JARQ 60 (2), 147-156 (2026) https://www.jircas.go.jp

Inter-Field Wheat Yield Variation Evaluated Using Time Series Satellite Imagery and Functional Data Analysis

Shuhei YAMAMOTO[1]*, Katashi KUBO[1], Shigeto FUJIMURA[1] and Hirotake MIYAJI[2]

[1] Tohoku Agricultural Research Center, National Agriculture and Food Research Organization, Fukushima, Japan
[2] Tohoku Agricultural Research Center, National Agriculture and Food Research Organization, Morioka, Japan

Abstract

The growing number of large agricultural corporations in Japan requires an efficient means to evaluate crop growth and yield formation for multiple fields. Using a combine harvester with a yield monitor sensor, we collected wheat yield data for 76 fields cultivated by an agricultural corporation. This study used functional regression and functional data analysis (FDA) to analyze the relation between inter-field wheat yields and growth using the time-series Normalized Difference Vegetation Index (NDVI) calculated from Sentinel-2 satellite imagery. The observed inter-field yield variation was 17–361 g m^{-2}. FDA applied functional regression with two explanatory variables: the estimated functional NDVI and the sowing date. The inter-field yield variation was explained by the coefficient of determination (R^2) = 0.39. The coefficient function indicated that functional NDVI affected yield. This study assessed the possibility of using satellite information for agricultural management and provided an alternative to analyzing time series data to infer crop growth and yield variation.

Discipline: Crop Science
Additional key words: agricultural fields, functional regression, Sentinel-2, yield monitor sensor

Introduction

Because of the insufficient farming population and increasing concentration of fields among managing organizations in Japan, agricultural operations cultivating multiple fields have become increasingly numerous in many regions (Umemoto 2019). One associated difficulty is inter-field yield variation, which decreases profits. Ishikawa et al. (2021) conducted an analysis on 637 fields in Japan and observed significant inter-field yield variation. They presented management procedures to improve rice productivity. The study findings implied that coping with inter-field yield variation is important for planning exit cultivation strategies and improving profits. Recognizing inter-field yield variation and evaluating the relation between crop growth and yield are important to manage widely diverse and dispersed fields and to improve yields for agricultural corporations. Some studies have analyzed inter-field yield variation, but information about the actual conditions of individual companies and farmers is scarce. This information might promote implementation of agricultural technologies.

Satellite-based remote sensing using vegetation indices (VIs) obtained from optical satellite imagery is suitable for collecting information over a wide area. VIs are composite intermediaries related to biophysical plant parameters such as biomass, leaf area, and leaf chlorophyll contents, which are applied mainly for crop classification, phenology estimation, and yield prediction (Gao et al. 2000, Inoue 2020, Zhang et al. 2020). Among the many VIs, the Normalized Difference Vegetation Index (NDVI) is the most used. Its ratio form reduces directional reflectance fluctuations (Tucker 1979, Huete et al. 2002, Jones & Vaughan 2010). Other VIs have been proposed to alleviate associated difficulties arising from the use of observed phenomena. The Enhanced Vegetation Index

*Corresponding author: yamamoto.shuhei418@naro.go.jp
Received 24 December 2024; accepted 15 July 2025; J-STAGE Advanced Epub 28 January 2026.
https://doi.org/10.6090/jarq.24J24

(EVI), which mitigates the influence of atmosphere and canopy background during monitoring, was developed originally for Moderate Resolution Imaging Spectroradiometry (MODIS) (Huete et al. 2002). Two-band EVI (EVI2), proposed as an accurate substitute for EVI, consists of two bands: red and near infrared (NIR) (Jiang et al. 2008). Recently, a free platform for satellite imagery has appeared with satellite VIs, facilitating frequent monitoring of vegetation at both national and regional levels (da Silva et al. 2020, Gong et al. 2020, Sonia et al. 2022, Rodigheri et al. 2023). Satellite-based remote sensing might be a useful tool for farmers to store continuous field data automatically and to use images from several satellites, which are freely available. Nevertheless, few studies have specifically examined areas containing multiple fields. Few have considered the practical usage of satellite imagery for farmers (Sato et al. 2020). Fukumoto & Shinohara (2023) used Sentinel-2 imagery to evaluate the protein contents of rice and described the possibility of providing predictions to farmers for the subsequent year's cultivation. Reviewing crop growth, yield, and quality to plan upcoming cultivation is key to increasing the use of satellite imagery.

Satellite-based remote sensing can provide discrete time-series data at approximate intervals of a few days or weeks. Time-series analysis of satellite imagery is suitable for tracing in- or inter-season crop growth processes, although the relation between VIs and actual vegetation is indirect in some cases (Akbari et al. 2020, Zeng et al. 2020, Cavalaris et al. 2021, Shammi & Meng 2021). Among several methods of conducting time-series analysis, Functional Data Analysis (FDA) is versatile, producing discrete observed values for use with general statistical methods (Ramsey 1996, Morris 2015, Reiss et al. 2017, Xu et al. 2018). FDA can resolve three significant difficulties imposed by sequential observed data: noise disturbance, increased dadimensions, and uneven observation periods (Matsui 2019). Moreover, FDA may yield clues to improving crop management by analyzing satellite data. Mancini et al. (2024) applied FDA to Sentinel-2 imagery and demonstrated its applicability, but assessments using FDA specifically emphasizing crop productivity of farmer fields are insufficient.

For this study, conducted in Iwate Prefecture, Japan, we used a combine harvester with a yield monitor sensor to collect inter-field wheat yield variation data for multiple fields cultivated. The managed field areas were diverse because two or three small fields were combined in many cases to produce large fields while neglecting fallow areas, thereby increasing operational efficiency. The study aimed to clarify the conditions of inter-field wheat yield variation in agriculture, to verify the potential application of satellite imagery, and to apply FDA to develop a method to determine the relationship between time series NDVIs and yields. The study's purposes are to advance the practical use of satellite imagery for crop production and to increase the variety of exploratory analytical approaches.

