



## Easy yield mapping for precision agriculture

Omran Alshihabi, Kristin Persson & Mats Söderström

To cite this article: Omran Alshihabi, Kristin Persson & Mats Söderström (2024) Easy yield mapping for precision agriculture, Acta Agriculturae Scandinavica, Section B — Soil & Plant Science, 74:1, 2411950, DOI: [10.1080/09064710.2024.2411950](https://doi.org/10.1080/09064710.2024.2411950)

To link to this article: <https://doi.org/10.1080/09064710.2024.2411950>



© 2024 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group



Published online: 03 Nov 2024.



Submit your article to this journal [↗](#)



Article views: 241



View related articles [↗](#)



View Crossmark data [↗](#)

## Easy yield mapping for precision agriculture

Omran Alshihabi, Kristin Persson and Mats Söderström

Department of Soil & Environment, Swedish University of Agricultural Sciences (SLU), Skara, Sweden

### ABSTRACT

Within precision agriculture, yield mapping is important in the evaluation of crop management and delineation of management zones. It can also be used to assess within-field yield potential, in order to guide different precision agriculture practices. However, some farmers do not have a yield monitoring system, and some who do may obtain incomplete or erroneous yield data. This study examined the accuracy with which winter wheat (*Triticum aestivum* L.) yield could be mapped in 18 fields in southern Sweden using a simple empirical relationship between Sentinel-2 (ESA, Paris, France) data, vegetation index (VI) maps and combined harvester data collected in nearby fields. The results showed that a decrease in map resolution to 40 m reduced the error in the yield maps obtained. Normalised difference water index (NDWI) was the most efficient VI, while a combination of satellite data from earlier and later plant development (booting and milk development stages) performed slightly better than data for other development stages and combinations. The best-performing model at a within-field scale (40-m resolution) had an average mean absolute error (MAE) of 0.40 tonnes ha<sup>-1</sup> in a leave-one-field-out cross validation. When the prediction model at field-means scale was applied on 69 farms in a 1055 km<sup>2</sup> area, MAE was 0.75 tonnes ha<sup>-1</sup> when comparing predictions with mean yields reported by farmers in a phone survey. Therefore, if adequate combined harvester and/or mean yield data are available, a modelling framework that translates satellite imagery into yield maps on-the-fly could be made available for different stakeholders via decision support systems for precision agriculture.

### ARTICLE HISTORY

Received 13 March 2024  
Accepted 30 September 2024

### KEYWORDS



Combine harvester; remote sensing; Sentinel-2; spatial variation; vegetation index; wheat

## Introduction

Knowledge of crop yield is essential for efficient use of natural resources in field crop production, as part of work towards the global Sustainable Development Goals (SDGs) (United Nations 2015). Yield monitors have been tested and used since the late 1980s and early 1990s and have been key in the development of precision agriculture, because they make it possible to define, quantify and characterise within-field variation in crop production (see early work by e.g. Schueller and Bae 1987; Schnug et al. 1993; Birrell et al. 1996; Stafford et al. 1996; Nissen and Söderström 1999). Yield monitors are mounted on combine harvesters and measure in real-time the amount of grain passing through the combine while the crop is being harvested (Reyns et al. 2002). However, it is still uncommon for yield maps to be a decisive component of the decision-making process in precision agriculture. From a more practical perspective, end-users cannot obtain the yield information until after the growing season,

which might be a limitation for decision-making. Maps of field fertility or yield stability (i.e. generalised maps based on yield maps for many years) have been developed in research projects over recent decades (e.g. Blackmore 2000; Ping and Dobermann 2005; Blasch et al. 2020) but, like single-year yield maps, they are under-utilised (Basso and Antle 2020). Fields are often split into constantly high-yielding zones, constantly low-yielding zones and unstable zones that produce different yields e.g. depending on whether the season is wet or dry (e.g. Delin 2005; Maestrini and Basso 2018 and 2021).

Yield data from combine harvesters often include large numbers of defective observations during the passage of the combine harvester inside fields, as a result of e.g. flow delay, filling and emptying times, abrupt speed changes or partially-used cutting bar (e.g. Blackmore and Moore 1999; Thylén et al. 2000; Simbahan et al. 2004; Sudduth and Drummond 2007). This may be one reason why many farmers who use the equipment find it difficult to create reliable yield maps.

**CONTACT** Omran Alshihabi  [omran.alshihabi@slu.se](mailto:omran.alshihabi@slu.se)  Department of Soil & Environment, Swedish University of Agricultural Sciences (SLU), P.O. Box 234, SE-532 23 Skara, Sweden

© 2024 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group  
This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

Moreover, there is currently no system available for actually using yield maps. A French survey found that best-equipped farmers used their systems only for real-time visualisation during harvest and that fewer than 25% were producing and using yield maps (Lachia et al. 2021).

Vegetation indices (VIs) from optical crop canopy measurements have since long been known to correlate with crop yield (e.g. Curran 1980; Ashcroft et al. 1990). New advances in satellite data acquisition and processing offer promise for satellite-based yield mapping as an alternative method to derive yield maps of individual fields at resolution relevant to precision agriculture. Current methods often rely on calibrated relationships between VIs and yield that are specific to individual locations and years, often with new ground measurements needed for each new setting (Panek et al. 2020). Some studies combine satellite data with ancillary information (Hunt et al. 2019; Franz et al. 2020; Shuai and Basso 2022). Both simple empirical relationships (e.g. Gaso et al. 2019; Söderström et al. 2021; Lyle et al. 2023) and more complex machine learning models (e.g. Liao et al. 2023; Perich et al. 2023) are used, while other studies, incorporate satellite data into mechanistic crop models to make spatial predictions of yield for agricultural fields (e.g. Bouras et al. 2023; Luo et al. 2023; Sadeh et al. 2024).

Ulfa et al. (2022) found only marginal differences in the performance of the different VIs tested, which mostly showed good predictions of the spatial pattern of yield, but only modest performance in terms of predictions of absolute yield (when assessed via within-field metrics with leave-one-field-out cross-validation). In terms of the crop development stage, data from the peak biomass stage showed the best performance in that study and combining data from the same VI but from multiple stages did not improve predictions. Results from stepwise analysis revealed that some simple combinations of different VIs from different stages could give better absolute yield predictions but did not outperform the single-index models alone, in terms of prediction of the spatial pattern of the high and low-yielding areas. Stettmer et al. (2022) examined three different site-specific yield-mapping methods (tractor-borne sensor, satellite, combine harvester) for winter wheat (*Triticum aestivum* L.) to evaluate their precision and suitability for delineation of management zones for site-specific crop management. The study showed differences in the precision and accuracy of the investigated methods due to the deviations in the absolute yields, and the results were only suitable for yield potential maps to a limited extent, and further research was required to improve the results. Vallentin et al. (2022) examined the

relationship between remote sensing data, soil and relief data and crop yield data in north-east Germany and concluded that the correlation is strongest when a field and its crop are spatially heterogeneous and when a suitable phenological stage of the crop is reached at the time of satellite imaging. They found that satellite images with higher resolution, such as RapidEye (Planet Labs PBC, San Francisco, CA, USA) and Sentinel-2 (ESA, Paris, France), perform better than the lower-resolution sensors of the Landsat 5, 7 and 8 series (NASA, Washington, CO, USA).

The aims of this study were to assess the accuracy with which winter wheat yield can be mapped by Sentinel-2 images through simple empirical models parameterised by combining yield data collected in a specific local area and to propose a modelling framework for after-harvest satellite-based yield mapping. The yield predictions were evaluated at different spatial resolutions and at different growth stages. In addition, yield predictions were evaluated when the models obtained were applied in predicting farm mean yields in a larger district in southern Sweden within the same satellite scene. Specific research questions were to determine:

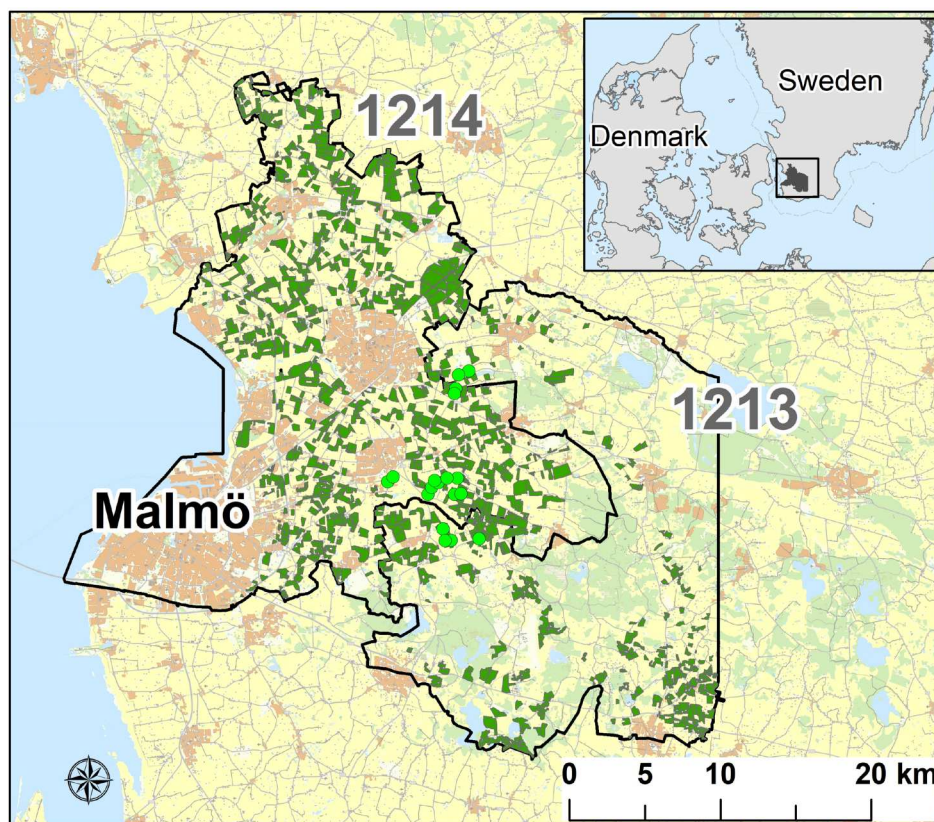
- (1) Differences in the magnitude of error of yield maps produced using different VIs and VI combinations.
- (2) The impact on model performance of using satellite images from different plant development stages and combinations of development stages.
- (3) The optimal spatial resolution for yield mapping using this approach (20, 40 m or field mean).
- (4) The accuracy of the models obtained when applied to estimate farm-average yield across a larger area.

