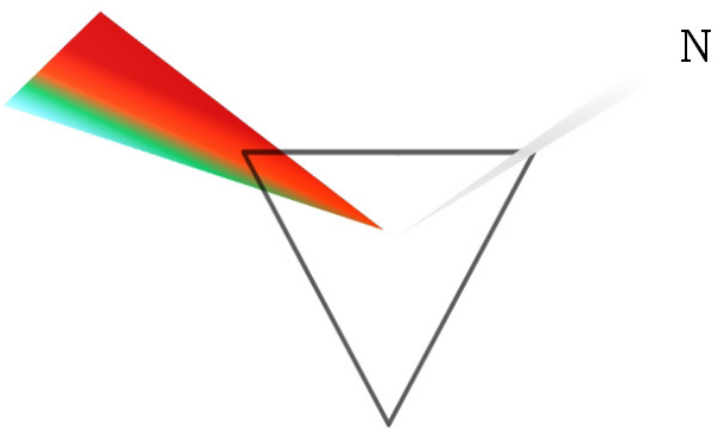
Acta Universitatis Agriculturae Sueciae

Doctoral Thesis No. 2022:59
Faculty of Natural Resources and Agricultural Sciences

Towards synthesis for nitrogen fertilisation using a decision support system

Sandra Wolters



Doctoral Thesis No. 2022:59 • Towards synthesis for nitrogen fertilisation using a decision support system • Sandra Wolters

Towards synthesis for nitrogen fertilisation using a decision support system

Sandra Wolters
Faculty of Natural Resources and Agricultural Sciences
Department of Soil and Environment
Skara

SLU
SWEDISH UNIVERSITY OF AGRICULTURAL SCIENCES

DOCTORAL THESIS
Skara 2022

Acta Universitatis Agriculturae Sueciae
2022:59

Cover: From (multi-)spectral data to N fertilisation. Drawing by S. Wolters.

ISSN 1652-6880
ISBN (print version) 978-91-7760-993-3
ISBN (electronic version) 978-91-7760-994-0

Skara
Print: SLU Service/Repro, Uppsala 2022

Towards synthesis for nitrogen fertilisation using a decision support system

Abstract

Nitrogen (N) fertilisation in crops can be made more efficient by moving from uniform application to meeting variable crop requirements within fields. Within field variable rate N fertilisation of winter wheat (*Triticum aestivum* L.) is practically feasible using information from web-based decision support systems (DSS). Data from different source platforms, such as satellite, unmanned aerial vehicle (UAV) or weather stations can be used for fertilisation planning. System output offers information that can be used to instruct variable rate fertilizer spreaders to increase or decrease fertilizer application rate on-the-go. In Sweden, satellite-based variable rate N fertilisation was available for winter wheat via a DSS, however, the existing module could be improved in different ways. In this thesis work, a new N-uptake model was estimated and opportunities using UAV-based modelling of grain quality were tested. Transferability of UAV-based models to a satellite data scale improved understanding of the complexity of data transfer from UAV-scale to a satellite scale for use in a DSS. Furthermore, it was possible to model crop phenology from historical data, which can improve accuracy of current implemented models, by taking timing of field operations in to account.

Keywords: phenology, precision agriculture, remote sensing, Sentinel-2, unmanned aerial vehicle

Author's address: Sandra Wolters, Swedish University of Agricultural Sciences, Department of Soil and Environment, Skara, Sweden

Abstract

Gödsling av kväve (N) kan göras mer effektiv om man övergår från jämna gödslingsgivor inom fält. Varierad tilldelning av N till höstvete (*Triticum aestivum* L.) inom fält är praktiskt möjligt genom att man använder information från webbaserade beslutsstödsystem (DSS). I sådana system kan data från olika källor, som till exempel satelliter, obemannade flygfarkoster (drönare; UAV) eller väderdata, användas som hjälpmedel för bedömning av tilldelningsbehovet. Informationen från ett DSS kan utgöras av tilldelningsfiler (VRA-filer), vilka kan innehålla instruktioner till gödningsspridare så att utmatningen ökas eller minskas allteftersom utrustningens position ändras på fältet. Utgångspunkten i föreliggande arbete har varit ett satellitdatabaserat DSS som används i Sverige där man bland annat kan generera VRA-filer för komplettering av N-behov i höstvete. Målsättningen var att utveckla olika empiriska modeller som kan användas för att förbättra N-behovsbedömning i ett sådant system. En ny modell för uppskattning av N-upptag i höstvete från satellitdata togs fram. Även en modell för beräkning av proteinhalten i höstvete i den kommande skörden med hjälp av spektrala data insamlade med UAV togs fram. I det arbetet testades möjligheten att överföra en UAV-data-modell till satellitdata, vilket ledde till en ökad förståelse för komplexiteten i att överföra modeller mellan sensorplattformar när man avser att skala upp modeller för användning i ett DSS. Dessutom undersöktes möjligheten att koppla samman historiska väderdata med ett omfattande dataset bestående av observationer av utvecklingsstadier i höstvete, för att ta fram en modell för skattning av höstvetets fenologiska utveckling under säsongen. En sådan modell skulle kunna förbättra tillämpningen av andra modeller i ett DSS.

