Modeling and simulation of agriculture production system considering seasonal variable information using big data analysis

Yusaku MATSUMOTO*, Hironori HIBINO**, Makoto KIMURA*** and Yosuke MIZUKAMI****

* Department of Industrial Administration, Graduate school of Science and Technology, Tokyo University of Science
  2641 Yamazaki, Noda-shi, Chiba 278-8510, Japan
** Department of Industrial Administration, Faculty of Science and Technology, Tokyo University of Science
  2641 Yamazaki, Noda-shi, Chiba 278-8510, Japan
*** HATAKE Company., Ltd.
  1470-5 Nishitakano, Tsukuba-shi, Ibaraki 300-2613, Japan
**** AGROPOLIS LLC.
  2675-1 Notsumachinotsuichi, Usuki-shi, Oita 875-0201, Japan

Received: 26 March 2019; Revised: 24 August 2019; Accepted: 23 October 2019

Abstract
The agricultural sector is required to produce a steady, year-round supply of crops to maintain the quality of life of consumers. The organization of agricultural production has, therefore, become increasingly important. This study aims to evaluate the agricultural production systems for storage shortages and quality loss while considering seasonal changes in crop growth. We first discuss the occurrence of storage shortages and quality loss. Next, we investigate the production and crop growth processes in fields used by an agricultural production corporation. We examined the statistical treatment of growth and the relationship between crop growth and the field environment using the data collected from field sensors. We then analyzed the data obtained in the observation. Based on the results of the analysis, we proposed a model for the work process in the field, which is a prediction algorithm for crop growth that considers seasonal changes. We then implemented the simulation model. We verified the accuracy of the proposed growth prediction algorithm through a case study and used the simulation to evaluate storage shortages and quality loss.

Keywords: Agricultural production system, Baby leaves, Crop growth, Soil temperature

1. Introduction

The agricultural sector is required to produce a steady, year-round supply of crops to maintain the quality of life of consumers. The organization of agricultural production has, therefore, become increasingly important. Several agricultural corporations have developed organizational systems with contributing factors such as the flow of crop supply and information related to demand and production control, where agriculture is considered as a production system. Cropping methods can be classified into three types, monocropping, double cropping, and multiple cropping (Sanseido, 2006). Crops grown by the monocropping and double cropping methods are generally treated as storage type crops (e.g., rice and potatoes). Crops cultivated by the multicropping method are year round multi rotational crops (annual crops) that are supplied on a daily basis and are regarded as flow type crops. Cultivation of the year-round, multi-rotational crops is a recurring process. Hence, it is necessary to space out the sowing dates and use multiple fields for crop growth. Further, the date of sowing, choice of sowing field, and sowing amount for year round, multi rotational crop fields are often determined empirically. For these reasons, production management technology is becoming a crucial factor in the cultivation of annual crops to ensure efficient
supply. The stable supply of crops has been made possible by the research carried out in plant factories, but the cost of supply is rather high (Takatsuki, 2010). Therefore, in agricultural production systems, plant factories are used only to a limited extent, and year-round crops are typically supplied through greenhouse cultivation or open-field cultivation. However, uncertainties in supply may arise in both greenhouse cultivation and open-field cultivation owing to the uncertainties in crop growth in the fields. Moreover, consumer demand for crops fluctuates depending on the price and the appearance of the crop (Datari, 2004), leading to a discrepancy between demand and supply. Typically, there may be variations in crop growth in different fields, and hence the period from sowing to harvesting or the amount of harvest may not be in line with predictions, resulting in storage shortages or quality loss. Quality loss occurs when a product that is meant for shipping grows (or ripens) beyond the stipulated standard due to the passage of time and therefore becomes unfit for shipping.

An information network system that stores and manages the data of greenhouse cultivation and open-field cultivation has been proposed in earlier studies on the agricultural production system (Hoshi et al., 2013; Takayama, 2013; Zhang et al., 2016). Using portable electronic devices and the Ubiquitous Environment Control System, Hoshi et al. (2013) developed a simple and inexpensive system that stores air temperature and humidity data, the amount of sunshine, and CO2 concentrations. Studies on using the accumulated data in agricultural production systems have also been conducted (Matsukura et al., 2015; Togami et al., 2011). Togami et al. (2011) analyzed information about the changes in the field environment and variations of the quality of oranges and studied the relationship between them. Matsukura et al. (2015) proposed an advanced evaluation technique using simulations to study the effect of the expansion in the scale of rice farming and the change in technology and skills of farmworkers on cost reduction. These studies have been used to research quality improvement and cost reduction using field information network systems and accumulated data. However, research on agricultural production systems has not progressed enough to tackle the uncertainties in supply, hindering the improvement of system efficiency. Therefore, there is an urgent need to develop methods to assist with production operations in advance based on the production schedule, evaluation of storage shortages, and quality loss. In an earlier study (Hibino et al., 2018), a simulation method was developed to anticipate storage shortages and quality loss in annual crops by considering production processes in the agricultural production system such as sowing and harvesting in the field and the uncertainties in the growth process. However, earlier studies did not consider the seasonal changes in crop growth, which depend on the seasonal changes in the field environment (Peng et al., 2017). In agricultural production systems, the period from sowing to harvesting varies with the seasonal changes in crop growth, and the number of harvests and the yield also vary accordingly. Storage shortages and quality loss also vary with the season, depending on the volume of the harvest. Thus, it is crucial to evaluate the agricultural production systems for storage shortages and quality loss, considering the seasonal changes in crop growth.