## Materials and methods

This work builds on a pilot study by Alshihabi et al. (2023) in which aspects of research questions 1–4 were addressed using top-of-atmosphere reflectance (L1C, Sentinel-2). The same combined harvester data were used in the present study, but the remote sensing data were bottom-of-atmosphere reflectance (L2A, Sentinel-2). In addition, the models obtained were evaluated for more VI combinations and crop development stages, and for farm mean yields in a surrounding yield district by comparison with independent data from a yield survey.

## Study area

The study was conducted in southernmost Sweden (Figure 1), in an area with about 50% arable land. Small-grain crops (winter wheat is most common), oilseed rape (*Brassica rapa* L.), sugarbeet (*Beta vulgaris*



**Figure 1.** Location of yield districts (SKO) 1213 and 1214 near the city of Malmö in southern Sweden, within the 33UUB Sentinel-2 tile. Arable land is yellow and fields with winter wheat in 2019 are dark green. The locations of the 18 combined harvested fields are marked with light green circles. Other colours: beige = built-up area; green = forest; blue = water; grey = road network; white = other. Background map based on Swedish Environmental Protection Agency (2023) and Swedish Board of Agriculture (2019).

subsp.) and potatoes (*Solanum tuberosum* L.) are commonly grown in the area.

### Yield data

Two types of yield data were used: (a) within-field yield data collected by a combine harvester, and (b) farm-average yield data collected via interviews with farmers performed by the National Statistics Bureau, Statistics Sweden (SCB, Örebro, Sweden). The within-field yield data were obtained from 18 fields cropped with winter wheat (Figure 1). These fields were distributed across a 10 km × 4 km area and all were managed by the same farmer, using a John Deere yield monitor (John Deere, Moline, IL, USA). The farm-average yield data were for two Swedish yield districts, SKO 1213 and 1214 (following the terminology of SCB) (Figure 1), which occupy an area of some 1055 km<sup>2</sup> within a single Sentinel-2 satellite scene (tile 33UUB). Sweden is divided into 107 SKOs for statistical purposes, with each SKO assumed to be relatively homogeneous with regard to the variation and level of crop yield from year to year. Therefore, it can be convenient to

study the responses of an agricultural system at that spatial scale. A yield survey is carried out every year by SCB within all 107 SKOs, where randomly selected farmers are invited to report the average yields of crops cultivated on their farms. In the present study, SCB data for 69 farms with a total of 494 winter wheat fields within a 1055 km<sup>2</sup> area were included, with yield reported as an average value for all fields on a farm.

### Satellite data

Satellite images from Sentinel-2 (L2A processing level, i.e. orthorectified surface reflectance) for Zadoks crop development stages DC 49–85 (Zadoks et al. 1974) in the fields were downloaded from ESA (<https://dataspace.copernicus.eu/>). Only images that were cloud-free for the selected fields on the four selected dates (4, 14, 29 June, 11 July 2019) were used. These acquisition dates coincided approximately with DC stages 49–54, 61–65, 71–75 and 83–85, respectively (based on data from the Swedish Board of Agriculture). The Sentinel-2 mission collects data in 13 bands, in the

visible to short-wave infrared region of the electromagnetic spectrum (ESA 2024).

### Data preparation

Combine harvester yield data were filtered to remove potential erroneous records. We used an interactive procedure based on the approach by Sudduth and Drummond (2007), taking into consideration registrations in the combination of cutting board width (only full cutting board was kept), speed, time after start and before stop, and yield. The settings (min or max) of the filtering were adapted to the data variation in each field, based on visual inspection and the histograms, in order to remove outliers without affecting areas of true yield variation. Unfortunately, we have no uncertainty estimation of the combined data. In this case, however, we parameterise and validate the models based on the same combine, which means what is predicted by the models is ‘yield as measured by this combine’. In a system, where multiple combines would be used together in a model, this would steer the model in different directions and to some extent be evened out. The best would of course be if all combined data used for model calibration would be well calibrated. That may be used as criteria for data sharing. The quality of the yield data was generally judged to be good and records were available for all parts of the fields (for a typical example, see Figure 2a). In order to combine yield data with Sentinel-2 data, two rasters with spatial resolutions 20 and 40, aligned with the Sentinel-2 grid, were generated from the combined harvester yield data. The raster cell values were calculated as a mean of all yield data records within each raster cell (typically, there were 14 measurements per 20 m grid cell and 56 measurements per cell in the 40 m grid), to enable as direct comparison as possible between the satellite data and yield data. The raster layers generated were

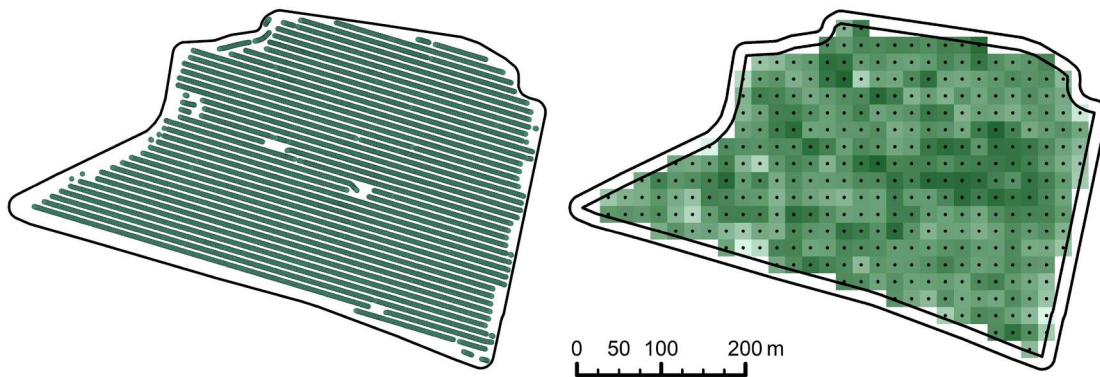
converted into point layers (pixel centroids), for which information from the different bands of Sentinel – 2 were extracted (re-sampled by bilinear method) and then clipped within buffer zones inside the field boundaries (15 m for the 20 m resolution raster and 25 m for the 40 m resolution raster), to ensure boundary cells were omitted and to avoid edge effects (Figure 2b).

Five normalised difference indices were calculated for each satellite image: (i) normalised difference vegetation index (NDVI), (ii) normalised difference water index (NDWI, combination of near-infrared and shortwave infrared bands), (iii) normalised difference red-edge vegetation index (NDRE75, bands in the red-edge region), (iv) normalised difference drought index (NDDI, based on NDVI and NDWI, believed to be sensitive to drought) and (v) green-red vegetation index (GRVI, based only on bands of visible light). The choices of the VIs were selected to correlate with different crop properties and to cover different regions of the electromagnetic spectrum: biomass and general crop vigour (NDVI, NDRE, and GRVI), water stress in the crop (NDWI, and NDDI). GRVI is based on visual bands, whereas the other indices also use infrared bands. The indices were computed using Equations 1-5, respectively, where  $\rho$  is reflectance and subscript numbers denote Sentinel-2 bands (Rouse et al. 1974; Adamsen et al. 1999; Sims and Gamon 2002; Gu et al. 2007; Zhang et al. 2017).

$$NDVI = \frac{\rho_8 - \rho_4}{\rho_8 + \rho_4} \quad (1)$$

$$NDWI = \frac{\rho_{8A} - \rho_{11}}{\rho_{8A} + \rho_{11}} \quad (2)$$

$$NDRE75 = \frac{\rho_7 - \rho_5}{\rho_7 + \rho_5} \quad (3)$$



**Figure 2.** (a) Left: filtered and quality-controlled combined yield recordings for one of the fields studied and (b) Right: masked point layer and the raster with a spatial resolution of 20 of the same field, where darker hue indicates higher yield.

$$NDDI = \frac{NDVI - NDWI}{NDVI + NDWI} \quad (4)$$

$$GRVI = \frac{\rho_3 - \rho_4}{\rho_3 + \rho_4} \quad (5)$$

### Data analyses

Univariate and multivariate linear regression models were used for yield predictions based on single VIs or pairwise combinations of VIs (e.g. NDVI-NDRE75, NDVI-NDWI, and NDWI-NDRE75). This was done for the different resolutions (20, 40 m and field mean) for the four individual plant development stages and for all combinations of 2–4 development stages. This resulted in a total of 675 date-VI combinations (225 per resolution). Models were evaluated by the leave-one-field-out cross-validation approach, using the R programming language (R Core Team 2023). For this, one entire field was left out and the other fields were used to develop a model to predict yield in the left-out field. The procedure was repeated until all fields had been left out once. Predicted values from the cross-validation were compared with the corresponding observed values, by computing mean absolute error (MAE) for all three spatial resolutions and coefficient of determination ( $r^2$ ) for a linear regression line between predicted and observed values for the 20 m and 40 m resolutions. This was done individually for each field. The evaluation metrics were computed using Equations 6 and 7, where  $o_i$  and  $p_i$  are observed and predicted values, respectively, for a pixel or a field ( $i$ ) and  $\bar{o}$  and  $\bar{p}$  are mean of all  $n$  observed or predicted pixel values, respectively.

$$MAE = \frac{\sum_{i=1}^n |p_i - o_i|}{n} \quad (6)$$

$$r^2 = \frac{\sum_{i=1}^n (o_i - \bar{o})(p_i - \bar{p})}{\sqrt{(\sum_{i=1}^n (o_i - \bar{o})^2 \sum_{i=1}^n (p_i - \bar{p})^2)}} \quad (7)$$

These two evaluation metrics have different strengths and weaknesses. So for identification of the best model, they were combined. Thus for each resolution and field, the mean of ranks of 225 values of MAE (smaller to larger) and  $r^2$  (larger to smaller) values were computed and then averaged for all fields. The model with the smallest mean rank values was considered to be the best model. The final models were parameterised using all data (no field left out) and deployed across the fields for the production of yield maps. The best model was also deployed on the mean VI values for each of the SCB farms. The yield values obtained in interviews with farmers were compared with these predicted yields, to assess the possibility of upscaling the local mapping model to the district level.