Materials and methods

1. Study area and yield data collection

The 80 targeted wheat fields (0.06 ha-1.4 ha, total 23.6 ha) in the study (Fig. 1) were in Kitakami City, Iwate Prefecture, Japan (39° 29′ N, 141° 12′ E). In 2022, the Nanbu komugi winter wheat cultivar was sown at the rate of 13 g m^{-2} from 30 September to 25 October. The sowing date was recorded for each field. Basal fertilizer was applied at the rate of N:P_2O_5:K_2O = 5:5:4 g m^{-2}. During March 17-28, N = 6 g m^{-2} was top-dressed. The sowing and fertilization settings were representative, varying with the field conditions and operators' judgment. After the plants matured in July, grain weight data were collected using a combine harvester with a yield monitor sensor (WRH1200; Kubota Corp., Osaka, Japan). These data were converted to yield data per field with 12.5% moisture content and 86% yield rate, which were chosen after inspection following harvesting. The machinery manufacturer reported the estimation accuracy to be ± 5% when more than 500 kg of grain was harvested (Kubota Corp. 2022), but low yields showed poor accuracy. The model measures grain weight with the load cell set at the bottom of the grain tank (Miyachi et al. 2014, Kubota Corp. 2022). The load cell method was

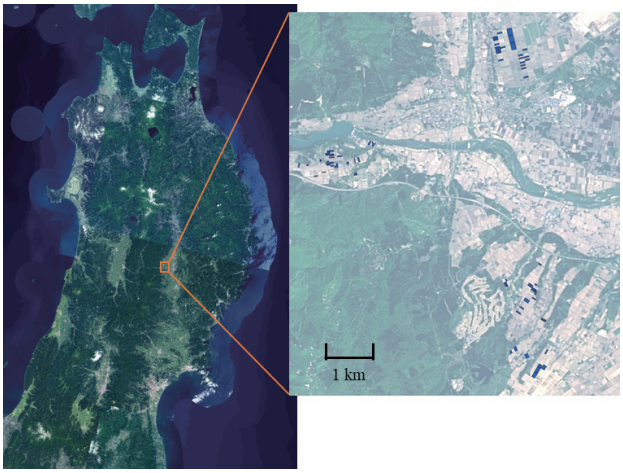

Fig. 1. Map of 80 target fields in Kitakami City, Iwate Prefecture, Japan
Blue quadrangles are target fields.

developed by Makino et al. (2007a), aiming at an estimation error of ± 5%. Models adopting this method show equivalent accuracy (Hirai et al. 2020). In fact, Makino et al. (2007b) validated the estimation accuracy for wheat grain weight using data from two combine harvesters with slightly different grain tanks and structures, demonstrating an estimation error of ± 5% for 90% of 36 targeted fields within a range of ± 8%. They assessed that accuracy from harvested grain weights of 31 kg-438 kg. Therefore, because of the grain weight range, we established a threshold of 30 kg to avoid inaccuracy.

2. Satellite data and NDVI

Sentinel-2 level-2A multispectral imagery was downloaded from the Google Earth Engine (GEE) platform (Alphabet Inc.). The imagery was corrected for atmosphere from level-1C using Sen2Cor (Main-Knorn et al. 2017). After manual removal of images including clouds and accumulated snow, seven days of data were left (28 March, 2 April, 27 April, 12 May, 17 May, 1 June, and 6 June) out of the 52 days observed during the wheat growing season, which covered the growth transition briefly despite the available rate of only 13%. Using those data, we calculated the time-series NDVI (n = 7) with a spatial resolution of 10 m × 10 m from two multispectral bands: band 4 (*Red*, 664.5 nm) and band 8 (near infrared: *NIR*, 835.1 nm).

$$NDVI = \frac{NIR - Red}{NIR + Red} \tag{1}$$

Field polygons were designed for each field and assigned mean pixel values using the zonal statistics command of QGIS 3.30 (QGIS Development Team 2023).

3. Functional regression with FDA

This study applied functional regression with FDA in two steps: (1) basis expansion and (2) regression using statistical methods as described below. The details of common principles were described by Matsui & Konishi (2011) and by Matsui (2019).

(1) Basis expansion

It is assumed that individual observed values were the sum of the original functional value and errors.

$$x_{i\alpha} = x_i(t_{i\alpha}) + e_{i\alpha} \tag{2}$$

In this equation, $x_{i\alpha}$ is the observed value at time α and data number i, realized from functional value $x_i(t_{i\alpha})$, which changes according to time $t_{i\alpha}$ and error $e_{i\alpha}$. To estimate target functional data $x_i(t)$, individual values were smoothed using the linear combination of several basis functions in basis expansion.

$$x_i(t) = \sum_{m=1}^{M} w_{im}\phi_m(t) = \boldsymbol{w}_i^T\boldsymbol{\phi}(t) \tag{3}$$

Here, $\phi_m(t)$ are basis functions at function number m, and $\boldsymbol{w}_i = (w_{i1}, \ldots, w_{iM})^T$ are the coefficients of $\boldsymbol{\phi}(t) = (\phi_1(t), \ldots, \phi_M(t))^T$. Estimating them gives functional data $x_i(t)$, for which T is the transpose. Many basis functions have been proposed. This study used the most general one, the B-spline basis function. We represented the observed NDVI as a linear combination of four B-spline basis functions.

(2) Regression

For sets of objective and explanatory variables $\{(y_i, x_i(t)); t \in \tau \in \mathbb{R}, i = 1, \ldots, n\}$, their relation is described as a functional regression.

$$y_i = \beta_0 + \int_\tau x_i(t)\beta_1(t)dt + \varepsilon_i \tag{4}$$

In this equation, y_i is the scalar objective variable, $x_i(t)$ denotes functional explanatory variables, β_0 is the intercept, $\beta_1(t)$ represents the coefficient function, and ε_i denotes the observation error with a mean of 0 and variance σ^2. As opposed to usual regression, $\beta_1(t)$ is a function described with basis functions.