Therefore, the aim of the present study is to evaluate agricultural production systems for storage shortages and quality loss, considering the seasonal changes in crop growth. We first discuss the occurrence of storage shortages and quality loss. Next, we investigate the production and crop growth processes in some of the fields used by an agricultural production corporation. We examined the statistical treatment of growth and the relationship between crop growth and the field environment using the data collected from field sensors. We then analyze the data obtained in the observation. Based on the results of the analysis, we propose a prediction algorithm for crop growth that considers seasonal changes. We then implement the simulation model and verify the accuracy of the proposed growth prediction algorithm through a case study, then use the simulation to evaluate storage shortages and quality loss.

2. Analysis of storage shortages and quality loss in the field

2.1 Causes of storage shortages and quality loss in the field

We first analyze the occurrence of storage shortages and quality loss. When there are no variations in crop growth, and the date and amount of the harvest take place as expected at the time of sowing, and there will be no storage shortages or quality loss. If there are variations in crop growth, the causes of storage shortages and quality loss in a given field can be summarized as follows.

Causes 1: If the harvested amount on the harvest day is less than the amount estimated at the time of sowing, there will be a storage shortage.

Causes 2: If the harvested amount on the harvest day is more than the amount estimated at the time of sowing,
adjustments may have to be made by reducing the amount harvested from other fields and delaying the harvest date, leading to quality loss.

**Cause 3:** If the growth value indicating the degree of growth (e.g., the size) on the scheduled harvest date does not meet the shipping standard, and hence the crop cannot be harvested, a storage shortage will occur.

**Cause 4:** If the growth value on the scheduled harvest date exceeds the shipping standard, and hence the crop cannot be harvested, quality loss will occur.

Cause 1 and Cause 2 occur because the harvest amount per unit area (unit yield) varies with the variations in crop growth. Cause 3 and Cause 4 occur because the state in which the growth value meets the shipping standards changes with the variations in crop growth.

### 2.2 Factors necessary to evaluate storage shortages and quality loss

Storage shortages and quality loss are caused by variations in the harvest amount and the growth state. The amount of harvest depends on the yield value, which indicates the amount of harvest per unit area. Therefore, the harvest amount varies because the yield value varies. The state in which the crop meets the shipping standards varies with the growth value, which indicates the degree of growth for yearly crops. Growth value depends on the number of days elapsed after sowing, and the speed at which the growth value changes after sowing depends on the change in the field environment, which varies with the season. Therefore, to evaluate the storage shortage and quality loss, it is necessary to investigate the growth process of the crop and the impact of seasonal changes in the field environment. In this study, we analyzed the relationship between the elapsed time after sowing and the field environment, the yield value, and the degree of growth (e.g., the growth value) for several seasons.

In the agricultural production system, crops are supplied by operating multiple fields. In the case of year round, multi rotational crops, many fields are maintained simultaneously, and the frequency of sowing, the sowing field, and the sowing amount are often determined empirically. It is necessary to consider field operations when evaluating storage shortages and quality loss. If, for instance, no suitable field is found at the sowing time, it could result in a future storage shortage. Therefore, the required number of holdings should be considered. In the operation of the field, the production process of the field units must be determined, and sowing must be assigned according to the operational or non-operational status of each field. In other words, for any one rotation of cultivation, a series of production processes, such as sowing, harvesting, and preparation for the next sowing after harvesting, must be specified. The relationship between the state of each of these production processes and their transitions needs to be identified because the production process and growth process are very closely related.

### 3. Field observation and analysis

In this section, we investigate the crop production and growth processes using fields that produce baby leaves. Baby leaf is an annual crop and is a generic term for young leafy vegetables harvested two to three weeks after germination. The target crop is the Pinogreen (young Komatsuna), which has a high content of baby leaves (HATAKE Company, 2016).

Observations of the production process were conducted from September 1, 2014, to May 11, 2016. Observations of the growth process were conducted from April 11, 2016, to October 23, 2017. Table 1 lists the observation conditions. The investigation of the growth process was carried out on a seasonal basis to observe the seasonal changes in the growth of the yearly crop. During the rainy season in June and July, the growth value of year-round crops is different from that in the other seasons. These two months are therefore regarded as the rainy season, and the year is divided into five seasons, spring, rainy season, summer, autumn, and winter. In this observation, we considered three seasons, rainy season, summer, and autumn.
Table 1 Field observation conditions

| Field owner       | HATAKE Company, Ltd. |
|-------------------|----------------------|
| Location          | Yoshinuma, Tsukuba, Ibaraki Prefecture |
| Area of house     | 230 $m^2$           |
| Target crop       | Baby leaf (Pino green) |
| Observation period (production process observation) | 12/September/2014–11/May/2016 |
| Observation period (growth process observation) | [Rainy season] 20/June/2016–18/July/2016 [Summer] 1/August/2017–31/ August /2017 [Autumn] 23/September/2017–23/October/2017 |

3.1 Production process

The production process in the field was investigated to verify the state of the field from sowing to harvesting and during the transition. The production process in the field consists of two states, the field state and the crop state. The field state itself consists of five states, plow, waiting for sowing, cultivation, reaping, and no cultivation. The crop state consists of three states, growth, suitable harvest, and unsuitable harvest. If a supply instruction occurs when the crop is in growth, it is not possible to meet the requirements, and storage shortages occur. If the crop is in an unsuitable harvest state, it does not meet the shipping standard, and quality loss occurs.

The states of the field and crop states are listed as follows.