In the present study, we tested relationships between crop yield and one or more, satellite-based vegetation indices from one or more occasions that could be used to map yields. The rationale behind this is that the crop vigour during the season should reflect the yield level, irrespective of the cause of the yield level. It did not form part of this study to explore the drivers of spatial yield variation in the eighteen fields.

## Results

### Descriptive statistics

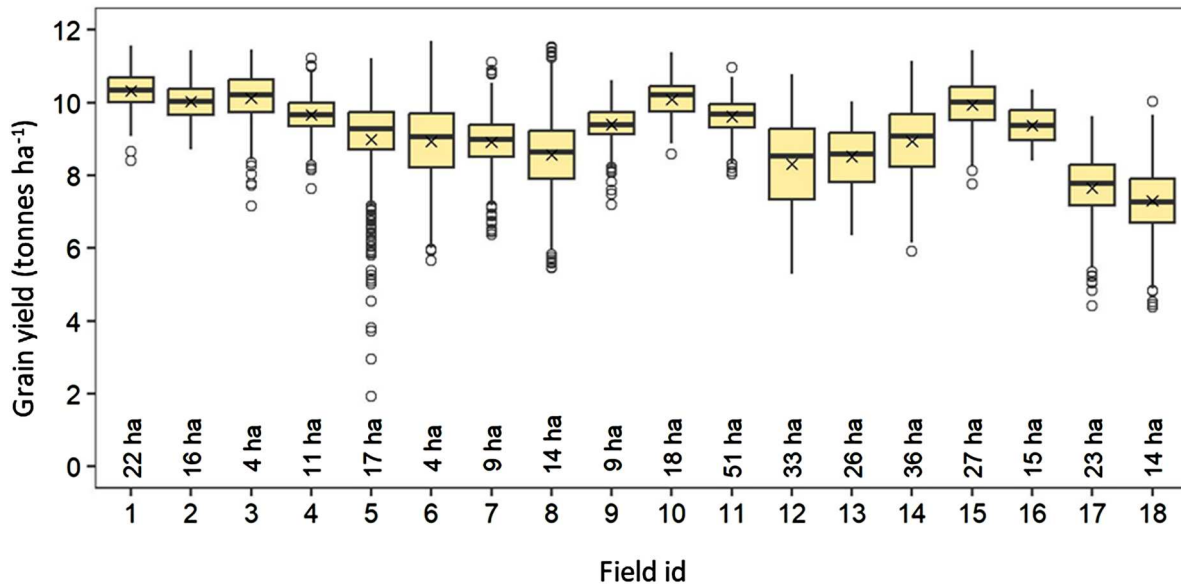
The fields mapped with the combine harvester ranged between 4 and 51 ha in size, with 1,500–16,000 yield records per field. Descriptive statistics on the yield values for the 18 fields are presented in Figure 3. The mean yield of the fields ranged from 7.4 to 10.2 tonnes ha<sup>-1</sup>. There were no very low-yielding fields, but there was some variation in the yield levels in different fields and in the magnitude of within-field variation. Thus the average yield level, measured as the degree of spread, varied from one field to another (as indicated by the midlines, interquartile ranges (boxes) and non-outlier ranges (whiskers) in Figure 3), regardless of field area. Some fields were homogeneous (e.g. field no. 16) and some were heterogeneous (e.g. field no. 5, where the yield was much lower in a sandy part of the field – the average yield was around 6.5 tonnes ha<sup>-1</sup> in that part compared to around 9 tonnes ha<sup>-1</sup> in the rest of the field).

### Overview of model performance

The results of the leave-one-field-out cross-validations showed in general acceptable performance of the model for practical application, but with some variation in prediction accuracy between date-VI combinations and resolution. Mean MAE for the 18 fields ranged between 0.49 and 0.81 tonnes ha<sup>-1</sup> for the 225 date-VI combinations at 20 m resolution, and between 0.40 and 0.75 tonnes ha<sup>-1</sup> at 40 m resolution. The corresponding ranges for mean  $r^2$  were 0.17–0.52 (20 m) and 0.20–0.60 (40 m). Figure 4 shows the ranks for the date-VI combinations at 40 m resolution. The best model was linear regression between NDWI from the first and the third image acquisition date (approximately corresponding to winter wheat booting and milk development).

### Performance of different VI combinations

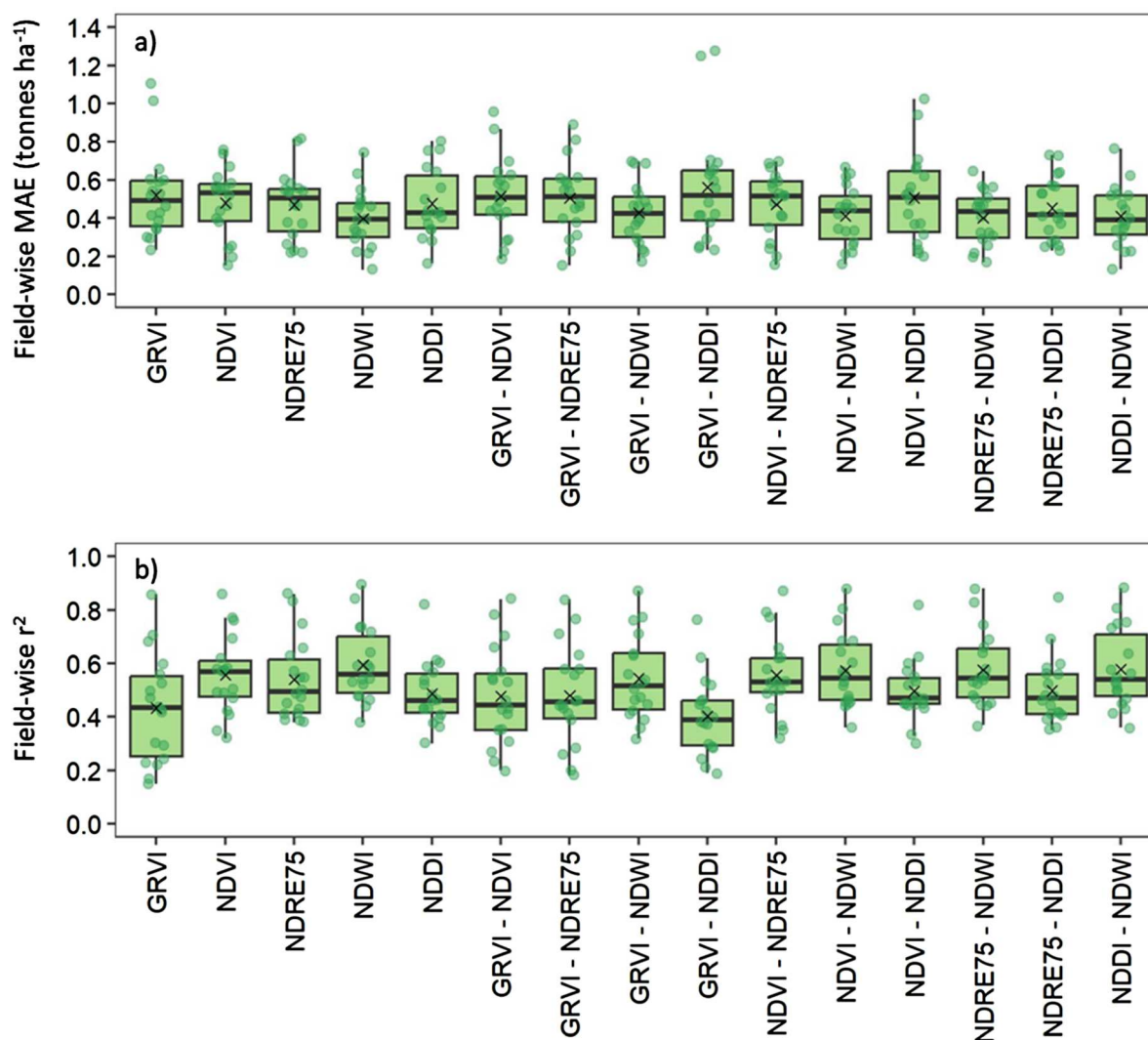
Figure 5 shows boxplots of MAE and  $r^2$  for the 18 fields for the different VI combinations in the best date-VI combination (dates 1 and 3, i.e. 4 and 29 June) and



**Figure 3.** Descriptive statistics on observed yield per field (nos. 1-18). A cross indicates mean, midline indicates median and a box indicates the interquartile range (IQR), where whiskers indicate the highest and lowest observation within  $1.5 \times$  IQR from the midline. Observations outside whiskers are outliers.