$$\beta_1(t) = \sum_{m=1}^{M} b_{1m}\phi_m(t) = \boldsymbol{b}_i^T\boldsymbol{\phi}(t) \tag{5}$$

where $\boldsymbol{b}_1 = (b_{11}, \ldots, b_{1M})^T$ is an unknown parameter. Based on Equations (4) and (5), the functional regression (6) can be arranged as

$$\begin{aligned} y_i &= \beta_0 + \int_\tau w_i^T\phi(t)\phi^T(t)b_1dt + \varepsilon_i \\ &= \beta_0 + \sum_{m=1}^{M} \boldsymbol{w}_i^T\,\Phi\boldsymbol{b}_1 + \varepsilon_i = \boldsymbol{z}_i^T\boldsymbol{b} + \varepsilon_i \end{aligned} \tag{6}$$

where $\Phi = \int_\tau w_i^T\phi(t)\phi^T(t)dt$ represents the matrix, $z_i = (1, w_i^T\Phi)^T$ is a known vector, and $\boldsymbol{b} = (\beta_0, b_1^T)^T$ is an unknown vector parameter. Parameters are estimated from the right side of (6) using a technique such as the least squares method. There can be more than two explanatory variables, a requirement of multiple regression.

$$y_i = \beta_0 + \sum_{j=1}^{p_1} x_{ij}\beta_j$$

$$+\sum_{j=p_1+1}^{p}\int_{\tau} x_{ij}(t)\beta_j(t)dt+\varepsilon_i \quad (7)$$

In this equation, $x_{ij}\,(j=1,\ .\ .\ .\ ,p_1)$ is a scalar explanatory variable and $x_{ij}(t)\,(j=p_1+1,\ .\ .\ .\ ,p)$ is a functional explanatory variable. β_j is the coefficient of the scalar explanatory variable, and $\beta_j\,(t)$ is the coefficient function, which changes temporally and indicates the contribution of the functional explanatory variable. For this study, the yield is a scalar objective variable, and functional NDVI is a functional explanatory variable. The sowing date was a scalar explanatory variable used as a confounding factor, affecting both yield and NDVI to eliminate error variance.

3. Weather data, statistics, and programming

The daily average temperature and maximum snow depth data were taken using the Automated Meteorological Data Acquisition System (AMeDAS) at the Kitakami Observatory (39° 17′ 17.0″ N, 141° 06′ 40.1″ E) and downloaded from the Japan Meteorological Agency website. Statistical metrics of the regression were evaluated using the coefficient of determination (R^2), mean absolute error (MAE), and root mean square error (RMSE).

$$R^2=1-\frac{\sum_{i=1}^{n}(y_i-\hat{y})^2}{\sum_{i=1}^{n}(y_i-\bar{y})^2} \quad (8)$$

$$MAE=\frac{1}{N}\sum_{i=1}^{N}|y_i-\hat{y}_i| \quad (9)$$

$$RMSE=\sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_i-\hat{y}_i)^2} \quad (10)$$

The R programming language, version 4.3.1 (R core team 2023), was used for standard statistical processes of FDA, especially using the fda.usc package (Febrero-Bande & Fuente 2012).

Results

1. Inter-field yield variation

Harvested grain weights were 2 kg-3,324 kg. The lower two values of 2 and 8 kg were cut off before conversion to yield, so the number of fields became 78. Figure 2 shows a box plot of yields for 78 fields, exhibiting considerable inter-field yield variation. In addition, 486 and 365 g m^{-2} were excluded as outliers based on the upper limit of the third quartile + 1.5 × interquartile range (IQR). The highest, third quartile, median, mean, first quartile, and lowest value were, respectively, 361, 222, 186, 181, 129, and 17 g m^{-2}. Thus, the number of target fields for analysis became 76. Figure 3 shows the negative relation between the sowing date and yield of 76 fields, where the sowing dates in 2022 were 30 September through 26 October. Correlation analysis showed that both have significant correlation coefficients (r) for all

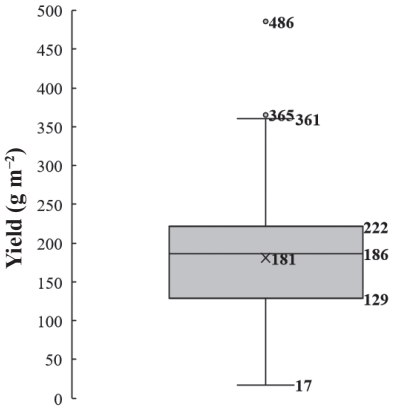

Fig. 2. Box plot of yield collected using a yield-monitoring harvester (n = 78)
The dots are outliers calculated based on the interquartile (IQR) value: the higher limit of the third quartile + 1.5 × IQR. The cross denotes the mean value.

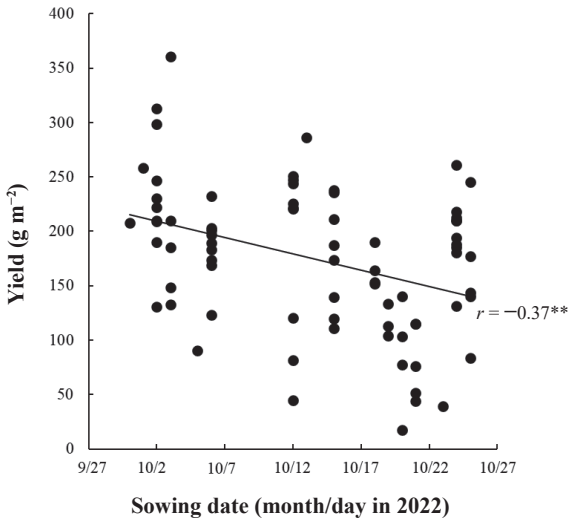

Fig. 3. Relation between sowing date in 2022 and yield (n = 76)
Significant modes of correlation coefficient are denoted as **, $p < 0.01$.