[Field state]
- Plough: Preparing the field environment for crop growth
- Wait for sowing: Waiting state from completion of curing until sowing time
- Cultivation: The time from sowing to harvesting, when the crop is present in the field
- Reaping: Harvesting of crops grown in the field
- No cultivation: Waiting state from the end of harvesting to the time of the next curing

[Crop state]
- Growth: The state after sowing until the crop growth value meets the shipping standard
- Suitable harvest: The state in which the growth value of the crop meets the shipping standard
- Unsuitable harvest: The state in which the crop has grown beyond the shipping standard

Figure 1 shows the production process and the growth process in the cultivated field.

![Production and growth processes in a cultivated field](image)

3.2 Observation of Growth Process

The field observations were conducted using sensors, while crop growth was manually measured to determine seasonal changes in yield and growth value after sowing (Fig. 2). The growth value, yield value, and field
environmental data were measured to formulate the relationship between the elapsed time after sowing and those measured values for each season. The measurement data was stored in a cloud system (Fujitsu F-SASS, 2016). The field environmental data acquired by the sensors includes humidity, air temperature, soil temperature, watering amount, and the amount of sunlight. Each sensor made measurements every 30 min for 30 days. The growth and yield values, which are difficult to obtain with sensors, were manually measured for 50 samples for 30 days.

3.2.1 Measurement of growth value

The growth value was measured to formulate the relationship between the elapsed time after sowing and the growth value for each season. The HATAKE Company determines the growth value of the baby leaves by measuring three parts of the leaf. The first part is the length from the tip of the leaf to the point where the leaf begins to spread, called the leaf blade. The second measurement is from the point where the leaf begins to spread to the point where it sprouts, called the petiole. The third measures the largest span of the leaf spread, which is known as the leaf width. As indicated in Fig. 3, the value of leaf blade is denoted by $\alpha$, the value of petiole by $\beta$, and the value of leaf width by $\gamma$. In this study, the growth value was evaluated based on the evaluation criteria of HATAKE, i.e., as $1.5 \times \alpha$. The value of $\alpha$ was measured in units of 0.1 cm. The measured growth values in the autumn of 2017 are shown in Fig. 4, which reveal the daily variations in the measured growth values.

---

Fig. 2: An outline of crop growth and the field environment investigation using the data collected from field sensors

---

Fig. 3: Measurement objects of leaf
3.2.2 Measurement of yield value

The yield values were measured to formulate a relationship between the growth value and the weight. The yield value measurement is investigated by the number of stocks observed per unit area. The number of stocks is obtained by investigating seven plots of the target field, each with an area of 1 m², and calculating the mean value of these seven plots. Figure 5 shows the relationship between weight value and growth value of the leaves investigated in seven plots of the target field. The weight value refers to the weight of the leaves in each stock of 1500 samples (50 samples*30 days). The yield value per unit area [kg/m²] is derived from the number of stocks (the stock density [stock/m²]) and the weight value per stock [g/stock] using Eq. (1).

\[
U_{st} = W_{st} \times R \times 10^{-3}
\]

\(U_{st}\): Yield value [kg/m²] in season \(s\), \(t\) days after sowing

\(W_{st}\): Weight value [g/stock] in season \(s\), \(t\) days after sowing

\(R\) : Stock density [stock/m²]
3.2.3 Field Environment

In order to clarify the relationship between the field environment and the growth value, the field environment and growth value were simultaneously measured.

Table 2 shows the correlation coefficient between the field environmental data obtained using sensors and the actual growth value in each season. The effective accumulated soil temperature is the value obtained by summing the daily average soil temperatures above the minimum soil temperature and below the maximum soil temperature required for growth. The minimum soil temperature required to grow Pino green is 5.0 °C, and the maximum soil temperature is 35.0 °C (Kumamoto Prefecture Vegetable Promotion Association, 2012). During the observation period, there were no occurrences of temperatures below 5.0 °C, while periods when the temperature was above 35.0 °C often occurred. Therefore, the period in which the soil temperature was higher than 35.0 °C was excluded from the effective accumulated soil temperature calculation. In this observation, the soil temperature was measured in 30 min increments. In Table 2, there are two items with high correlation coefficients, the accumulated air temperature and effective accumulated soil temperature. These measurements show a relationship between the field environment and growth value. Table 2 shows a high correlation coefficient between the measured growth value and the effective accumulated soil temperature, as well as the accumulated air temperature during the three seasons.

Table 2 Correlation coefficients between field and crop growth

|                | Accumulated air temperature | Air temperature | Amount of sunlight | Soil moisture | Effective accumulated soil temperature | Soil EC | Amount of watering |
|----------------|-----------------------------|-----------------|--------------------|--------------|----------------------------------------|---------|-------------------|
| 2016 Rainy season | 0.95                        | -0.27           | 0.22               | -0.09        | 0.96                                   | -0.33   | -0.15             |
| 2017 Summer     | 0.96                        | -0.05           | 0.04               | -0.45        | 0.96                                   | -0.19   | -0.23             |
| 2017 Autumn     | 0.95                        | -0.27           | 0.04               | -0.41        | 0.96                                   | -0.27   | 0                 |

4. Field analysis

The field data obtained using sensors were analyzed to develop an advanced evaluation simulation of the agricultural production system for the assessment of storage shortages and quality loss, considering the seasonal changes in crop growth.