Date 1, 2, 3, 4	132	70	74	<b>38</b>	118	114	113	90	100	74	64	67	68	74	82
Date 2, 3, 4	144	112	122	62	141	125	126	65	126	117	60	94	68	111	64
Date 1, 3, 4	129	69	75	38	117	116	110	61	115	68	38	66	40	68	48
Date 1, 2, 4	173	103	99	64	124	148	133	130	134	95	92	88	95	93	122
Date 1, 2, 3	136	74	76	38	113	118	112	95	98	75	71	67	75	73	112
Date 3, 4	144	121	127	62	152	142	144	111	127	120	100	112	105	117	111
Date 2, 4	182	148	147	105	151	179	189	131	179	155	112	153	111	148	109
Date 2, 3	149	111	117	62	137	128	127	65	156	113	59	129	63	134	66
Date 1, 4	177	99	98	62	125	168	174	121	168	104	79	122	74	115	83
Date 1, 3	132	72	75	<b>37</b>	113	117	108	63	142	69	43	98	42	96	48
Date 1, 2	199	127	120	84	113	172	156	143	149	119	113	111	107	105	110
Date 4	192	183	195	133	177	187	192	122	148	184	131	127	125	144	126
Date 3	149	119	121	63	150	123	117	68	108	117	68	79	68	92	64
Date 2	205	169	159	140	154	170	162	140	153	161	143	149	140	146	142
Date 1	202	126	126	84	123	115	123	86	115	124	82	86	85	106	74
	GRVI	NDVI	NDRE75	NDWI	NDDI	GRVI - NDVI	GRVI - NDRE75	GRVI - NDWI	GRVI - NDDI	NDVI - NDRE75	NDVI - NDWI	NDVI - NDDI	NDRE75 - NDWI	NDRE75 - NDDI	NDDI - NDWI

**Figure 4.** Mean of ranks of mean absolute error (MAE) and  $r^2$  for the 225 date-vegetation index (VI) combinations at 40 m data resolution. The predictor set with the smallest (i.e. best) mean rank is marked with a bold red frame. Thin red frames indicate the predictor set for which field-wise MAE and  $r^2$  values at the field level are presented in Figures 5 and 6.



**Figure 5.** Boxplots of (a) field-wise mean absolute error (MAE) and (b) coefficient of determination ( $r^2$ ) from leave-one-field-out cross-validation for models based on all vegetation index (VI) combinations. Best date combination (4 June together with 29 June) and best resolution (40 m) only. A cross indicates mean, midline indicates median and box indicates the interquartile range (IQR), where whiskers indicate the highest and lowest observation within  $1.5 \times \text{IQR}$  from the midline. Jittered points are individual values.

best within-field resolution (40 m). The effect of date on the satellite image for NDWI is shown in Figure 6. The performance of VIs and VI combinations other than NDWI was relatively similar based on MAE, with somewhat lower but comparable  $r^2$  for GRVI. Mean MAE value varied between  $0.40 \text{ tonnes ha}^{-1}$  for NDWI to  $0.48 \text{ tonnes ha}^{-1}$  for NDRE75 and  $0.52 \text{ tonnes ha}^{-1}$  for NDVI, the corresponding  $r^2$  values were 0.59, 0.56 and 0.43. Combining two VIs did not substantially improve the performance compared with using the indices on their own (Figure 5).

Figure 6 presents boxplots of MAE and  $r^2$  for the 18 fields for the different date combinations in the best index (NDWI) and best within-field resolution (40 m). The best performance in predicting yield was obtained for date 1 and then date 3 (DC 49–54 and DC71–75,

respectively). Combining more than one date only improved the models very slightly in terms of mean MAE and  $r^2$ .

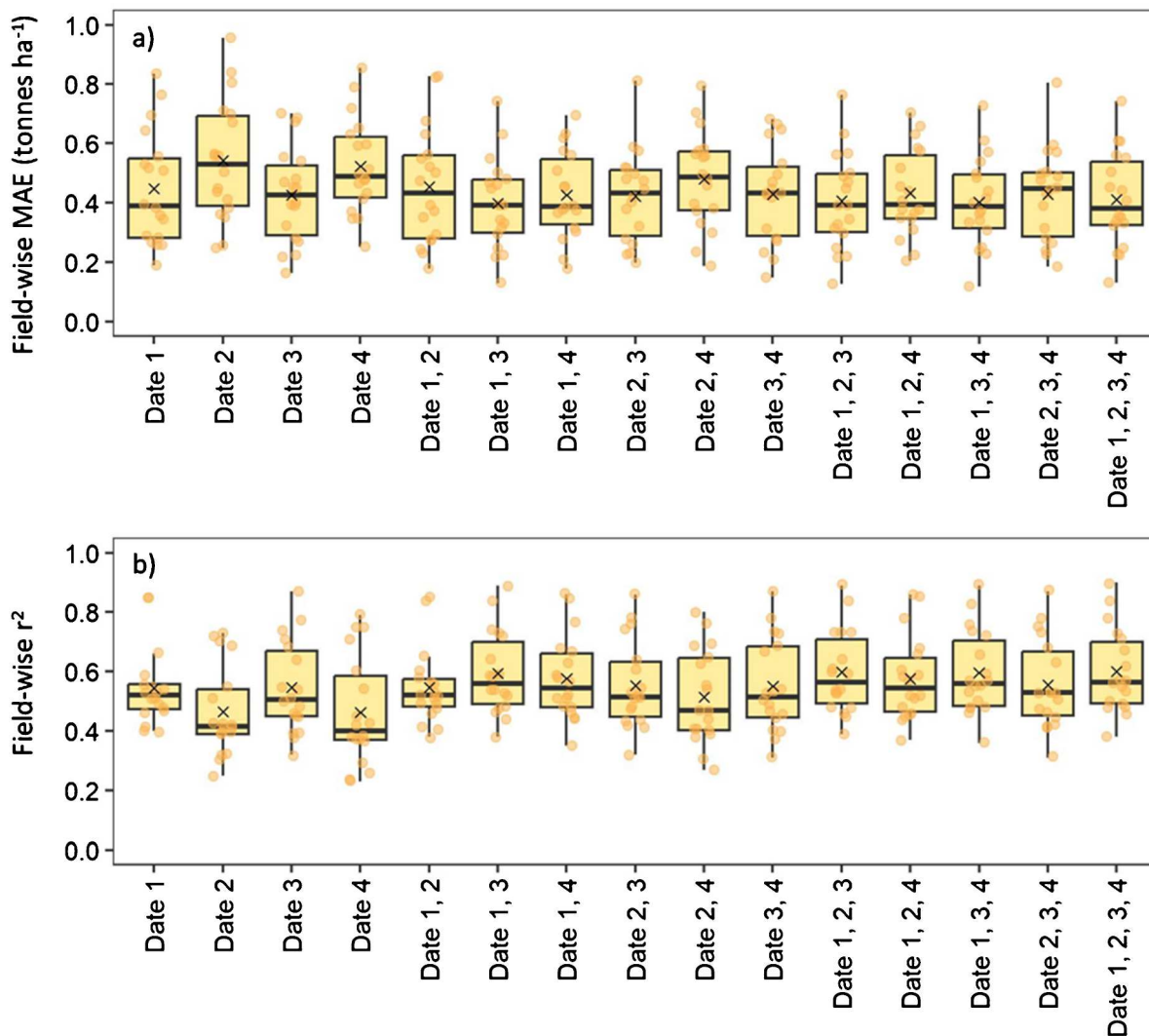
#### Performance at different resolutions

For the best date-VI combination (dates 1 and 3, NDWI), the MAE of field means was in general better than that of pixels, while both MAE and  $r^2$  values were better at 40 m resolution than at 20 m resolution (Figure 7).

#### Spatial patterns of observed and predicted yields

For the best model (based on NDWI from booting and milk maturity, 40 m resolution) observed yields and predicted yields (from the leave-one-field-out cross-





**Figure 6.** Boxplots of (a) field-wise mean absolute error (MAE) and (b) coefficient of determination ( $r^2$ ) from leave-one-field-out cross-validation for models based on all combinations of satellite image dates. Best index and best resolution only (NDWI, 40 m). A cross indicates mean, midline indicates median and a box indicates the interquartile range (IQR), where whiskers indicate the highest and lowest observation within  $1.5 \times$  IQR from the midline. Jittered points are individual values. Date 1: 4 June (Zadoks plant development stage DC 49-54); date 2: 14 June (DC61-65); date 3: 29 June (DC71-75); date 4: 11 July (DC83-85).

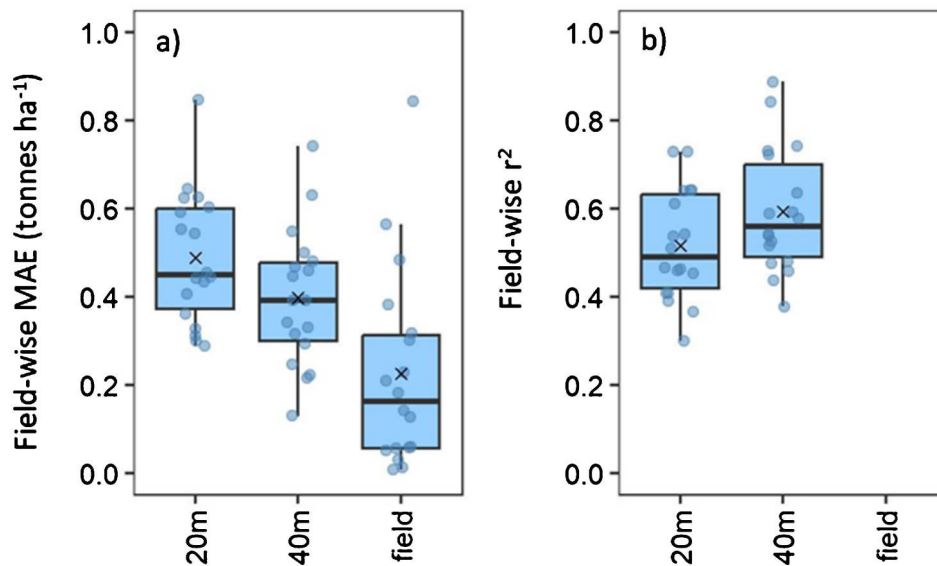
validation) were mapped together so that spatial variation patterns could be visually compared (example in Figure 8).

