THE INTERNATIONAL SOCIETY OF PRECISION AGRICULTURE PRESENTS THE
13th INTERNATIONAL CONFERENCE ON
PRECISION AGRICULTURE
July 31-August 4, 2016 • St. Louis, Missouri USA

Analysis of high-yield conditions using a rice yield predictive model

Y. Hirai, T. Yamakawa, E. Inoue, T. Okayasu, M. Mitsuoka

Faculty of Agriculture, Kyushu University, 6-10-1 Hakozaki, Higashi-ku, Fukuoka 812-8581, Japan

A paper from the Proceedings of the
13th International Conference on Precision Agriculture
July 31 – August 4, 2016
St. Louis, Missouri, USA

Abstract. *Rice production in Japan is facing problems of yield and quality instability owing to recent climate changes and a decline in rice prices, and possible competition with foreign inexpensive rice. Thus, it is becoming more important to stably achieve high yield and quality, while reducing production costs. Various data, including crop growth, farmer's management styles, yield and quality, has recently become accessible in actual fields using advanced information and communication technologies. Those data can be effectively used to aid farmer's decision-making on their management. In this study, we built predictive models of brown rice yield (yield) using 85 data sets collected from 2010 to 2015 at 21 paddy fields in Itoshima city, Fukuoka prefecture, Japan. In the paddy fields, rice was cultivated under various environmental conditions and management styles. Support vector machine was applied to build the models to predict three yield classes (low, middle, high) that ranged from 2.98 to 6.17 t ha*$^{-1}$*. The models were built using all the combinations of four explanatory variables: number of spikelets, inorganic nitrogen (N) supply from panicle initiation to mid-ripening dates, and average values of sunshine hours and air temperature from heading to mid-ripening dates. A yield predictive model had the highest classification accuracy of 79% when the model selected two variables: the number of spikelets and inorganic N supply. The model identified high-yield conditions ranging from 27,216 to 30,146 m*$^{-2}$* for number of spikelets and from 58.1 to 72.1 kg ha*$^{-1}$* for inorganic N supply. The results indicated that shortage of inorganic N supply was one of the major causes to lower yield. The results also suggested that applying second topdressing was necessary to achieve high yield in the target region.*

Keywords. *fertilizer application, nitrogen supply, ripening, support vector machine, weather, yield capacity.*

Introduction

Rice production in Japan is facing problems of yield and quality instability owing to recent climate changes and a decline in rice prices, and possible competition with foreign inexpensive rice. Thus, it is required to stably achieve high yield and quality, while reducing production costs. However, current farmer's management mainly relies on their experience and knowledge, making yield and quality unstable due to unavoidable trial and error. Furthermore, current farmer's management is not always optimized under different local climates and soil conditions. One of the reasons that experience-based management has been persistent in actual fields was mainly because of complex relationships among crop growth, weather condition, and nutrient and water supply. This makes difficult to provide quantitative guidelines for farmer's management such as fertilizer application.

Crop growth models are available as a possible method to provide quantitative guidelines for farmer's management (Yoshida and Horie, 2010). Crop growth models allow to predict straw weight and rough rice yield by mathematically expressing processes of crop nitrogen (hereafter, N) uptake and photosynthesis. The models can relate crop growth and yield with farmer's management such as fertilizer application. However, the models require various data for their model calibration that include N content in leaf and leaf area index measured under different conditions of fertilizer application and weather. This model calibration is a big obstacle to apply the models in actual fields. Furthermore, crop growth models, which are normally calibrated under limited experimental conditions, are likely to deteriorate their predictive accuracy in the application to actual fields involving diverse environmental conditions and farmer's management styles.

Various data, including crop growth, farmer's management, yield and quality, has recently become accessible in actual fields of rice production because of the advance in information and communication technologies. Although those data collected in actual fields are limited in terms of their details, they reflect different rice production conducted under diverse environmental conditions and management styles. Data collected in actual fields can be effectively used to aid farmer's decision-making on their management if we can extract useful information and find quantitative relationships between yield or quality and influential factors from the data.

The objective of this study was to identify conditions where high yield was achieved (hereafter, high-yield conditions), focusing on factors including yield capacity, inorganic N supply during the reproductive growth stage, and weather conditions during the ripening period. We applied a pattern recognition method, support vector machine (SVM), to build predictive models for brown rice yield (hereafter, yield). The models were built using 85 data sets collected from 2010 to 2015 at 21 paddy fields in Itoshima city, Fukuoka prefecture. The best combination of explanatory variables were determined based on classification accuracy. High-yield conditions were finally identified by using the predictive model with the highest classification accuracy.