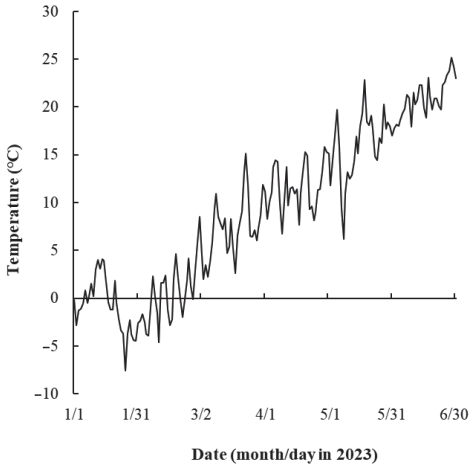

Fig. 4. Temperature during the January-June 2023 period

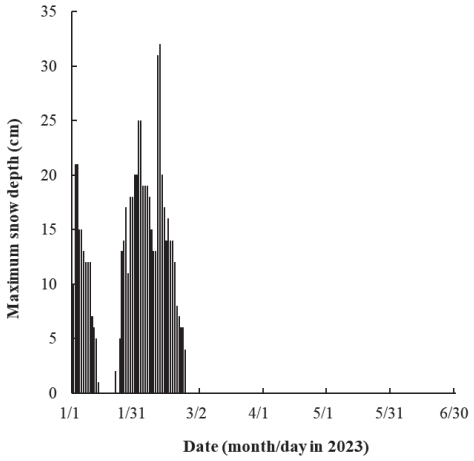

Fig. 5. Maximum snow depths during the January-June 2023 period

plots of $r = -0.37$ ($p < 0.01$). Therefore, the effect of sowing date on yield should not be ignored.

2. Weather data and observed NDVI data

Figures 4 and 5 show daily average temperatures and maximum snow depths from 1 January through 30 June in 2023, respectively. The January–March period was mainly wintery, with low temperatures and heavy snow. After that, temperatures increased. The remaining snow had disappeared entirely by April, promoting wheat growth. NDVI exhibited temporal changes, with transition patterns such as wheat growth, which accelerated from late March after the snow melted to flowering in mid-May, after which it dropped at maturity early June (Fig. 6). A few fields exhibited higher sustained values of NDVI, with some fields replaced in

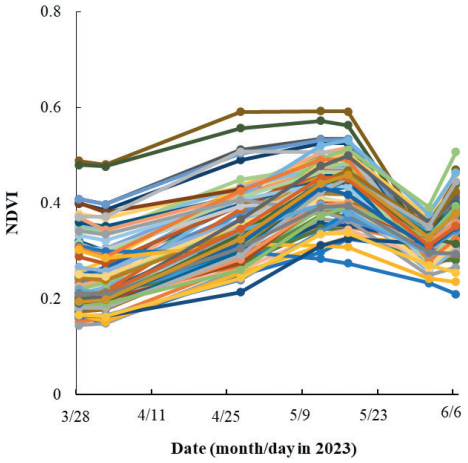

Fig. 6. Time series NDVI for each field shown in different colors (n = 76)
Points represent values observed in 2023, connected by solid straight lines.

NDVI order over time. The NDVI values were found to correlate with sowing dates, with significant negative correlation coefficients during development, which became uncorrelated with time (Table 1). Therefore, the sowing date was inferred to be a confounding factor affecting both yield and NDVI.

3. Basis expansion and functional regression with FDA

Using a combination of four B-spline basis functions (Fig. 7), time series NDVIs were represented well. Figure 8 shows transition patterns replaced by smoothed curves for typical higher and lower yields. The values of R^2, MAE, and RMSE for all fields were 0.89, 0.019, and 0.026, respectively, indicating that the point-to-curve conversion was suitable and that basis expansion was effective as the first step of FDA. To conduct functional regression for Equation (7), the sowing date and functional NDVI were used as two explanatory variables. The effect of the former was eliminated statistically as a confounding factor. Yield was the objective variable. The least squares method was applied to minimize residuals between the model output and the yield to obtain the regression result below.

$$yield = -0.063\ sowing\ date + [0.36\ Bspl_1 - 0.72\ Bspl_2 + 0.55\ Bspl_3 - 0.099\ Bspl_4]$$
$$(R^2 = 0.39, p < 0.001) \quad (11)$$

where $Bspl_{1\sim4}$ represent the four B-spline basis functions by date, as shown in Figure 7. This result shows that the

Table 1. Correlation coefficients between sowing dates and NDVI for each Sentinel-2 observation date in 2023

Observed day (DOY)	87	92	117	132	137	152	157
NDVI	−0.67***	−0.67***	−0.58***	−0.35**	−0.29**	0.29*	0.11n.s.

Asterisks to the right of a coefficient indicate the statistical significance of the correlation analysis:
***, $p < 0.001$; **, $p < 0.01$; *, $p < 0.05$; n.s., not significant.

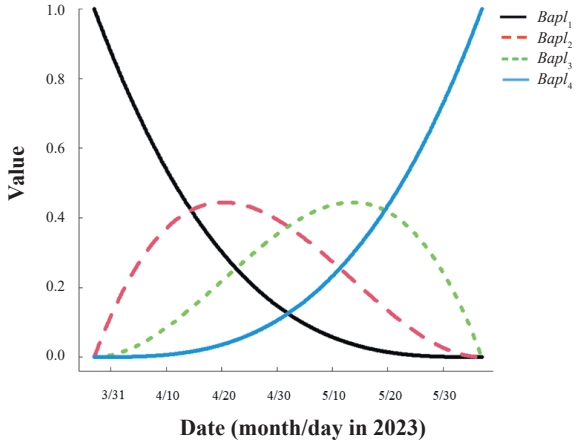

Fig. 7. Four B-spline basis functions ($Bspl_1$–$Bspl_4$) used for basis expansion
The Y axis represents the B-spline function values.

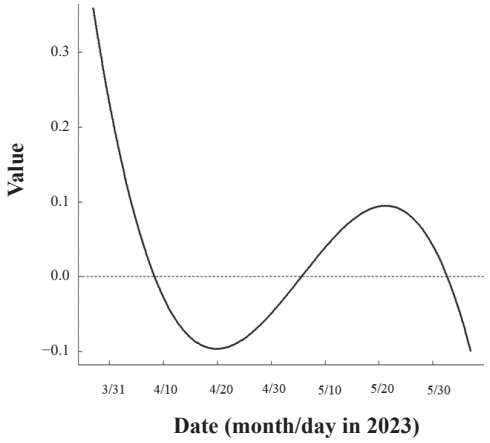

Fig. 9. Coefficient function of functional NDVI as an explanatory variable

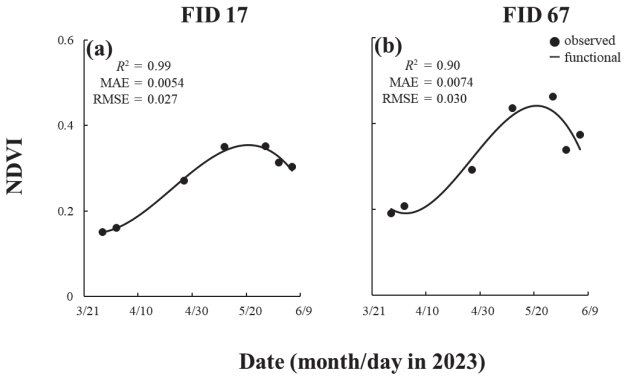

Fig. 8. Examples of observed and functionalized NDVI in 2023
Points and solid lines show observed and functionalized values, respectively. Field IDs (FIDs) 17 and 67 are representatives of lower and higher yields among all fields, respectively.