### Evaluating the best model for the development of management zones

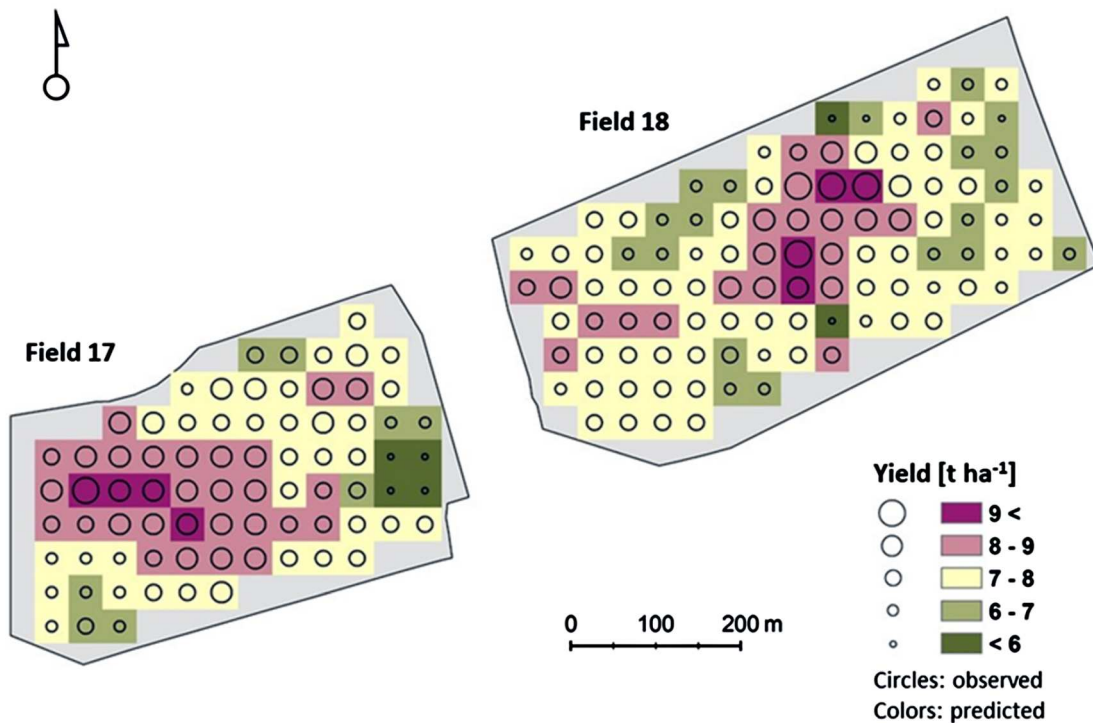
In practice, it is common to use yield maps to delineate management zones. In Figure 9, the predicted yield (aggregated mean value for four zones per field) is plotted against the observed yield for the 18 fields. As can be seen from the plot, the highest predicted yield fell within the zone with the highest observed yield and the lowest predicted yield fell within the zone with the lowest observed yield. However, some fields showed signs of model bias (consistent over-prediction).

### Test on independent farms

The field mean yield prediction model for the best date-VI combination (date 1 and 3, NDWI) was used to predict the mean yield of all 69 farms in the SCB yield survey, representing a larger region. The results showed relatively good agreement (points largely following the 1:1 line; Figure 10). MAE was 0.75 tonnes ha<sup>-1</sup> and  $r^2$  was 0.25. There seems to be a slight so-called conditional bias with overestimation of low values and underestimation of high values, a common effect in empirical modelling. In this case, it means that the satellite-based yield maps do not show the full variation range. However, when interpreting this diagram, one should bear in mind that the observed data are uncertain and it is unknown



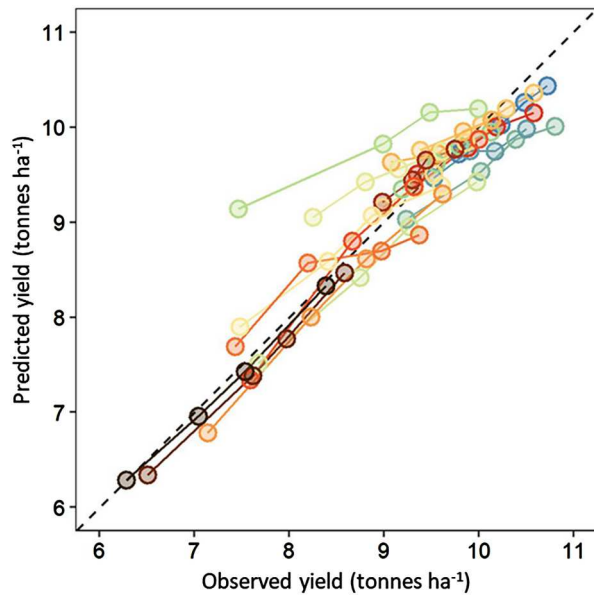
**Figure 7.** Boxplots of (a) field-wise mean absolute error (MAE) and (b) coefficient of determination ( $r^2$ ) from leave-one-field-out cross-validation for models for the three resolutions tested (20, 40 m, field mean). Best index (NDWI) and date combination (4 June together with 29 June) only. A cross indicates mean, midline indicates median and box indicates the interquartile range (IQR), where whiskers indicate the highest and lowest observation within  $1.5 \times$  IQR from the midline. Jittered points are individual values. For whole fields, there is only one observed and one predicted value, so the MAE is just the absolute error.



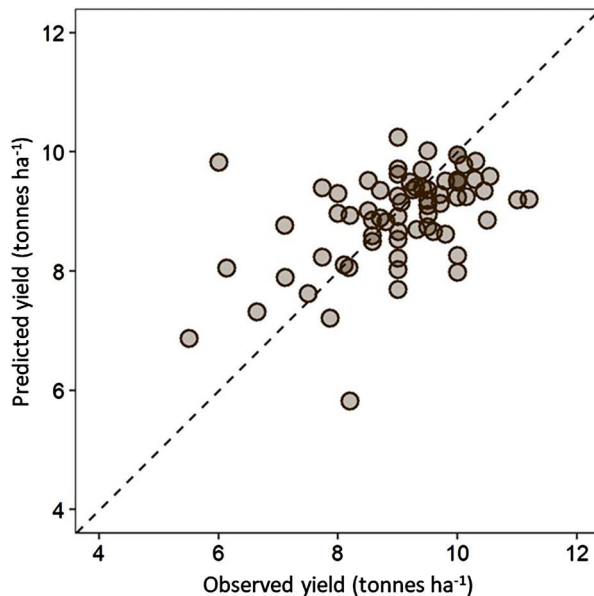
**Figure 8.** Observed and predicted yields for the 40 m grid of two example fields. Circles show combined data and colours show yield predictions from the leave-one-field-out cross-validation. The model was based on the best index (NDWI) and date combination (4 June together with 29 June).

whether the predicted or observed values are more correct. In this case, the ground truth is based on phone interviews with farmers, who have assessed their farm average winter wheat yield in tonnes per

hectare. More information on the data collection and the national statistics that are based on it are given by SCB, Statistics Sweden (2019) and the Swedish Board of Agriculture (2019).



**Figure 9.** Relationship between predicted and observed yield. Zones were generated by splitting observed yield into four zones with equal numbers of observations. Points represent zone means. Lines connect the four zones in the same field. Predictions based on the best index (NDWI), best date combination (4 June and 29 June) and best resolution (40 m).



**Figure 10.** Relationship between predicted and observed yield for 69 farms surveyed by Statistics Sweden (SCB) within a 1055 km<sup>2</sup> area in southern Sweden. Predictions based on the best index (NDWI) and date combination (4 June and 29 June), models developed based on field mean values.

## Discussion

### Model accuracy

This study showed that univariate linear regression models can achieve reasonable accuracy in predicting

winter wheat yield when using one of the following VIs in this study: NDVI, NDRE75, NDWI, and NDDI. The best performance was achieved with NDWI for the date combination 4 June and 29 June (DC 49-54 and DC71-75, respectively) (Figure 4). A plausible explanation for the superior performance of NDWI is that it is calculated using a combination of two bands, the near infrared (NIR) region, that mainly reflects the leaves' structure, and the shortwave infrared (SWIR) band, which reflects changes in both the vegetation structure and water content (see also Gao 1996). To what extent this is common in this type of modelling remains to be tested further. The visual index GRVI showed slightly weaker performance than indices including reflectance bands of longer wavelengths (Figure 4). This is similar to what is indicated in e.g. Bar-meier et al. (2017). In Sentinel-2 based modelling, this is not a limitation, but if one would choose to develop models for drone-based yield mapping, the recommendation based on our results would be to invest in a (more expensive) camera not only registering visible wavelength bands but also bands in the red edge and NIR or NIR and SWIR regions.