4.1 Generation of growth curves

In order to facilitate the mathematical modeling of year-round crop growth, the daily measured growth value was averaged for each season, and an actual growth curve was generated from the daily average of the actual growth value. The growth curve is typically approximated by a logistic function (Tozaki et al., 2011; Matsuda et al., 2011; Sato et al., 2002). Therefore, the measured growth curve was approximated by the logistic function using the least-squares method. The general equation of the logistic function is shown in Eq. (2).

\[
h = \frac{K}{1 + be^{-ct}}, \tag{2}
\]

where \( t \) denotes the number of elapsed days since sowing and \( h \) is the growth value of the crop. The three coefficients of Eq. (2), are defined as follows: \( K \) determines the convergence value of the growth curve, \( b \) determines the initial value of growth, and \( c \) determines the slope of the logistic function. The growth curve approximated by the logistic function is called the approximate growth curve. Table 3 lists the parameters of the logistic function in the approximate growth curve of each season. The correlation coefficient and the contribution ratio of the actual growth curve and the approximate growth curve in each season are presented in Table 4, which shows that the contribution ratio is over 0.96. This means that the measured growth curve can be accurately approximated by the logistic function. Figure 6 shows the measured growth curve and the approximate growth curve for the autumn of 2017.
Table 3 Parameters of the logistic function in the approximate growth curve

|          | 2016 Rainy season | 2017 Summer | 2017 Autumn |
|----------|-------------------|-------------|-------------|
| K        | 10.94             | 11.06       | 10.56       |
| b        | 340.12            | 160.79      | 1414.17     |
| c        | 1.58              | 1.39        | 2.00        |

Table 4 Correlation coefficient and contribution ratio between the measured growth curve and the approximate curve in each season

|          | 2016 Rainy season | 2017 Summer | 2017 Autumn |
|----------|-------------------|-------------|-------------|
| Correlation coefficient | 0.98 | 0.99 | 0.99 |
| Contribution ratio       | 0.96 | 0.98 | 0.98 |

Fig. 6 Measured growth curve and approximate growth curve

4.2 Variations in measured growth value

The observation of the measured growth values shows variations in daily growth. To analyze these variations, we treated the data as normally distributed. Specifically, we examined the variations in the measured growth value at K, which is the convergence value of the approximate growth curve in each season, and at K/2 in the approximate growth curve.

Normality was tested using the D’Agostino-Pearson test, which calculates the test statistic $k^2$ from the skewness and kurtosis of the distribution, to detect normality when the sample size is small. When the test value $k^2 \leq 5.291$, the data follows a normal distribution at the 5% significance level (D’Agostino et al., 1990).

Table 5 shows the growth value variations for each season and the corresponding $k^2$ value at K of the logistic function. Since the values in Table 5 are less than or equal to 5.291, the variations in the growth value at K/2 and the convergence value K follow a normal distribution.

Table 5 Test result of growth value K/2 and convergence value K

|          | 2016 Rainy season | 2017 Summer | 2017 Autumn |
|----------|-------------------|-------------|-------------|
| Growth value K/2 | 0.190 | 0.61 | 1.40 |
| Yield value K | 0.077 | 0.40 | 0.58 |
4.3 Relationship between growth value and weight value

To generate accurate predictions of growth and yield values in the simulation, it is necessary to link these two values. In other words, once a predicted growth value and the predicted growth curve are generated by the simulation, the predicted yield value must be generated in accordance with those values. Linking the predicted growth and yield values requires clarification of the relationship between the actual growth value and the crop weight.

Figure 7 shows the relationship between the growth value and the yield value measured in the field. In Fig. 7, the relationship between the measured growth value and the measured weight value is represented by a quadratic function in which the x-axis represents the measured growth value and the y-axis represents the measured yield value. The relationship between the actual growth value and the actual weight value enables a highly accurate prediction of the yield value. The quadratic function representing the relationship between the actual growth value and the actual weight value is shown in Eq. (2). Table 6 shows the correlation coefficient between the approximate curve and the measured value as well as the contribution ratio.

If the weight is W and the growth value is h, then the weight curve is expressed by Eq. (3).

\[
W = Ah^2 + Bh + C
\]  

(3)

A, B, and C in Eq. (3) denote the growth values obtained from the field and the coefficients obtained by approximating the weight per stock by a quadratic function. The values of A, B, and C, obtained in this observation, were \(4.7 \times 10^{-3}\), \(1.5 \times 10^{-2}\), and \(-5.9 \times 10^{-2}\), respectively. Further, the number of stocks per square meter was measured as 2741 stocks/m² to predict the yield value from the weight per stock.

![Fig. 7 Relationship between actual measured growth value and measured single yield value](image)

Table 6 Correlation coefficient between approximate function and measured value and contribution ratio in the relationship between growth value and yield value

|                     |          |          |
|---------------------|----------|----------|
| Correlation coefficient | 0.93     |          |
| Contribution ratio   | 0.88     |          |

4.4 Relationship between effective accumulated soil temperature and growth curve

From Table 2, there are two items with a high correlation coefficient of the relationship between the field environment and growth value: the accumulated air temperature and the effective accumulated soil temperature. The correlation coefficient between the measured growth value and the effective accumulated soil temperature in three seasons is greater than the value of the correlation coefficient between the measured growth value and the accumulated air temperature. Hence, the effective accumulated soil temperature is considered more suitable for
expressing the growth values of yearly crops. The growth values in three seasons and the transition of the effective accumulated soil temperature are shown in Fig. 8.

Figure 9 shows the transition of the effective accumulated soil temperature in each season. Figure 10 illustrates the relationship between the number of elapsed days after sowing and the approximate growth curve for each season. Figures 9 and 10 demonstrate that the effective accumulated soil temperature and the growth are different in each season. Figure 11 illustrates the relationship between the effective accumulated soil temperature from the date of sowing and the approximate growth curve in each season.