sowing date accounts for 39% of inter-field yield variation and functional NDVI converted to a linear combination of four B-spline basis functions placed in square brackets in Equation (11). The absolute value of the standardized partial regression coefficient is larger at each B-spline basis function than at the sowing date. In addition, the coefficient of determination of the single regression of yield and sowing date, as calculated from the relation in Figure 3, was 0.14. Therefore, functional NDVI was the primary factor affecting inter-field yield variation. To estimate the combined effect of the four B-spline basis functions, a coefficient function was calculated using Equation (11) with the fda.usc algorithms provided in R (Fig. 9). The temporal change of the coefficient function signifies positive (around early April and mid-May) or adverse effects (around mid-April and early June) on inter-field yield variation by NDVI.

Discussion

For this study, we assessed wheat yields in multiple farm fields and analyzed the relation between inter-field yield variation and satellite-based time series NDVI. FDA processing was applied to satellite data for two reasons. First, to eliminate fluctuations of time-series NDVI, smoothing was conducted by basis expansion. The fluctuations (Fig. 6) were probably caused by noise associated with atmospheric effects that were specific to the day (Jones & Vaughan 2010). The extent to which atmospheric effects were removed or left in satellite imagery after atmospheric pre-processing remains uncertain (Caselles & Lopez Garcia 1989, Rumora et al. 2020, Valdivieso-Ros et al. 2021). Smoothing is one means of reducing unnecessary effects reflected in the time-series data. Second, the FDA method was proposed

to manage a set of sequential data, suggesting a relation between inter-field yield variation and functional NDVI, as demonstrated by Equation (11). Regression analysis with a time-series variable that is analogous to dynamic crop growth can be done because understanding in-season whole growth processes offers potential insights into dry matter production. For instance, Yamamoto et al. (2023) sequentially arranged discrete soybean LAI data using nonlinear growth functions. This study was unable to present an explicit idea of dry matter production or yield constraints, but the regression results, excluding sowing date effects, implied that inter-field yield variation depends on differences in wheat growth dynamics. The coefficient function (Fig. 9) implied that NDVI had positive effects on yield in early April and mid-May. This finding might indicate the importance of obtaining the number of seedlings during early growth and maintaining vigor in mid-growth (Miyama et al. 1989). By contrast, adverse effects appeared in mid-April and in early June. The former may have been caused by lodging due to excess growth resulting from the longer period of early growth, and the latter due to a temporal mismatch between harvesting and maturity corresponding to the sowing date (Ogiuchi et al. 2004). Weather, drainage, and soil physiochemical properties linked to field locations affect wheat seedling and maturity. Farmers will be able to use those findings along with their own cultivation experiences to plan for subsequent years.

Among many other VIs, NDVI was selected because of its strong relation to crop growth (Aparicio et al. 2000, Cavalaris et al. 2021, Tenreiro et al. 2021). However, calibration using ground-based data was absent. Moreover, the conducted evaluation was relative. Wheat NDVI was generally low because of growth constraints, but NDVI often saturates as LAI increases (Doraiswamy et al. 2004, Herrmann et al. 2011, Bajocco et al. 2022). A careful assessment based on calibration is needed for large LAI. NDVI analysis without calibration makes it difficult to establish specific, optimal techniques for farmers at present. An important challenge is how satellite VIs are calibrated using ground truth data on biophysical crop parameters over different scales. Several studies have investigated calibration, demonstrating that biophysical parameters, mainly LAI, are linked only roughly with satellite VIs for some crops and trees (Ishii et al. 1999, Chen et al. 2002, Colombo et al. 2003, Morisette et al. 2006, Hasegawa et al. 2013). However, the number of verifications remains inadequate, especially for temporal crop growth in fields. Hashimoto et al. (2009) proposed a method to detect wheat LAI time-series patterns using their original index, aiming at establishing a robust relation based on biophysical processes by ground-observed reflectance. To improve the utility of this study's findings, a versatile method that is supported theoretically with biophysical information must be constructed, along with expanded use of satellite imagery for wide areas.

Yield data collected using a combine harvester with a yield monitor sensor provided information revealing actual wheat inter-field yield variation of 17-361 g m^{-2} (Figs. 2, 3), which has been rarely reported. With the constraints described above, the yield of the Nambu Komugi cultivar tends to be more unstable than others because it is a relatively old cultivar involving unimproved breeding characteristics, such as long culm leading to lodging or sensitivity to some diseases (Ishida et al. 1986, National Agriculture and Food Research Organization 2006). For example, Kowata et al. (2016) reported the average yield of Nambu Komugi in Iwate Prefecture as approximately 150 g m^{-2} during 2008-2015, whereas the yield of Yuki Chikara, the other primary cultivar there, was higher than 200 g m^{-2} during the entire period. The appropriate period for sowing winter wheat, such as Nambu Komugi in Iwate Prefecture, is early or mid-October in the subject area (Ogiuchi et al. 2004). If sowing is delayed, yield becomes low partly because of inadequate growth before snowfall, which causes snow mold damage (Yukawa et al. 1987). The snow tolerance of Nambu Komugi is strong, but snow mold damage was reported (Yukawa et al. 1987) and observed on parts of the fields in 2023. A negative correlation between yield and sowing date (Fig. 3) was presumed to be responsible for snow mold damage as well as inappropriate harvesting time against maturity. Moreover, precipitation was often observed immediately after sowing (data not shown), causing excessive soil moisture damage, which may have decreased the number of viable seedlings before snow accumulated in autumn 2022. This effect was reported from another study investigating wheat production in the same area (Yamamoto et al. 2025). Although the sowing date was late, the yields in 15 fields sown on October 24 and 25 were higher. These higher yields are partly attributable to good weather lasting about one week after sowing, in contrast to other days affected by rainfall.