No obvious improvement was obtained by using multiple linear regression with a combination of VIs, or by using complex VIs such as NDDI (Figure 5), supporting previous findings by e.g. Ulfa et al. (2022). In a modelling framework, like the one tested here, it is important that models are robust and work well in other fields than those used for model calibration. A less complex model type limits the risk for overfitting, compared to a more flexible model type. The present results show that simple univariate linear regression is good enough for the present purpose.

The models developed in our study, showed the best performance in the prediction of field means, while at the within-field scale, 40 m resolution was better than 20 m resolution (Figure 7). The improvement in the prediction model from within-field scale to mean field scale was much greater than the improvement between resolutions at within-field scales (20, 40 m) (Figure 7a). One possible reason why aggregation improved the results is that errors due to data uncertainty at the finer scale were cancelled out by aggregation. This could be due to positional uncertainty in both combined harvester data and satellite data and/or errors in the recorded values. Similar results have been found in earlier studies at the within-field scale in multi-aggregation of cell size for NDVI maps (e.g. Martínez-Beltrán et al. 2009).

The approach developed in the present study was capable of predicting winter wheat yield with reasonable performance even at advanced development stages (e.g. DC 83-85), but the best performance was

achieved for a combination of two development stages (DC 49-54 and DC 71-75) (Figure 6). Also, previous studies have shown that combining multiple images during a season improves the accuracy of yield prediction (e.g. Hunt et al. 2019). Note that the accuracy of the yield maps obtained will depend on the absolute calibration of the combined harvester yield monitoring equipment since if there is a bias in these values it will propagate to the satellite-based maps.

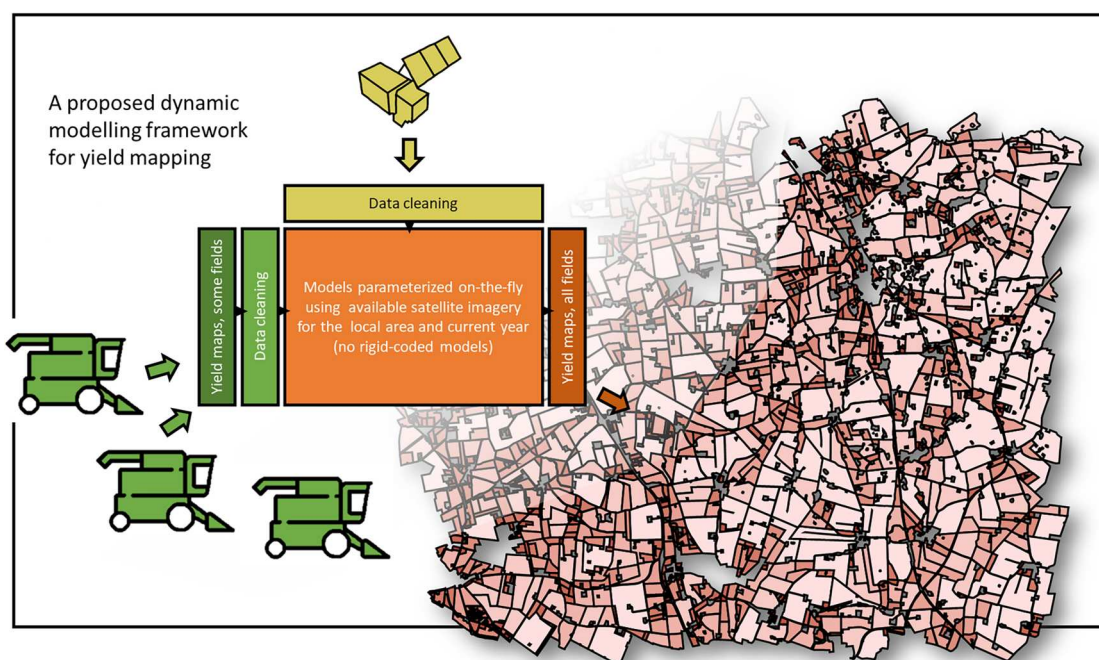
### ***A proposed dynamic modelling framework for yield mapping***

The approach for calibrating VI maps from Sentinel-2 with combined harvest data for yield mapping in neighbouring fields showed promising results (Figures 4, 9, 10). The results suggest that a satellite image-based system where farmers can benefit from collected combined harvest yield data not only in the fields where these data are collected but also in neighbouring fields or in parts of fields with data gaps, could be viable. In order to develop a similar yield mapping strategy in other areas and for other crops, a dynamic modelling framework for yield mapping is proposed in Figure 11. This framework could also allow other farmers in the neighbourhood to issue yield maps for their fields with the same crop in the same satellite image. This framework has similarities to the work presented by Filippi et al. (2019), although that focused on yield prediction during the season, rather than yield mapping after

harvest and for previous years. In addition to providing farmers with useful data, remote-sensing-based yield mapping frameworks at regional or national scale, can support the production of national statistics. This was proposed Brandt et al. (2024), who also tested a scalable machine-learning-based method for this purpose. Their modelling was coarser than the present study (1 km) but also our method could potentially be useful in this way. However, before using our approach at regional and national scales, for any purpose, it needs to be further validated. One benefit of parameterising models on-the-fly and picking the best model for the current year and local neighbourhood based on cross-validation, instead of relying on rigid-coded models, is that the system overcomes the problems caused by satellite imagery not being available for the same dates or developmental stages in different areas and years. Another benefit of the presently proposed framework is that potential differences in VI values between images are not translated into yield differences when each image is converted individually into yield maps.

### ***Potential use of maps for delineation of management zones and yield stability maps***

It was demonstrated in the present one-year study that satellite-based yield maps derived from the current method can be used to delineate management zones with different yield levels. General biases were observed in some fields, possibly because of differences in cultivar



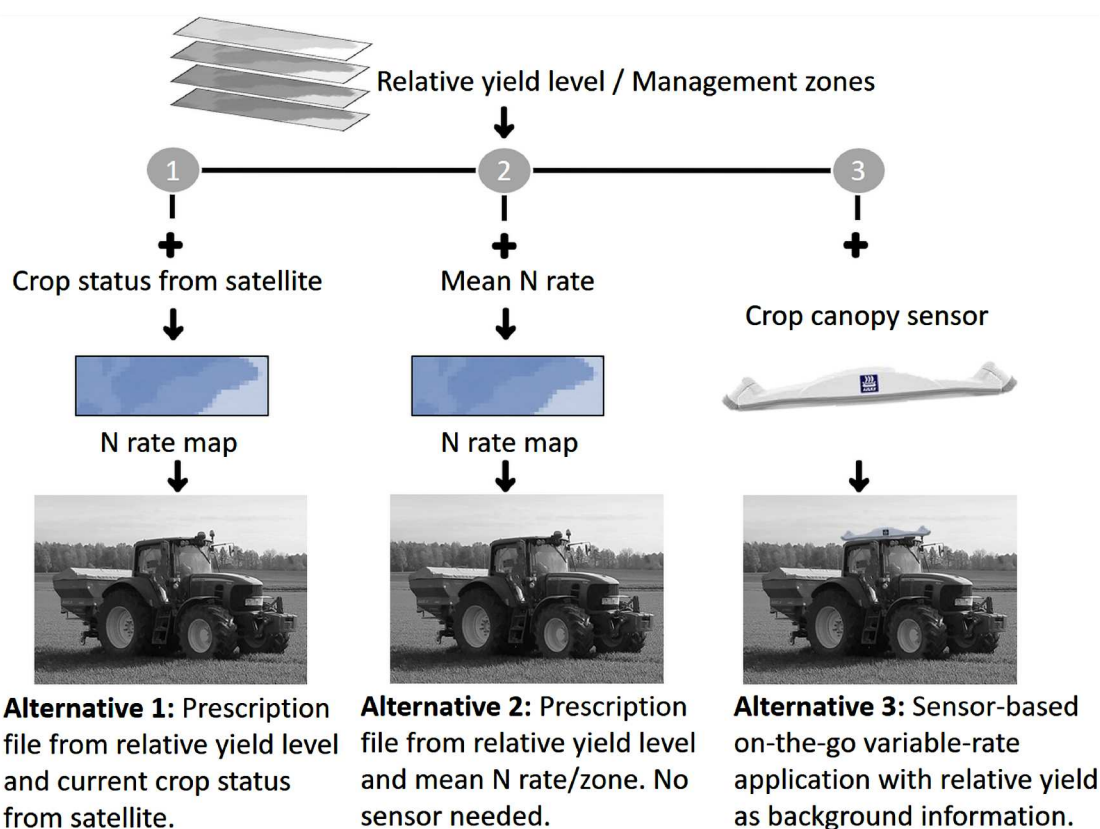
**Figure 11.** Proposed dynamic modelling framework for yield mapping.

and sowing date between the 18 fields (Figure 9). If the true mean yield of the fields is known, one could potentially compensate for field-specific biases and obtain a more accurate yield map. A method for this has been proposed by Söderström et al. (2021).

Satellite-based yield mapping offers possibilities to develop yield stability maps based on several years, which enable better management of natural resources through e.g. precision nutrient management (Basso and Antle 2020). By applying the approach developed in this study, historical yield data from combined harvesters and statistical reports can be used to generate time series of yield maps (space-time data cubes of yield), not only in fields for which combined harvester data or field means are available, but also in neighbouring fields lacking such data. This can be of great value for farmers, advisors and local planners. Such time series of yield maps can be combined into yield stability maps, relative yield level maps or management zones, which e.g. can be used to refine supplementary nitrogen (N) fertilisation in split application cropping systems (Figure 12). In-season adaptation of N is generally considered efficient in terms of N use efficiency, and reaching a target yield quality and quantity (e.g. Raun and

Johnson 1999; European Parliament 2016; Piikki et al. 2022). This depends on difficulties in foreseeing the weather and the amount of soil N mineralisation early in the season. In most cases, the expected yield is a crucial variable in determining the amount of N to apply, and this is something that normally is determined from field to field based on experience (historical data).