Figure 11 also shows that the approximate growth curve of each season may be treated as a single seasonal growth curve when the effective accumulated soil temperature is the independent variable. The general formula of the seasonal approximation growth curve is shown in Eq. (4). Further, the formula for calculating the effective accumulated soil temperature in the seasonal growth curve is given by Eq. (5).

\[ h_{st} = \frac{K}{1 + b \cdot e^{c \cdot T_{accum}}} \]  
\[ T_{accum} = \sum_{t=1}^{n} T_{st} \]

*\( h_{st} \): Effective accumulated soil temperature [°C] in season s, t days after sowing

*\( T_{accum} \): Effective accumulated soil temperature [°C] in season s, t days after sowing

*\( T_{st} \): Soil temperature [°C] of each day in season s, t days after sowing

Table 7 lists the parameters of the approximate growth curve for each season and the approximate seasonal growth curve of the logistic function. Table 8 shows the error rate of the mean growth value, the standard deviation of the seasonal approximation growth curve, and the measured growth curve. Table 9 presents the correlation coefficients between the approximate growth curve and the seasonal growth curve, as well as the contribution ratios. The contribution ratio of the seasonal approximation growth curve and the approximation curve of each season is 0.99 in the rainy season of 2016, 0.99 in the summer of 2017, and 0.99 in the autumn of 2016. This indicates that the seasonal changes in the growth curve of the year-round crop can be predicted using the soil temperature data.

| Season       | 2016 Rainy season | 2017 Summer | 2017 Autumn | Seasonal approximate growth curves |
|--------------|-------------------|-------------|-------------|----------------------------------|
| K            | 10.94             | 11.06       | 10.56       | 10.65                            |
| b            | 340.12            | 160.79      | 1414.17     | 494.22                           |
| c            | 1.58              | 1.39        | 2.00        | 1.72                             |

| Year         | Error rate of mean growth value | Error rate of standard deviation |
|--------------|---------------------------------|----------------------------------|
| 2016 Rainy season | 1.9%                            | 6.4%                             |
| 2017 Summer     | 2.8%                            | 6.5%                             |
| 2017 Autumn     | 3.2%                            | 0.9%                             |

| Year         | Correlation coefficient | Contribution ratio |
|--------------|-------------------------|---------------------|
| 2016 Rainy season | 0.99                    | 0.99                |
| 2017 Summer     | 0.99                    | 0.99                |
| 2017 Autumn     | 0.99                    | 0.99                |
Fig. 8 Effective accumulated soil temperature and predicted growth value based on the number of elapsed days after sowing

Fig. 9 Transition of accumulated soil temperature in each season

Fig. 10 Relationship between the number of elapsed days after sowing and approximate growth curves in each season
5. Simulation to evaluate storage shortages and quality loss

5.1 Overview

Using the state transition models of the proposed production and growth processes and sensor data, we propose a simulation method to evaluate storage shortages and quality loss. By using this simulation, if there are shortages in the field or a desire to increase the number of sowing fields, it will be possible to predict the sufficient number of fields and the harvest yield. An overview of the simulation method is shown in Fig. 12. The simulation consists of two-state transition models, pertaining to the field state and the crop state, and a growth prediction algorithm using the growth observation results after sowing. Figure 13 summarizes the two-state transition models. There are multiple fields in the agricultural production system and each field has a state transition model. Any given field undergoes a transition through the five states listed in Section 3.1. Depending on the growth value, the crop state transitions through three states: growth, suitable harvest, and unsuitable harvest. If the growth value does not reach the stipulated standard, the crop cannot be harvested.

The field state transitions to the reaping according to the supply instruction. If the growth value meets the criteria, the crop can be harvested, after which it is shipped. If the growth value exceeds the standard, it is discarded as quality loss.

After harvesting, the field state transitions to the fallow state. It then transitions to the plough following the instruction for curing and to the next production process for sowing. The growth prediction algorithm will be discussed in detail in Section 5.2.

Sowing is carried out on vacant fields that are in wait for sowing, in accordance with a previously prepared sowing plan that formulates the number of sowing fields on a daily basis. Demand is regarded as input information. For the shipping standard growth value, the zone is entered as the model input.

Thus, factors such as storage shortage, storage shortage ratio, amount of quality loss, quality loss ratio, field operation rate, sowing frequency, and harvest amount, which are necessary to evaluate the agricultural production systems, are considered as model outputs. The output data obtained from the simulation are examined, and the input values are reset for the next simulation.
5.2 Growth prediction algorithm based on effective accumulated soil temperature

We propose a growth prediction algorithm, based on the estimated soil temperature, that will predict the crop growth and yield values necessary for the evaluation of storage shortages and quality loss from growth and yield values, and from the results of the analysis. Storage shortages occur when the growth value does not meet the shipping standard at the time of demand, and the crop cannot be reaped. Quality loss occurs when the growth value of the crop exceeds the shipping standard before harvesting is completed. Therefore, an appropriate harvest state is predicted by estimating the growth value in the simulation, and storage shortages and quality loss due to fluctuations in harvest date are evaluated. The amount of harvest varies depending on the growth value of the crop at the time of harvesting. However, the relationship between the growth value and the yield value follows a quadratic curve. Therefore, the amount of harvest is predicted by estimating the yield value in the simulation, and storage shortages and quality loss due to the difference between demand and supply are evaluated.