A combine harvester with a yield monitor sensor is indispensable for yield investigation, having a ± 5% measurement accuracy when harvested grain weight exceeds a certain threshold. However, values with errors attributable to mechanical structures of machinery or a harvesting accident must be eliminated (Hidaka et al. 2009). For example, Hirai et al. (2020) reported that estimation errors are increased by an uneven distribution of harvested grain in a grain tank. From this study, and based on information from other studies, two lower grain

weights were excluded. Moreover, two higher yields were eliminated by an upper limit based on an outlier method. Also, asking workers about factors inhibiting harvesting operations may have biased the data. Even though some inspection of extreme values is warranted, analysis of yield variation in fields using a combine harvester with a yield monitor sensor has been demonstrated to be practical, with high applicability to research (Tanaka et al. 2023, Tanaka et al. 2024).

Large-scale farming is expected to make agriculture more profitable. The farming concept might be suitable for farmland in vast flat areas, but not in hilly and mountainous areas, presenting geographical difficulties when combining several farmland parcels (Yamashita & Kanahira 2022, Kawasaki et al. 2024). If farm fields are bounded and expanded only in flat areas with favorable conditions, then inter-field yield variation similar to that shown by this study (Fig. 2) might disturb the balance of the expected costs and benefits of efficient cultivation. By determining field conditions for a wide area before cultivation and by sensing crop growth during a season, farmers can improve their cultivation planning based on collected data. Data-driven approaches in agriculture using information and communications technologies such as remote sensing, GIS, cloud service, machinery, and machine learning (artificial intelligence, AI) have attracted special attention recently (Shiratani 2019, Inoue 2023, Nishimura 2023, Shimizu et al. 2024). Several services that collect data for use with data-driven approaches have become available, but their modes of use and analysis of collected data for practical cultivation and improved productivity remain insufficient for several reasons (Kawamura et al. 2020, Takahashi et al. 2020, Hashimoto et al. 2021, Ohdan 2023). Variable fertilization for rice based on UAV remote sensing data has become common, for instance (Kanmuri et al. 2018, Otani 2020), but other uses of collected data depend on the practitioner. The FDA process using the yield data and satellite imagery for this study was not applied directly, but site-specific management has become important for sustainable farming (Toriyama et al. 2003, Kato & Hayashi 2021). Evaluating inter-field yield variation related to dynamic crop growth, as clarified in this study, can contribute to future studies, eventually supporting and facilitating farm field management.

Acknowledgements

This study was partially supported by the On-farm Demonstration Trials of Smart Agriculture from the Ministry of Agriculture, Forestry and Fisheries (funding by the National Agriculture and Food Research Organization).

References

Akbari, E. et al. (2020) Crop mapping using random forest and particle swarm optimization based on multi-temporal Sentinel-2. *Remote Sens*., **12**, 1449.

Aparicio, N. et al. (2000) Spectral vegetation indices as nondestructive tools for determining durum wheat yield. *Agron. J*., **92**, 83-91.

Bajocco, S. et al. (2022) On the use of NDVI to estimate LAI in field crops: Implementing a conversion equation library. *Remote Sens*., **14**, 3554.

Caselles, V. & Lopez Garcia, M. J. (1989) An alternative simple approach to estimate atmospheric correction in multitemporal studies. *Int. J. Remote Sens*., **10**, 1127-1134.

Cavalaris, C. et al. (2021) Modeling of durum wheat yield based on Sentinel-2 imagery. *Agronomy*, **11**, 1486.

Chen, J. M. et al. (2002) Derivation and validation of Canada-wide coarse-resolution leaf area index maps using high-resolution satellite imagery and ground measurements. *Remote Sens. Environ*., **80**, 165-184.

Colombo, R. et al. (2003) Retrieval of leaf area index in different vegetation types using high resolution satellite data. *Remote Sens. Environ*., **86**, 120-131.

da Silva Junior, C. A. et al. (2020) Mapping soybean planting area in midwest Brazil with remotely sensed images and phenology-based algorithm using the Google Earth Engine platform. *Compt. Electron. Agric*., **169**, 105194.

Doraiswamy, P. C. et al. (2004) Crop condition and yield simulations using Landsat and MODIS. *Remote Sens. Environ*., **92**, 548-559.

Febrero-Bande, M. & Fuente, M. O. (2012) Statistical computing in functional data analysis: The R package fda. usc. *J. Stat. Softw*., **51**, 1-28.

Fukumoto, M. & Shinohara, K. (2023) Influence of different rice growth stages on NDVI in the latter half of growth and relative evaluation of rice yield and rice protein content by NDVI. *Nogyo noson kogakkai ronbunshu* (*Trans. JSIDRE*.), **31**, 9-18 [In Japanese with English summary].

Gao, X. et al. (2000) Optical–biophysical relationships of vegetation spectra without background contamination. *Remote Sens. Environ*., **74**, 609-620.

Gong, Z. et al. (2020) Vegetation dynamics and phenological shifts in long-term NDVI time series in Inner Mongolia, China. *JARQ*, **54**, 101-112.

Hasegawa, K. et al. (2013) Estimation of leaf area index of *Cryptomeria japonica* using various methods: A case study of Aso District, Kumamoto Prefecture. *Chigaku zasshi* (*J. Geog*.), **122**, 875-891 [In Japanese with English summary].

Hashimoto, N. et al. (2009) Study of a method for extracting LAI time-series patterns for the estimations of crop phenology. *Nihon remote sensing gakkaishi* (*J. Remote Sens. Soc. Jpn*.), **29**, 381-391 [In Japanese with English summary].

Hashimoto, N. et al. (2021) Detection of leaf area expansion promoted by additional fertilizer application on farmer's paddy fields using multispectral images. *Nihon sakumotsu gakkai kiji* (*Jpn J. Crop Sci*.), **90**, 211-221 [In Japanese with English summary].