Materials and methods

Yield predictive models were built by using 85 data sets collected from 2010 to 2015 at 21 paddy fields in the city of Itoshima (33°29′N – 33°37′N, 130°08′E – 130°15′E), Fukuoka Prefecture, Japan. The Hinohikari cultivar was grown in all surveyed fields. The purpose of building predictive models was to identify high-yield conditions for yield capacity, inorganic N supply during the reproductive growth stage, and weather during the ripening period. Thus, the following four explanatory variables were selected as candidates for building the models: number of spikelets, inorganic N supply from panicle initiation to mid-ripening dates (hereafter, inorganic N supply), and the average values of sunshine hours and air temperature from heading to mid-ripening dates (hereafter, ripening period). Here, panicle initiation and mid-ripening dates were defined as dates 30 days before heading and after heading, respectively. Table 1 presents four candidate explanatory variables and the range of their values.

Sources of inorganic N supply included soil, manure, basal fertilizer and topdressing. Inorganic N supply through the mineralization of organic N in soil and manure was estimated by the kinetic method (Sugihara et al., 1986). Release of inorganic N from basal fertilizer and topdressing was estimated by applying different formulas for each of three fractions of N in fertilizer: quick release N, slow release N, and organic N, which have different inorganic N release patterns. Inorganic N supply from a quick release fraction was estimated by a predictive equation of fertilizer N release (Shibahara et al., 2000). Inorganic N supply from a slow release fraction was estimated by an improved simulation model for nutrient release with Gaussian correction (Kobayashi et al., 2000). Inorganic N supply from an organic fraction was estimated by the kinetic method (Sugihara et al., 1986).

Table 1. Four candidate explanatory variables and the range of their values

Candidate explanatory variables	Range of values
Number of spikelets (m^{-2})	18,428–35,419 (25,828±2,976)
Inorganic N supply from panicle initiation to mid-ripening dates (kg ha^{-1})	19.1–99.9 (49.2±12.6)
Average temperature from heading to mid-ripening dates (°C)	21.5–27.3 (24.0±1.2)
Average sunshine hours from heading to mid-ripening dates (h)	3.3–7.1 (5.4±1.0)

Numbers in parentheses represent mean±standard deviation.

SVM was applied to build models to predict three classes of yield. The yield was discretized into three classes, low, middle, and high, which corresponded to data included in the lowest, middle, and highest interval, respectively, when the range of yield was split into three equal intervals. Table 2 shows the class interval and the number of data sets included in each yield class. Predictive models were built using all possible combinations of candidate explanatory variables presented in table 1. The predictive models were evaluated by classification accuracy in leave-one-out cross validation. Classification accuracy was calculated by the ratio of correct classification to total data sets. The predictive model with the highest classification accuracy was used to identify high-yield conditions.

Table 2. Class interval and the number of data sets included in each yield class

Yield class	Interval (ton ha^{-1})	Number of data sets
low	2.98 – 4.04	22
middle	4.04 – 5.11	55
high	5.11 – 6.17	8

Results

Table 3 shows classification accuracy of yield predictive models. The yield predictive model, including two explanatory variables that are number of spikelets and inorganic N supply, had the highest classification accuracy, 79%. Predictive models with only one variable also had relatively high classification accuracy ranging from 65 – 73%.

Figure 1 shows predicted conditions of number of spikelets and inorganic N supply corresponding to three yield classes and observations. High-yield conditions were mainly identified in the range from 27,216 to 30,146 m^{-2} for number of spikelets and from 58.1 to 72.1 kg ha^{-1} for inorganic N supply. As the number of spikelets and inorganic N supply decreases, yield were down to middle and low classes, depending on how much the explanatory variables were different from high-yield conditions.

Discussion

The number of spikelets was the most influential to classification accuracy among the four candidate variables. The predictive model, which only included number of spikelets as an explanatory variable, had classification accuracy of 73%. This result was because data sets categorized in middle yield accounted for 65% of total data sets. Thus, classification accuracy became high even when only the smallest and the largest numbers of spikelets in the middle class were used as separating lines to classify three yield classes. However, including inorganic N supply allowed to identify high-yield conditions that applied to 4 out of 8 high-yield cases. This improved classification accuracy from 73 to 79%.

Table 3. Classification accuracy of yield predictive models with different combinations of explanatory variables

Variables	Number of variables	Number of spikelets	Inorganic N supply	Average temperature	Average sunshine hours	Correction rate (%)
Selection	4	✓	✓	✓	✓	77
	3	✓	✓	✓		73
	3	✓	✓		✓	77
	3	✓		✓	✓	72
	3		✓	✓	✓	70
	2	✓	✓			79
	2	✓		✓		74
	2	✓			✓	72
	2		✓	✓		69
	2		✓		✓	73
	2			✓	✓	66
	1	✓				73
	1		✓			67
	1			✓		65
	1				✓	65

✓ represents that a corresponding variable was selected as an explanatory variable of a predictive model.