The growth prediction algorithm based on the accumulated soil temperature is composed of three algorithms: predicted growth curve generation algorithm, predicted suitable harvest state generation algorithm, and predicted yield value generation algorithm. The predicted growth curve generation algorithm uses the growth values obtained from the field to generate a predicted growth curve that considers the variations in growth. The predicted suitable harvest state generation algorithm uses the generated predicted growth curve and the field-defined crop shipping standard to generate a predicted reaping state. The yield value prediction algorithm uses the generated predicted growth curve and the weight curve obtained from the field to generate a predicted yield value at the harvest time. Using these algorithms, the reaping state and the yield value required for the prediction of storage shortages and quality loss are estimated.

5.2.1 Predicted growth curve generation algorithm
We propose an algorithm to generate a predicted growth curve using the growth curve obtained from the growth observation and the variation in the growth value. The coefficients $K_r$, $b_f$, $c_r$ of the predicted growth curve are generated from the coefficients $K$, $b$, $c$ in the growth curve obtained from the field observations and the measured midpoints ($t_{mf}$, $h_{mf}$). Further, to generate $c_r$, a prediction midpoint ($t_{mf}$, $h_{mr}$) is also generated. The variation is assigned on the growth curve when generating the prediction midpoint, and thus the growth variation factor is included in the generation of the predicted growth curve. The predicted midpoint is derived by generating a normal random number from the mean and standard deviation of the measured midpoints.

**STEP 1.1** Derive the accumulated soil temperature $t_{mf}$ at the measurement midpoint from the seasonal approximation growth curve obtained by the growth process investigation. $t_{mf}$ takes the value of the accumulated soil temperature when the growth value is $K/2$ [cm].

**STEP 1.2** Generate the growth value $h_{mr}$ of the predicted midpoint at the accumulated soil temperature $t_{mf}$. $h_{mr}$ takes the value of the normal random number generated based on the mean and standard deviation of $h_{mf}$ obtained from the seasonal approximation growth curve.

**STEP 1.3** Generate the coefficient $K_r$ of the predicted growth curve. $K_r$ takes the value of the normal random number generated based on the mean and standard deviation of $K$ obtained from the seasonal approximation growth curve.

**STEP 1.4** Generate the coefficient $b_f$ of the predicted growth curve. $b_f$ takes the value of $b$ obtained from the seasonal growth curve.

**STEP 1.5** To determine the coefficient $c_r$ of the predicted growth curve, substitute $t_{mf}$, $h_{mr}$, $K_r$, $b_f$ in Eq. (2) and solve for $c$. Then, $c_r$ is expressed by Eq. (6).

$$c_r = \frac{1}{t_{mf}} \left[ \log b_f - \log \left( \frac{K_r}{h_{mr}} - 1 \right) \right]$$  \hspace{1cm} (6)

**STEP 1.6** Generate Eq. (4), which is the predicted growth curve based on the accumulated soil temperature by substituting the values obtained for $K_r$, $b_f$, $c_r$ in Eq. (1).

$$h = \frac{K_r}{s + b_f e^{-c_r t}}$$ \hspace{1cm} (7)

### 5.2.2 Predicted suitable harvest state generation algorithm

We propose an algorithm to predict the suitable harvest state $T_{ship}$ from the generated predicted growth curve. This algorithm that predicts the change in the reaping state, due to the growth value of crops, is defined as follows.

**STEP 2.1** Enter the lower specification limit $h_{lf}$ and the upper specification limit $h_{hf}$ that meet the crop shipping standard. The standard value is the value used in the actual field.

**STEP 2.2** By substituting $h_{lf}$ and $h_{hf}$ into the predicted growth curve Eq. (8), determine the soil temperature $t_{tr}$ at the start of the harvesting season and the corresponding time $T_{tr}$, and the soil temperature at the end of the harvesting season $t_{hr}$ and the corresponding time $T_{hr}$. $t_{tr}$ is expressed by Eq. (8) and $t_{hr}$ by Eq. (9).

$$t_{tr} = \frac{1}{c_r} \left[ \log b_f - \log \left( \frac{K_r}{h_{lf}} - 1 \right) \right]$$  \hspace{1cm} (8)

$$t_{hr} = \frac{1}{c_r} \left[ \log b_f - \log \left( \frac{K_r}{h_{hf}} - 1 \right) \right]$$  \hspace{1cm} (9)

**STEP 2.3** The reaping state $T_{ship}$ is calculated using $T_{tr}$ and $T_{hr}$ with Eq. (10).
\[ T_{ship} = T_{hr} - T_{tr} \]  

5.2.3 Predicted yield generation algorithm

The growth value and the weight per stock are related by a quadratic curve, and the yield value changes depending on the value of the weight per stock. Therefore, the algorithm to derive the predicted yield value \( U_{cut} \) from the generated predicted growth curve and the weight curve obtained from the field observations is presented below.

**STEP 3.1** Enter the harvest time \( T_{cut} \), which changes with the harvesting state \( T_{ship} \) and the reaping plan of the field.

**STEP 3.2** Determine the growth value \( h_{cut} \) at the harvest time by substituting \( T_{cut} \) into Eq. (4).

**STEP 3.3** Determine the predicted yield value \( U_{cut} \) using the weight curve obtained by approximating the field-obtained value by a quadratic curve and \( h_{cut} \). The predicted weight value \( W_{cut} \) is expressed by Eq. (11).

\[ W_{cut} = A h_{cut}^2 + B h_{cut} + C \]  

The predicted yield value \( U_{cut} \) is obtained by multiplying the predicted weight value \( W_{cut} \) by the number of stocks per unit area from the field observations.

6. Case study

A case study was carried out to verify the effectiveness of the proposed simulation. The objectives of the case study are as follows.

Objective 1: To verify whether the predicted growth curve generated from the proposed algorithm is representative of the transition and variation in each season in the field and determine its accuracy.