Herrmann, I. et al. (2011) LAI assessment of wheat and potato crops by VENμS and Sentinel-2 bands. *Remote Sens. Environ*., **115**, 2141-2151.

Hidaka, Y. et al. (2009) Development of combine harvester with yield monitoring function (part 4) – schematic of practical model and result of field monitor test –. *Nogyo kikai gakkaishi* (*J. JSAM*.), **71**, 60-68 [In Japanese with English summary].

Hirai, Y. et al. (2020) Calibration equations for the weight of rough rice layers of various shapes piled in a grain tank. *Nogyo shokuryo kogakkaishi* (*J. JSAM*.), **82**, 362-369 [In Japanese with English summary].

Huete, A. et al. (2002) Overview of the radiometric and biophysical performance of the MODIS vegetation indices. *Remote Sens. Environ*., **83**, 195-213.

Inoue, Y. (2020) Satellite- and drone-based remote sensing of crops and soils for smart farming – a review. *Soil Sci. Plant Nutr*., **66**, 798-810.

Inoue, Y. (2023) Utility and caveats of sensing and data science for smartification of crop production – remote sensing, AI, big data, and phenotyping –. *Nihon sakumotsu gakkai kiji* (*Jpn. J. Crop Sci*.), **92**, 91-103 [In Japanese with English summary].

Ishida, R. et al. (1986) Adaptability Test of Wheat Cultivars in Snowy District. *Hokuriku sakumotsu gakkaiho* (*Hokuriku Crop Sci*.), **21**, 29-31 [In Japanese].

Ishii, T. et al. (1999) Estimation of leaf area index using remote sensing data. *Suimon, mizu shigen gakkaishi* (*J. Jpn. Soc. Hydro. Water Resour*.), **12**, 210-220 [In Japanese with English summary].

Ishikawa, T. et al. (2021) Construction and utilization of field-specific datasets for data-driven large-scale rice cultivation. *Nihon sakumotsu gakkai kiji* (*Jpn. J. Crop Sci*.), **90**, 222-229 [In Japanese with English summary].

Jiang, Z. et al. (2008) Development of a two-band enhanced vegetation index without a blue band. *Remote Sens. Environ*., **112**, 3833-3845.

Jones, H. G. & Vaughan, R. A. (2013) *Shokusei no remote sensing* (First edition). Morikita Shuppan, Tokyo, Japan, pp. 175-177 [In Japanese].

Kanmuri, H. et al. (2018) Demonstration experiment of dry direct seeding method of rice in 5.8 ha paddy field expanded from standard 2 ha paddy field. *Nogyo noson kogakkai ronbunshu* (*Trans. JSIDRE*.), **86**, 1121-1124 [In Japanese].

Kato, Y. & Hayashi, K. (2021) Japan and IRRI: Contributions to International Rice Research for sustainable development, lessons learned and ways forward. *JARQ*, **55**, 483-487.

Kawamura, T. et al. (2022) Research data platform for data-driven agriculture. *Joho shori gakkaishi* (*J. Inform. Proc*.), **2**, 21-31 [In Japanese with English summary].

Kawasaki, Y. et al. (2024) Narrow row and dense planting of new black soybean cultivar 'Kuroshofuku' in southwestern Japan. *Nihon sakumotsu gakkai kiji* (*Jpn. J. Crop Sci*.), **93**, 132-139 [In Japanese with English summary].

Kowata, H. et al. (2016) Wheat grain characteristics and factors affecting the quality of main cultivars in Iwate Prefecture. *Nihon sakumotsu gakkai Tohoku shibukaiho* (*Tohoku J. Crop Sci*.), **59**, 1-4 [In Japanese].

Kubota Corp. (2022) WORLD WRH1200. https://agriculture.kubota.co.jp/img_sys/catalog/5-20-2-0006-04%20(1).pdf. Accessed on 22 November 2024 [In Japanese].

Main-Knorn, M. et al. (2017) Sen2Cor for Sentinel-2. *Image Signal Process. Remote Sens. XXIII*, **10427**, 37-48.

Makino, E. et al. (2007a) Development of combine harvester with yield monitoring function (part 1) – basic design of the yield monitoring system and continuous measurement of yield and moisture –. *Nogyo kikai gakkaishi* (*J. JSAM*.), **69**, 79-88 [In Japanese with English summary].

Makino, E. et al. (2007b). Development of combine harvester with yield monitoring function (part 3) – monitoring yield and map production in paddy and wheat fields –. *Nogyo kikai gakkaishi* (*J. JSAM*.), **69**, 95-103 [In Japanese with English summary].

Mancini, A. et al. (2024) Time series from Sentinel-2 for organic durum wheat yield prediction using functional data analysis and deep learning. *Agronomy*, **14**, 109.

Matsui, H. (2019) Statistical modeling via functional data analysis. *Tokei suri* (*Proc. Inst. Statist. Math*.), **67**, 73-96 [In Japanese with English summary].

Matsui, H. & Konishi, S. (2011) Variable selection for functional regression models via the L1 regularization. *Comput. Stat. Data Anal*., **55**, 3304-3310.

Miyachi, K. et al. (2014) Development of the farming support system "Kubota Smart Agricultural System (KSAS)". *Nogyo shokuryo kogakkai* (*J. JSAM*.), **76**, 284-288 [In Japanese].

Miyama, M. et al. (1989) A Suitable Nitrogen Fertilization Method for Several Crop Plants Based on the Optimum Nitrogen Contents; 3. The Optimum Amount of Nitrogen for Wheat Crop to be Absorbed at Each Growth Stage. *Nihon dojo hiryo kagaku zasshi* (*Jpn. J. Soil Sci. Plant Nutr*.), **60**, 106-115 [In Japanese with English summary].

Morisette, J. T. et al. (2006) Validation of global moderate-resolution LAI products: A framework proposed within the CEOS land product validation subgroup. *IEEE Trans. Geosci. Remote Sens*., **44**, 1804-1817.

Morris, J. S. (2015) Functional regression. *Ann. Rev. Stat. Appl*., **2**, 321-359.

National Agriculture and Food Research Organization (2006) *Nogyo gijutsu jiten*: *NAROPEDIA*. https://lib.ruralnet.or.jp/nrpd/#koumoku=13897. Accessed on 5 June 2025. [In Japanese]

Nishimura, K. (2023) Development of QAgriSupport, a GIS-based agricultural production process management system, and foregis, a mobile application. *JARQ*, **57**, 79-87.