Three alternative approaches to take advantage of the relative yield level maps in this context are shown in Figure 12. As ‘Alternative 1’ the N rate is determined from satellite imagery during the growing season. There are a number of agricultural decision support systems available that provide this service (e.g. Söderström et al. 2017), but often the expected yield is a guessed field average. The relative yield level map generated here can distribute this guessed field average into a more reasonable value based on historical yield outcomes. Similarly, if advanced tractor-mounted crop sensors (Raun et al. 2001; Reusch 2003) are used (‘Alternative 3’), the relative yield map could function as a background map that adjusts the prescribed N rate on-the-go. Commonly, an average expected yield is also used in that case. Management zones of different yield levels can also be used directly without



**Figure 12.** Three alternative approaches are to use a generated relative yield level map as described in this study for supplementary nitrogen (N) fertilisation. Alternative 1: in a satellite image-based decision support system; Alternative 2: manual use; and Alternative 3: as background information in a tractor-sensor system.

the requirement of advanced equipment – a paper map and manual adjustment can be sufficient provided that the driver knows the vehicle position reasonably well ('Alternative 2'). This alternative could also be developed further, with on-farm trials such as zero or max N plots in each yield level zone to guide the in-season N rate determination (Johnson and Raun 2003).

## Conclusions

A simple but useful approach for mapping winter wheat yield in fields lacking measured yield data, based on combined harvester data from other fields in the neighbourhood combined with VI maps from Sentinel-2, was evaluated. The results showed that simple linear models based on NDVI, NDRE75, GRVI and NDWI can give relatively good prediction accuracy. Using VI combinations or a more complex index (NDDI) did not improve the modelling performance. However, performance was improved by lowering the resolution and by using a combination of satellite images from relatively early and late wheat development stages. The best model obtained was based on NDWI at two crop development stages, booting and milk maturity. The MAE for this model was 0.40 tonnes ha<sup>-1</sup> in leave-one-field-out cross-validation at 40 m map resolution. For the field mean predictions MAE was much lower (0.23 tonnes ha<sup>-1</sup>), while application of the field mean model to predict the farm mean yield in a larger area had MAE of 0.75 tonnes ha<sup>-1</sup>, indicating promise for precision agriculture and also other applications. The approach of individually parameterising models for the current year and local neighbourhood can provide farmers, advisors and authorities with spatial yield data (which are not always available today), and offers new opportunities for improving precision agriculture practices and natural resource management.

## Acknowledgements

Thanks to the farmers surveyed for making their combined yield data available, and to Henrik Stadig who organised the data sharing.

## Disclosure statement

No potential conflict of interest was reported by the authors.

## Funding

This work was supported by the Formas – a Swedish Research Council for Sustainable Development through the national research programme for food under Grant 2019-02280; and Västra Götalandsregionen together with the Swedish

University of Agricultural Sciences under Grant RUN2021-00020.Svenska Forskningsrådet Formas

## Data availability statement

Data used in this study are not openly available, for inquiry please contact the authors.

## Author contribution

The authors confirm their contribution to the paper as follows: All authors contributed to the study conception and design, data preparation, analysis and interpretation of results, as well as manuscript preparation. All authors reviewed the results and approved the final version of the manuscript.

## Ethical approval

No ethical approval was required for the present study, as it did not involve human/animal subjects and did not include case reports.

## Notes on contributors

**Omran Alshihabi**, researcher/lecturer at the Swedish University of Agricultural Sciences (SLU), focuses on water management in agriculture and developing digital applications in precision agriculture. Holds a PhD in civil engineering from Polytech-Lille (France).

**Kristin Persson**, holds a position as a Senior Lecturer with (SLU), where she conducts research in precision agriculture and pedometrics. She holds a PhD in Environmental Sciences with specialisation in plant physiology (University of Gothenburg).

**Mats Söderström**, Associate Professor in Soil Science at SLU, focuses on applied research in digital soil mapping and precision agriculture often involving proximal and remote sensing of soil and crops.

## References

- Adamsen FJ, Pinter PJ, Barnes EM, LaMorte RL, Wall GW, Leavitt SW, Kimball BA. 1999. Measuring wheat senescence with a digital camera. *Crop Sci.* 39:719–724. doi:10.2135/cropsci1999.0011183X003900030019x.
- Alshihabi O, Persson K, Söderström M. 2023. Post-processing yield maps of winter wheat using data from satellites and combines. In: Stafford J.V., editor. *Precision agriculture '23*. Wageningen, The Netherlands: Wageningen Academic Publishers; p. 523–530. doi:10.3920/978-90-8686-947-3\_66.
- Ashcroft PM, Catt JA, Curran PJ, Munden J, Webster R. 1990. The relation between reflected radiation and yield on the broadbalk winter wheat experiment. *Remote Sens.* 11 (10):1821–1836.
- Barmeier G, Hofer K, Schmidhalter U. 2017. Mid-season prediction of grain yield and protein content of spring barley

- cultivars using high-throughput spectral sensing. *Eur J Agron.* 90:108–116. doi:10.1016/j.eja.2017.07.005.
- Basso B, Antle J. 2020. Digital agriculture to design sustainable agricultural systems. *Nat Sustain.* 3:254–256. doi:10.1038/s41893-020-0510-0.
- Birrell SJ, Sudduth KA, Borgelt SC. 1996. Comparison of sensors and techniques for crop yield mapping. *Comput Electron Agric.* 14(2-3):215–233. doi:10.1016/0168-1699(95)00049-6.
- Blackmore S. 2000. The interpretation of trends from multiple yield maps. *Comput Electron Agric.* 26:37–51. doi:10.1016/S0168-1699(99)00075-7.
- Blackmore S, Moore M. 1999. Remedial correction of yield map data. *Precision Agriculture.* 1:53–66. doi:10.1023/A:1009969601387.
- Blasch G, Li Z, Taylor JA. 2020. Multi-temporal yield pattern analysis method for deriving yield zones in crop production systems. *Precis Agric.* 21(6):1263–1290. doi:10.1007/s11119-020-09719-1.
- Bouras EH, Olsson PO, Thapa S, Díaz JM, Albertsson J, Eklundh L. 2023. Wheat yield estimation at high spatial resolution through the assimilation of sentinel-2 data into a crop growth model. *Remote Sens.* 15(18):4425. doi:10.3390/rs15184425.
- Brandt P, Beyer F, Borrmann P, Möller M, Gerighausen H. 2024. Ensemble learning-based crop yield estimation: a scalable approach for supporting agricultural statistics. *Gisci Remote Sens.* 61(1):2367808. doi:10.1080/15481603.2024.2367808.
- Curran P. 1980. Multispectral remote sensing of vegetation amount. *Prog Phys Geography: Earth and Environ.* 4(3):315–341. doi:10.1177/030913338000400301.
- Delin S. 2005. Site-specific nitrogen fertilization demand in relation to plant available soil nitrogen and water [Acta Universitatis Agriculturae Sueciae, doctoral thesis. 2005:6. <https://publications.slu.se/?file=publ/show&id=12509>
- ESA. 2024. Multispectral Instrument (MSI) Overview. <https://sentinels.copernicus.eu/en/web/sentinel/technical-guides/sentinel-2-msi/msi-instrument> (verified 2024-09-05).
- European Parliament. 2016. Precision agriculture and the future of farming in Europe: scientific foresight study. European Parliament. <https://data.europa.eu/doi/10.2861020809> (verified 2024-09-18).
- Filippi P, Jones EJ, Wimalathunge NS, Somarathna PDSN, Pozza LE, Ugbaje SU, Jephcott TG, Paterson SE, Whelan BM, Bishop TF. 2019. An approach to forecast grain crop yield using multi-layered, multi-farm data sets and machine learning. *Precision Agriculture.* 20:1015–1029. doi:10.1007/s11119-018-09628-4.
- Franz TE, Pokal S, Gibson JP, Zhou Y, Gholizadeh H, Tenorio FA, Wardlow B. 2020. The role of topography, soil, and remotely sensed vegetation condition towards predicting crop yield. *Field Crops Res.* 252:107788. doi:10.1016/j.fcr.2020.107788.
- Gao BC. 1996. NDWI—a normalized difference water index for remote sensing of vegetation liquid water from space. *Remote Sens Environ.* 58(3):257–266. doi:10.1016/S0034-4257(96)00067-3.
- Gaso DV, Berger AG, Ciganda VS. 2019. Predicting wheat grain yield and spatial variability at field scale using a simple regression or a crop model in conjunction with landsat images. *Comput Electron Agric.* 159:75–83. doi:10.1016/j.compag.2019.02.026.
- Gu Y, Brown JF, Verdin JP, Wardlow B. 2007. A five-year analysis of MODIS NDVI and NDWI for grassland drought assessment over the central Great Plains of the United States. *Geophys Res Lett.* 34(6):L06407. doi:10.1029/2006GL029127.
- Hunt ML, Blackburn GA, Carrasco L, Redhead JW, Rowland CS. 2019. High resolution wheat yield mapping using Sentinel-2. *Remote Sens Environ.* 233:111410. doi:10.1016/j.rse.2019.111410.
- Johnson GV, Raun WR. 2003. Nitrogen response index as a guide to fertilizer management. *J Plant Nutr.* 26(2):249–262. doi:10.1081/PLN-120017134.
- Lachia N, Pichon L, Marcq P, Taylor J, Tisseyre B. 2021. Why are yield sensors seldom used by farmers—a French case study. In: Stafford J.V., editor. *Precision agriculture '21*. Wageningen, The Netherlands: Wageningen Academic Publishers; p. 745–751. doi:10.3920/978-90-8686-916-9\_89.
- Liao C, Wang J, Shan B, Song Y, He Y, Dong T. 2023. Near real-time yield forecasting of winter wheat using sentinel-2 imagery at the early stages. *Precision Agriculture.* 24(3):807–829. doi:10.1007/s11119-022-09975-3.
- Luo L, Sun S, Xue J, Gao Z, Zhao J, Yin Y, Luan X. 2023. Crop yield estimation based on assimilation of crop models and remote sensing data: A systematic evaluation. *Agric Sys.* 210:103711. doi:10.1016/j.agry.2023.103711.
- Lyle G, Clarke K, Kilpatrick A, Summers DM, Ostendorf B. 2023. A spatial and temporal evaluation of broad-scale yield predictions created from yield mapping technology and landsat satellite imagery in the Australian Mediterranean dryland cropping region. *ISPRS Int J Geoinf.* 12(2):50. doi:10.3390/ijgi12020050.
- Maestrini B, Basso B. 2018. Predicting spatial patterns of within-field crop yield variability. *Field Crops Res.* 219:106–112. doi:10.1016/j.fcr.2018.01.028.
- Maestrini B, Basso B. 2021. Subfield crop yields and temporal stability in thousands of US Midwest fields. *Precision Agriculture.* 22:1749–1767. doi:10.1007/s11119-021-09810-1.
- Martínez-Beltrán C, Osann Jochum MA, Calera A, Meliá J. 2009. Multisensor comparison of NDVI for a semi-arid environment in Spain. *Int J Remote Sens.* 30(5):1355–1384. doi:10.1080/01431160802509025.
- Nissen K, Söderström M. 1999. Mapping in precision farming—from the farmer's perspective. In: Stafford J.V., editor. *Precision agriculture '99*. Sheffield, UK: Sheffield Academic Press; p. 655–664.
- Panek E, Gozdowski D, Stępień M, Samborski S, Ruciński D, Buszke B. 2020. Within-Field relationships between satellite-derived vegetation indices, grain yield and spike number of winter wheat and triticale. *Agronomy.* 10(11):1842. doi:10.3390/agronomy10111842.
- Perich G, Turkoglu MO, Graf LV, Wegner JD, Aasen H, Walter A, Liebisch F. 2023. Pixel-based yield mapping and prediction from Sentinel-2 using spectral indices and neural networks. *Field Crops Res.* 292:108824. doi:10.1016/j.fcr.2023.108824.
- Piikki K, Söderström M, Stadig H. 2022. Remote sensing and on-farm experiments for determining in-season nitrogen rates in winter wheat – options for implementation, model accuracy and remaining challenges. *Field Crops Res.* 289:108742. doi:10.1016/j.fcr.2022.108742.
- Ping JL, Dobermann A. 2005. Processing of yield map data. *Precis Agric.* 6:193–212. doi:10.1007/s11119-005-1035-2.