Objective 2: Verify that the proposed simulation is capable of evaluating, in advance, the storage shortages and the quality loss.

The data obtained from HATAKE Company, Ltd. were used as the input for the simulation. In this case study, the WITNESS Monte Carlo simulator and Visual Basic Extension were used [19].

The input conditions used for the simulation are listed in Table 10. Input conditions are divided into field conditions and crop conditions. Table 11 presents the information on soil temperature observations used as inputs in the case study. Figure 14 shows the soil temperature data for Imakashima, Tsukuba, Ibaraki Prefecture from June 20, 2015 to October 23, 2015. The data obtained by the growth process observation were used as the input crop conditions.

6.1 Verification of growth prediction algorithm based on effective accumulated soil temperature

We verified the prediction accuracy of the growth and yield values generated by the proposed growth algorithm. The predicted growth curve was generated by simulating 1000 turns of sowing. Accuracy was verified by comparing the growth value obtained by approximating the measured values from the growth process observation to the logistic function. The yield value was also compared with the predicted growth and yield values from the simulation. Crop conditions were used as the input conditions (Table 10).

6.1.1 Verification of prediction accuracy of growth value

The prediction accuracy of the growth value was verified by comparing the seasonal approximation growth curve from the growth process observation with the output of the growth curve prediction algorithm. The values generated by the growth prediction algorithm are shown in Fig. 15. The comparison of the predicted growth value obtained from the simulation with the approximated measured growth value showed that the error rate of the
predicted growth value had a mean of 2.6% and standard deviation of 3.2%.

**6.1.2 Verification of prediction accuracy of yield value**

The prediction accuracy of the yield value was verified by comparing the approximated actual measurement yield value from the growth process observation with the output of the yield value prediction algorithm. Figure 16 shows the relationship between the accumulated soil temperature and the predicted yield value generated by the growth prediction algorithm. The relationship between the growth value and the predicted yield value is shown in Fig. 17. The comparison of the approximated measured growth value with the predicted yield value obtained from the simulation showed that the error rate of the predicted yield value had a mean of 5.4 [%] in the rainy season of 2016, 3.9 [%] in the summer of 2017, 5.8 [%] in the autumn of 2017, and standard deviation of 8.4 [%] in the rainy season of 2016, 0.4 [%] in the summer of 2017 and 7.7 [%] in the autumn of 2017.

| Table 10 Input conditions of the simulation |
|--------------------------------------------|
| **Field state information**                |
| Number of days of curing [d]               |
| Cultivated area [m²]                       |
| Number of fields [Building]                |
| Number of sowings per day [times]          |
| Harvesting state [h]                       |
| Shipping standard upper limit [cm]         |
| Shipping standard lower limit [cm]         |
| Scheduled harvesting date [d]              |
| Demand per day [kg/d]                      |
| Soil temperature information [°C]          |
| **Crop state information**                |
| Growth value of the measurement midpoint (mean) [cm] |
| Growth value of the measurement midpoint (standard deviation) |
| Elapsed days of actual measurement midpoint [d] |
| Mean of convergence value K of growth curve [cm] |
| Standard deviation of Convergence value K of growth curve |
| Initial value b of growth curve            |
| Slope c of growth curve                    |
| Coefficient A of single yield curve        |
| Coefficient B of single yield curve        |
| Coefficient C of single yield curve        |
Table 11 Information on soil temperature observations

| Case Study | Period                          | Data acquisition location       |
|------------|---------------------------------|---------------------------------|
| 6.1 Accuracy verification of growth prediction algorithm based on the accumulated soil temperature | [Rainy season] 20/June/2016-18/July/2017 [Summer] 1/August/2017-31/August/2017 [Autumn] 23/September/2017-23/October/2017 | Yoshinuma, Tsukuba, Ibaraki Prefecture |
| 6.2 Evaluation of storage shortages and quality loss by simulation | 20/June/2015–23/October/2015 | Imakashima, Tsukuba, Ibaraki Prefecture |

6.2 Evaluation of storage shortage and quality loss by simulation

We verified the possibility of evaluating storage shortages and quality loss using the proposed simulation method. Simulations were carried out on 150 fields. Table 12, 13 lists the input conditions. The amount of harvest, storage shortages, and quality loss, when the crop growth varies with the season, are shown in Fig. 18, and their implications are shown in Fig. 19. In the seasonal variations in crop growth, if the soil temperature decreases, the growth state becomes longer, resulting in increased storage shortages and reduced quality loss. As the soil temperature rises, the growth state is shortened, storage shortages decrease, and the quality loss increases, which is agreement with the field observations. Thus, the proposed state transition model and the growth prediction algorithm can be used to determine the storage shortages and quality loss in the field by considering the growth variations.
Table 12 Input conditions (Verification of growth prediction algorithm accuracy based on accumulated soil temperature)

| Parameter            | Value   |
|----------------------|---------|
| K mean               | 10.65   |
| K standard deviation | 1.19    |
| b                    | 494.22  |
| c                    | 1.72    |
| A                    | 4.7×10⁻³ |
| B                    | 1.5×10⁻² |
| C                    | -5.9×10⁻² |

Fig. 15 Predicted growth values generated by the growth curve generation algorithm