Ogiuchi, K. et al. (2004) Optimum Seeding Date and Seeding Density for Winter-Seeding Cultivation of Winter Wheat in Iwate. *Nihon sakumotsu gakkai kiji* (*Jpn. J. Crop Sci*.), **73**, 396-401 [In Japanese with English summary].

Ohdan, H. (2023) Drone-based remote sensing in agriculture production. *Nihon gazo gakkaishi* (*J. Image. Soc. Jpn*.), **62**, 219-224.

Otani, R. (2020) Application of advanced agricultural methods to restore agricultural areas devastated by a Tsunami on the Sendai Plain. *Nogyo keizai kenkyu* (*J. Rural Soc. Econ*.), **38**, 18-23 [In Japanese with English summary].

QGIS Development Team (2023) QGIS Geographic Information System. *Open-Source Geospatial Foundation Project*. http://qgis.osgeo.org. Accessed on 20 November 2024.

Ramsey, J. O. (1996) Principal differential analysis: Data reduction by differential operators. *J. Royal Stat. Society – Series B*, **58**, 495-508.

R Core Team (2023) *R*: *A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.

Reiss, P. T. et al. (2017) Method for scalar-on-function regression. *Int. Stat. Rev.*, **85**, 228-249.

Rodigheri, G. et al. (2023) Estimating crop sowing and harvesting dates using satellite vegetation index: A comparative analysis. *Remote Sens*., **15**, 5366.

Rumora, L. et al. (2020) Impact of various atmospheric corrections on Sentinel-2 land cover classification accuracy using machine learning classifiers. *ISPRS Int. J. Geo-Inf*., **9**, 277.

Sato, T. et al. (2020) Determinants of Cropping: Using Object-Based Remote Sensing Analysis to Construct Field Plot-level Polygon Panel Data. *Nogyo joho kenkyu* (*Agric. Inform. Res.*), 14-23 [In Japanese with English summary].

Shammi, S. A. & Meng, Q. (2021) Use time series NDVI and EVI to develop dynamic crop growth metrics for yield modeling. *Ecol. Indic*., **121**, 107124.

Shimizu, Y. et al. (2024) Achieving both yield increase and farm area expansion in large-scale paddy farming by accumulating and utilizing agricultural data and smart agricultural machinery. *Nogyo joho kenkyu* (*Agric. Inform. Res.*), **33**, 14-26 [In Japanese with English summary].

Shiratani, E. (2019) R&Ds for Society 5.0 in agriculture. *Nogyo noson kogakkai ronbunshu* (*Trans. JSIDRE.*), **87**, 741-744 [In Japanese].

Sonia et al. (2022) Geospatial evaluation of cropping pattern and cropping intensity using multi temporal harmonized product of Sentinel-2 dataset on Google Earth Engine. *Appl. Sci*., **12**, 12583.

Takahashi, T. et al. (2020) New method of analysis and measures of soil physical properties for large-scale managements of paddy fields. *Nihon dojo hiryo kagaku zasshi* (*Jpn. J. Soil Sci. Plant Nutr.)*, **91**, 172-177 [In Japanese].

Tanaka, T. S. T. et al. (2023) Toward an effective approach for on-farm experimentation: Lessons learned from a case study of fertilizer application optimization in Japan. *Prec. Agr*., **24**, 2044-2060.

Tanaka, T. S. T. et al. (2024) Can machine learning models provide accurate fertilizer recommendations? *Prec. Agr*., **25**, 1839-1856.

Tenreiro, T. R. et al. (2021) Using NDVI for the assessment of canopy cover in agricultural crops within modelling research. *Comput. Electron. Agric*., **182**, 106038.

Toriyama, K. et al. (2003) Development of a site-specific nitrogen management system for paddy rice. *JARQ*, **37**, 213-218.

Tucker, C. J. (1979) Red and photographic infrared linear combinations for monitoring vegetation. *Remote Sens. Environ*., **8**, 127-150.

Umemoto, M. (2019) Technological innovation in Japanese agriculture – Progress and prospect –. *Nogyo keizai kenkyu* (*J. Rural Econ.*), **91**, 207-220 [In Japanese with English summary].

Valdivieso-Ros, C. et al. (2021) Effect of different atmospheric correction algorithms on Sentinel-2 imagery classification accuracy in a semiarid Mediterranean area. *Remote Sens*., **13**, 1770.

Xu, Y. et al. (2018) Functional modeling of plant growth dynamics. *Plant Phenome. J*., **1**, 1-10.

Yamamoto, S. et al. (2023) Evaluation of LAI dynamics by using plant canopy analyzer and its relationship to yield variation of soybean in farmer field. *Agriculture*, **13**, 609.

Yamamoto, S. et al (2025) Analysis of Wheat Productivity Improvement with Open Ditch and Underdrain in Farmer Fields Using a Yield Monitoring Combine Harvester and UAV Remote Sensing: Evaluation of On-site Demonstration Experiment for Upland Agricultural Corporation Fields in Kitakami City, Iwate Prefecture, *Nihon sakumotsu gakkai kiji* (*Jpn. J. Crop Sci*.), **94**. 148-158 [In Japanese with English summary].

Yamashita, R. & Kanahira, K. (2022) Diversity of perception and expectation on agricultural management entities differentiating the evaluation of farmland accumulation in the two prefectures, Hokuriku district. *Nogyo noson kogakkaishi* (*Trans. JSIDRE.*), **90**, 495-498 [In Japanese].

Yukawa, T. et al. (1987) Relationship between Seeding Time, Growth in Autumn, and Snow Damage of Wheat. *Hokuriiku sakumotsu gakkaiho* (*Hokuriku Crop Sci*.), **22**, 57-61 [In Japanese].

Zeng, L. et al. (2020) A review of vegetation phenological metrics extraction using time-series, multispectral satellite data. *Remote Sens. Environ*., **237**, 111511.

Zhang, C. et al. (2020) High-resolution satellite imagery applications in crop phenotyping: An overview. *Compt. Electron. Agric*., **175**, 105584.