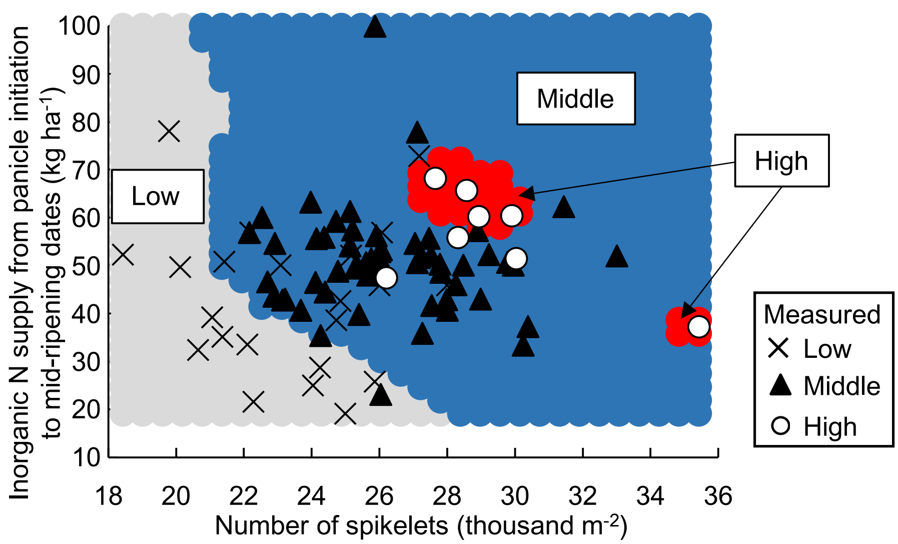

Fig 1. Predicted conditions of number of spikelets and inorganic N supply corresponding to three yield classes and observations. Areas filled with grey, blue, and red colors represent the range of values for number of spikelets and inorganic N supply that correspond to a low, middle, and high yield class, respectively.

Furthermore, many cases remained in the middle yield class although the number of spikelets was in the range corresponding to high-yield class. This result reflected that shortage of inorganic N supply lowered yield. One of the reasons for the shortage of inorganic N supply was due to low soil fertility found in several paddy fields. Topdressing is normally applied two times in the target region. However, some farmers didn't apply second topdressing because they were concerned over negative effects on palatability and disease occurrence. High yield seemed to be achieved when inorganic N was more supplied from soil under high temperature throughout the growth period and from increased amount of second topdressing under longer sunshine hours during the ripening period. This result indicates that applying second topdressing is necessary to achieve high yield in the target region. However, applying second topdressing needs to be determined by considering possible negative effects on palatability and disease occurrence.

For the causes of small number of spikelets that lowered yield, we need to further investigate its relationships with inorganic N supply and weather conditions (temperature, sunshine hours, precipitation) during the vegetative growth stage.

Conclusions

The objective of this study was to identify high-yield conditions, focusing on factors including yield capacity, inorganic N supply during the reproductive growth stage, and weather conditions during the ripening period. The yield predictive model built using data sets collected in actual fields had the highest classification accuracy of 79% when the model selected two variables: the number of spikelets and inorganic N supply from panicle initiation to mid-ripening dates. The model identified high-yield conditions ranging from 27,216 to 30,146 m^{-2} for number of spikelets and from 58.1 to 72.1 kg ha^{-1} for inorganic N supply. Many cases remained in the middle yield class although the number of spikelets was in the range corresponding to the high-yield class. This result reflected that shortage of inorganic N supply lowered yield. Reasons for the shortage of inorganic N supply were due to low soil fertility and the omission of second topdressing. This suggests that applying second topdressing is necessary to achieve high yield in the target region, although negative effects on palatability and disease occurrence need to be considered.

Acknowledgements

This work was supported by JSPS KAKENHI Grant Number 15K07678.

References

Kobayashi, A., Fujisawa, E., Kubo, S., Hanyuu, T. (1997). A mechanism of nutrient release from resin-coated fertilizers and its estimation by kinetic methods 3. An improved simulation model for nutrient release from coated urea with Gaussian correction. *Jpn. J. Soil Sci. Plant Nutr.*, 68, 487–492.

Shibahara, F., Takehisa, K., Hasegawa, K. (2000). Development of a simple model for nitrogen uptake and fertilizer application of rice crop. *Jpn. J. Soil Sci. Plant Nutr.*, 71, 898–902.

Sugihara, S., Konno, T., Ishii, K. (1986). Kinetics of mineralization of organic nitrogen in soil. Bulletin of National Institute for Agro-Environmental Sciences, 1, 127–166.

Yoshida, H., Horie, T. (2010). A model for simulating plant N accumulation, growth and yield of diverse rice genotypes grown under different soil and climatic conditions. Field Crops Research, 117, 122-130.