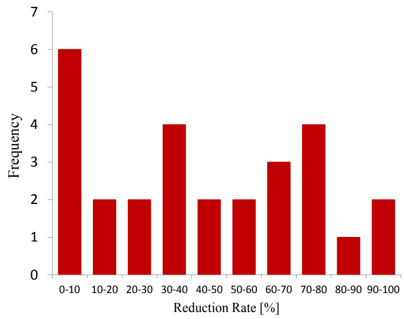

Fig. 8. Reduction rate.

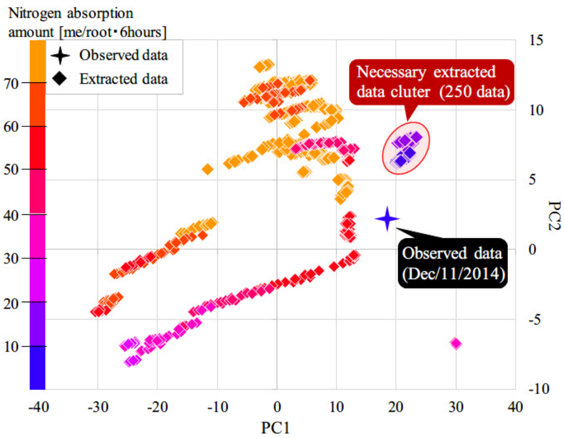

Fig. 9. Observed data and extracted data by SW-SVR.

can learn the latest environment by rebuilding. For that reason, it is important to rebuild with the natural environmental prediction which changes the properties.

However, for some periods the prediction error increased by using SW-SVR. Among them, the prediction result of Dec/11/2014, the prediction error for SW-SVR increased substantially compared with that of SVR. We reveal the reason that SW-SVR increased the prediction error in the period by analyzing training data used for the prediction in Dec/11/2014. **Fig. 9** indicates difference in properties of the observed data at Dec/11/2014 and the training data extracted by SW-SVR. Fig. 9 converted each data into the two-dimensional space by Principal Component Analysis (PCA), Fig. 9 expresses the overview of each data property. SW-SVR extracted not only 250 training data as necessary training data but also 818 training data. SW-SVR increased the prediction errors in the period because of learning unnecessary training data for the predictions in spite of extracting training data.

The reason why apposite extraction training data is not performed is because the necessary independent variables are lacking, we consider that environment similar to the environment of the prediction object is not expressed adequately. Although SDC estimates the similarity by calculating the distance between independent variables of observed data and ones of training data, relationship of the distance and the similarity is lost when appropriate independent variables is not given. For example, when nitrogen concentration was not used for independent variables in this experiment, since an environmental condition was changed rapidly by exchanging solution at Jan/1/2015, it is difficult to distinguish observed data after Jan/1/2015 from data before Jan/1/2015. Thus, SDC extracts the training data of environment which is not similar to current environment. In future, we aim to further improve the prediction accuracy by considering the items which are relevant to the nitrogen absorption amount.

Finally, the utility of the prediction control by SW-SVR is evaluated. As a result, the fruits cultivated with the proposed system had 1.53 times of sugar content compared with usual fruits, and the proposed system cultivated the fruits with additional value. In future, further improving the prediction accuracy aims for improvement of quality.

6. Conclusion

In this study, in order to reproduce prediction control performed by expert farmers' cultivation without human intervention, we propose a smart greenhouse environmental control system based on SW-SVR. The proposed system performs prediction control based on accurate predictions in real time. By applying new SW-SVR methodology for micrometeorological data prediction, the proposed system ensured both high accuracy and real time prediction which device control systems for agriculture require. The prototype system collected environmental information such as plant growth using wireless scattered light sensors that measure leaf area size indirectly. Experimental results using the prototype system, show SW-SVR reduced prediction error of nitrogen absorption amount by 59.44% as MAE and 52.89% as RMSE compared with SVR, and reduced training data by 43.07% on average. Furthermore, the sugar content of tomatoes cultivated by the prototype system increased 1.54 times compared with usual tomatoes.

In future work, we elucidate independent variables required for agricultural data prediction. In particular, when agricultural dependent variables such as nitrogen absorption amount are utilized, it is difficult to reveal the required independent variables. We aim to reproduce expert farmers' cultivation accurately by considering structures that can extract expert farmers' implicit knowledge so as to reveal appropriate independent variables. Meanwhile, apart from nitrogen absorption, we take into account the items such as photosynthesis and soil moisture.

Acknowledgements

This study was partially supported by the funding for 2013-2014 MIC SCOPE and JSPS KAKENHI (26660198), Japan.

References

[1] Kim, Y., et al., "Software design for wireless sensor-based site-specific irrigation," Computers and Electronics in Agriculture, vol.66, pp.159-165, 2009.

[2] Park, D.H., et al., "Wireless sensor network-based greenhouse environment monitoring and automatic control system for dew condensation prevention," MDPI Sensors, vol.11 (4), pp.3640-3651, 2011.

[3] Khan, R., et al., "Wireless sensor network based irrigation management system for container grown crops in Pakistan," World Applied Sciences Journal, vol24 (8), pp.1111-1118, 2013.

[4] Suzuki, Y., et al., "Proposal to sliding window-based support vector regression," International Conference on Knowledge-Based and Intelligent Information & Engineering Systems (KES2014), pp.1615-1624, 2014.

[5] Sapankevych, N.I., et al., " Time series prediction using support vector machines: a survey," IEEE Computational Intelligence Magazine, vol.4 (2), pp.24-38, 2009.

[6] Guo, Y., et al., "Time series prediction method based on LS-SVR with modified gaussian RBF," Neural Information Processing, LNCS, vol.7664, pp.9-17, 2012.

[7] Elattar, E.E., et al., "Electric load forecasting based on locally weighted support vector regression," IEEE trans. on Systems, Man, and Cybernetics, vol.40 (4), pp.438-447, 2010.

[8] Vapnik, V., "The nature of statistical learning theory," Springer, 1999.

[9] Cauwenberghs, G., et al., "Incremental and decremental support vector machine learning," Advances in Neural Information Processing Systems 13: Proceedings of the 2000 Conference, Vol. 13, MIT Press, 2001.

[10] Koby C., et al., "Online passive-aggressive algorithms," The Journal of Machine Learning Research 7, 551-585, 2006.

[11] Jubatus : Distributed Online Machine Learning Framework, Jubatus[Online], Available <http://jubat.us/en/> (Accessed Apr/01/2015).

[12] Suzuki, Y., et al., "Analysis of support vector regression model for micrometeorological data prediction," Computer Science and Information Technology, vol.3 (2), pp.37-48, 2015.

[13] Ibayashi, H., et al. "Highly reliable wireless environmental control system for home gardening," Consumer Electronics (GCCE), 2014 IEEE 3rd Global Conference on. IEEE, 2014.

[14] Suzuki, Y., et al., "SW-SVR improved by short-distance data collection method," Mobile Computing and Ubiquitous Communications (MBL), vaol.2014MBL_73 No.9, pp.1-8, 2014.

[15] Oishi, N., "Development of the scattered light multi-sensor unit to get greenhouse environment and plant ecology information," Japanese Society of Agriculture Biological and Environmental Engineers and Scientists 2014.

[16] Japan Meteorological Agency, Japan Meteorological Agency[Online], Available <http://www.jma.go.jp/jma/indexe.html> (Accessed Mar/01/2015)

[17] Japan Weather Association, Japan Weather Association[Online], Available < https://www.jwa.or.jp/english/> (Accessed May/01/2015)