Fig. 16 Predicted yield value vs. accumulated soil temperature plot generated by the growth prediction algorithm
### Table 13 Input conditions

| Field state information | Crop state information |
|-------------------------|------------------------|
| Number of curing days [d] | Growth value of the measurement midpoint (mean) [cm] | 5.32 |
| Cultivated area [m²] | Growth value of the measurement midpoint (standard deviation) | 0.86 |
| Number of fields [Building] | Accumulated soil temperature of measured midpoint | 361.43 |
| Number of sowings per day [times] | Mean of K | 10.65 |
| Harvesting state [h] | Standard deviation of K | 1.19 |
| Shipping standard upper limit [cm] | B | 494.22 |
| Shipping standard lower limit [cm] | C | 1.72 |
| Demand per day [kg/d] | A | 4.7×10⁻³ |
| | B | 1.5×10⁻² |
| | C | -5.9×10⁻² |

![Fig. 17 Predicted yield value generated by the growth prediction algorithm vs. growth value](image-url)
7. Conclusion

In this study, we evaluated the shortages and quality loss that occur in the growth and reaping stages of the agricultural production system while considering the seasonal changes in crop growth. The field data were analyzed using sensors. We then proposed a model for the working process of the field and a growth prediction algorithm and implemented these changes in the simulation. We verified the accuracy of the proposed algorithm and determined the operation of the state transition model through a case study. We derived the changes in storage shortages and quality loss in the rainy season, summer, and autumn. Future prospects of this study include the evaluation of storage shortages and quality loss for the entire year including the winter and spring seasons, considering the seasonal variations in the yield value.

Acknowledgements

We would like to thank Mr. Ryohei Hamasato of HATAKE Company for his multifaceted cooperation in the field surveys. We would like to thank Mr. Hiroshi Nogawa and Mr. Shuichiro Suetsugu of Fujitsu FSAS Inc. for their specific and continuous supports for IoT-related technologies. We would like to thank Mr. Toshihiro Matsui and Mr. Shogo Shimizu who was graduated from Hibino Laboratory at Tokyo University of Science, for their participation in the field surveys and analysis.
References

D’Agostino, R.B., Belanger,A. and D’Agostino, Jr.R.B., A Suggestion for Using Powerful and Informative Tests of Normality, The American Statistician, Vol.44, No.4(1990), pp.316-321.

Dai, H., An analysis of Consumption at an Establishment for Direct Sales, Scientific reports of the Faculty of Agriculture, Okayama University, Vol.93, No.1(2004), pp.77–81(in Japanese).

Fujitsu FSAS, FUJITSU Intelligent Society Solution, Mina no WA!, available from <http://www.fujitsu.com/jp/group/fsas/solutions/business-technology/intelligent-society/chienowa/>,(accessed on May 25, 2016)(in Japanese).

HATAKE company , HATAKE COMPANY, available from <http://hatake.co.jp/>, (accessed on October 10, 2016) (in Japanese).

Hibino, H., Matsumoto. Y., Kubo. N., Kimura. M., Mizukami. Y., Modelling and Simulation of Agricultural Production System Based on IoT Cultivated Fields Information, 2017 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM 2017)

Hoshi, T., Ohata, R., Watanabe, K., Ueda, S. and Nanseki, T., Measures of URCS application for checking the environment in the greenhouse aiming to establish a one-stop management system by mobile electronic equipment, Agricultural Information Research, Vol.22, No.1, (2013) pp.193-200(in Japanese).

ITOCHU Techno-Solutions Corporation, WITNESS, engineering-eye, available from <http://www.engineering-eye.com/WITNESS/>, (accessed on June 20, 2016)(in Japanese).

Kumamoto Prefecture Vegetable Promotion Association, Japanese mustard spinach, <http://www.k-engei.net/contents/koushu_standard/131 コマツナ.pdf>, (accessed on August 16, 2019)(in Japanese).

Matsuda, S., Matsubara, K. and Ozawa, K., Proposal of model for Ginkgo's yellow leaves using the logistic function, Hosei University Tama Research Report, Vol. 26 (2011), pp. 21–25(in Japanese).

Matsukura, S., Nanseki, T., Fujii, Y., Sato, M., Chomei, Y. and Miyazumi, M., Improvement of technology and skill in large-scale rice cultivation management and cost reduction effect of scale expansion, Agricultural Information Research, Vol.24, No.2(2015), pp.35-45(in Japanese).

Zhang, P., Economic impacts of climate change on agriculture: The importance of additional climatic variable other than temperature and precipitation, Journal of Environmental Economics and Management, Vol.83 No.2(2017), pp.8-31.

Sanseido ed., Daijirin third edition (2006), p.1576(in Japanese).

Sato, K. and Takahaschi,T., Estimation of Soil Nitrogen Absorption by Paddy Rice Using a Logistics Curve, Journal of the science of soil and manure, Vol. 73, No.3(2002), pp. 293-295(in Japanese).

Takatsuji, M., Present status of completely-controlled plant factory, Journal of Science and High Technology in Agriculture , Vol.22, No.1(2010), pp.2-7(in Japanese).

Takayama, K., Second generation speaking plant approach, Journal of Science and High Technology in Agriculture, Vol.25, No.4 (2013), pp.165-174(in Japanese).

Togami, T., Itoh, Y., Hashimoto, A. and Kameoka, K., Measurement of field environment using sensor network aimed at high quality mandarin orange production, Agricultural Information Research, Vol.20, No.3(2011), pp.110-121(in Japanese).

Tozaki, T., Seasonal stem growth and growth curve of water shield, Journal of Agricultural Engineering Society, Vol. 52, No.2(2011), pp. 121-127(in Japanese).

Zhang, Y., Wang, L. and Duan, Y., Agricultural information dissemination using ICTs: A review and analysis of information dissemination models in China, Information processing in agriculture, Vol.3, No.1(2016), pp.17-29.