- Raun WR, Johnson GV. 1999. Improving nitrogen use efficiency for cereal production. *Agron J.* 91(3):357–363. doi:10.2134/agronj1999.00021962009100030001x.
- Raun WR, Solie JB, Johnson GV, Stone ML, Lukina EV, Thomason WE, Schepers JS. 2001. In-season prediction of potential grain yield in winter wheat using canopy reflectance. *Agron J.* 93(1):131–138. doi:10.2134/agronj2001.931131x.
- R Core Team. 2023. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. URL <https://www.R-project.org/>.
- Reusch S. 2003. Optimisation of oblique-view remote measurement of crop N-uptake under changing irradiance conditions. *Precision agriculture: Papers from the 4th European Conference on Precision Agriculture: Wageningen Academic Publishers.* 573–578.
- Reyns P, Missotten B, Ramon H, De Baerdemaeker J. 2002. A review of combine sensors for precision farming. *Precis Agric.* 3:169–182. doi:10.1023/A:1013823603735.
- Rouse JW, Haas RH, Schell JA, Deering DW. 1974. Monitoring vegetation systems in the Great Plains with Erts. NASA Special Publication. 351:309.
- Sadeh Y, Zhu X, Dunkerley D, Walker JP, Chen Y, Chenu K. 2024. Versatile crop yield estimator. *Agron Sustain Develop.* 44(4):1–22. doi:10.1007/s13593-024-00974-4.
- SCB, Statistics Sweden. 2019. Kvalitetsdeklaration Skörd av spannmål, trindsäd och oljeväxter. [https://www.scb.se/contentassets/462f77b26e0d4febbaab8b204d960f5/jo0601\\_kd\\_2019\\_gj\\_20191213.pdf](https://www.scb.se/contentassets/462f77b26e0d4febbaab8b204d960f5/jo0601_kd_2019_gj_20191213.pdf) (verified 2024-07-17; in Swedish).
- Schnug E, Murphy D, Evans E, Haneklaus S, Lamp J. 1993. Yield mapping and application of yield maps to computer-aided local resource management. In: Robert P. C., Rust R. H., Larson W.E., editors. *Proceedings of soil specific crop management: A workshop on research and development issues.* WI, USA: American Society of Agronomy, Crop Science Society of America, Soil Science Society of America; p. 87–93. doi:10.2134/1993.soilspecifccrop.c7.
- Schueller JK, Bae YH. 1987. Spatially attributed automatic combine data acquisition. *Comput Electron Agric.* 2(2):119–127. doi:10.1016/0168-1699(87)90022-6.
- Shuai G, Basso B. 2022. Subfield maize yield prediction improves when in-season crop water deficit is included in remote sensing imagery-based models. *Remote Sens Environ.* 272:112938. doi:10.1016/j.rse.2022.112938.
- Simbahan GC, Dobermann A, Ping JL. 2004. Screening yield monitor data improves grain yield maps. *Agron J.* 96(4):1091–1102. doi:10.2134/agronj2004.1091.
- Sims DA, Gamon JA. 2002. Relationships between leaf pigment content and spectral reflectance across a wide range of species, leaf structures and developmental stages. *Remote Sens Environ.* 81(2):337–354. doi:10.1016/S0034-4257(02)00010-X.
- Söderström M, Piikki K, Stadig H. 2021. Yield maps for everyone - scaling drone models for satellite-based decision support. In: Stafford J.V., editor. *Precision agriculture '21.* Wageningen, The Netherlands: Wageningen Academic Publishers; p. 911–918. doi:10.3920/978-90-8686-916-9\_109.
- Söderström M, Piikki K, Stenberg M, Stadig H, Martinsson J. 2017. Producing nitrogen (N) uptake maps in winter wheat by combining proximal crop measurements with sentinel-2 and DMC satellite images in a decision support system for farmers. *Acta Agriculturae Scandinavica, Section B — Soil & Plant Science.* 67(7):637–650. doi:10.1080/09064710.2017.1324044.
- Stafford JV, Ambler B, Lark RM, Catt J. 1996. Mapping and interpreting the yield variation in cereal crops. *Comput Electron Agric.* 14(2-3):101–119. doi:10.1016/0168-1699(95)00042-9.
- Stettmer M, Mittermayer M, Maidl F-X, Schwarzensteiner J, Hülsbergen K-J, Bernhardt H. 2022. Three methods of site-specific yield mapping as a data source for the delineation of management zones in winter wheat. *Agriculture.* 12:1128. doi:10.3390/agriculture12081128.
- Sudduth K, Drummond ST. 2007. Yield editor: software for removing errors from crop yield maps. *Agron J.* 99(6):1471–1482. doi:10.2134/agronj2006.0326.
- Swedish Board of Agriculture. 2019. Standard yields for yield survey districts, counties and the whole country in 20.019. <https://jordbruksverket.se/download/18.514d3694172cce07237d52f6/1624262130865/JO15SM1901.pdf> (verified 2024-07-17; in Swedish).
- Swedish Environmental Protection Agency. 2023. Nationella Marktäckdata (land cover of Sweden). Available online: <https://www.naturvardsverket.se/verktug-och-tjanster/kartor-och-karttjanster/nationella-marktackedata/ladda-ner-nationella-marktackedata/>.
- Thylén L, Algerbo PA, Giebel A. 2000. An expert filter removing erroneous yield data. In: Robert P. C., Rust R. H., Larson W.E., editor. *Proceedings of the 5th international conference on precision agriculture.* Madison, WI, USA: American Society of Agronomy; p. 1–9.
- Ulf F, Orton TG, Dang YP, Menzies NW. 2022. Developing and testing remote-sensing indices to represent within-field variation of wheat yields: assessment of the variation explained by simple models. *Agronomy.* 12:384. doi:10.3390/agronomy12020384.
- United Nations. 2015. Transforming our World: The 2030 Agenda for Sustainable Development. A/RES/70/1. <https://sdgs.un.org/sites/default/files/publications/21252030%20Agenda%20for%20Sustainable%20Development%20web.pdf> (verified 2024-09-05).
- Vallentin C, Harfenmeister K, Itzerott S, Kleinschmit B, Conrad C, Spengler D. 2022. Suitability of satellite remote sensing data for yield estimation in northeast Germany. *Precision Agriculture.* 23:52–82. doi:10.1007/s11119-021-09827-6.
- Zadoks JC, Chang TT, Konzak CF. 1974. A decimal code for the growth stages of cereals. *Weed Res.* 14(6):415–421. doi:10.1111/j.1365-3180.1974.tb01084.x.
- Zhang T, Su J, Liu C, Chen W, Liu H, Liu G. 2017. Band selection in sentinel-2 satellite for agriculture applications. *Proceedings of 2017 23rd International Conference on Automation and Computing (ICAC), September 2017.* 1–6. doi:10.23919/ICAC.2017.8081990.