Nyckelord: drönare, fenologi, precisonsodling, fjärranalys, Sentinel-2

Author’s address: Sandra Wolters, Swedish University of Agricultural Sciences, Department of Soil and Environment, Skara, Sweden

Contents

List of publications 9

List of tables 11

List of figures 13

Abbreviations 17

1. Introduction 19
   1.1 Aim and objectives 20

2. Background 23
   2.1 Precision agriculture 23
      2.1.1 Spatial variation 23
      2.1.2 Challenges for precision agriculture 24
   2.2 The importance of decision support systems 24
   2.3 Fertilisation of winter wheat 25
      2.3.1 Winter wheat 26
      2.3.2 Crop chlorophyll content and N-uptake 26
      2.3.3 Determination of grain seed quality 26
      2.3.4 Phenology and timing of fertilisation 27
   2.4 Sensing instruments, data and platforms 29
      2.4.1 Reflectance 29
      2.4.2 Composites and vegetation indices 31
      2.4.3 Active and passive sensing 32
      2.4.4 Sensor mounts 32
      2.4.5 Sensor geometry and view angle 35
   2.5 Options for collecting field information 36
      2.5.1 Field sampling 36
      2.5.2 Field trials 36
   2.6 Multi-source data 36

3. Materials and methods 39
3.1 Workflow 39
3.2 Study area 41
3.3 Remote sensing 41
3.3.1 Sentinel-2 41
3.3.2 Unmanned aerial vehicle 43
3.4 Mesan grid 44
3.5 Field observations 45
3.6 Vegetation indices 47
3.7 Statistics 48
3.8 Software 49
3.9 Data and method summary 50

4. Results 51
4.1 Cultivar specific models (Paper I, II & III) 51
4.2 Performance by growing stage (Paper I & III) 52
4.3 UAV-based CP predictions (Paper II) 53
4.4 Sentinel-2 data processing levels (Paper I) 55
4.5 Upscaling a prediction model (Paper I) 57
4.6 Multi-source data (Paper I & II) 59
4.6.1 Index model VS. vehicle measured N-uptake (Paper I) 59
4.6.2 UAV model transfer (Paper II) 60
4.7 Predicting crop phenology (Paper III) 60
4.7.1 Modelling error of growth stage classes (Paper III) 64
4.8 Implementation 65

5. Discussion 67
5.1 Progress in the N fertilisation complex 67
5.2 Calibration for N-uptake 68
5.3 Field trials for CP prediction 69
5.4 Transfer from UAV- to satellite data scale 71
5.5 Approaching phenology 71
5.6 Research for direct DSS implementation 72
5.7 Using fast vs. elaborate algorithms 73

6. Conclusions and future outlook 75

References 77

Popular science summary 87

Populärvetenskaplig sammanfattning 89

Acknowledgements 91

Appendix 93

List of publications

This thesis is based on the work contained in the following Papers, referred to by Roman numerals in the text:

I. Wolters, S., Söderström, M., Piikki, K., Reese, H., Stenberg, M. (2021). Upscaling proximal sensor N-uptake predictions in winter wheat (*Triticum aestivum* L.) with Sentinel-2 satellite data for use in a decision support system. *Precision agriculture*, 1263-1283. doi: 10.1007/s11119-020-09783-7

II. Wolters, S., Söderström, M., Piikki, K., Börjesson, T., Pettersson, C. (2022). Predicting grain protein concentration in winter wheat (*Triticum aestivum* L.) based on unpiloted aerial vehicle multispectral optical remote sensing. *Acta agriculturae scandinavica*, *Section B: Soil and plant Science*, 72, 788-802. doi:10.1080/09064710.2022.2085165

III. Wolters, S., Marzec-Schmidt, K., Börjesson, T., Söderström, M., Piikki, K. (2022). Crop phenology forecasting for decision support in precision agriculture: A procedure towards practical implementation (Manuscript).

The contribution of Sandra Wolters to the Papers included in this thesis was as follows:

I. Conceptualisation and method: Söderström, M., Wolters, S., Piikki, K.; Data collection and preparation: Stenberg, M., Wolters, S.; Data analysis and visualisation: Wolters, S.; First draft preparation: Wolters, S.; Review and editing: Wolters, S., Söderström, M., Piikki, K., Reese, H. and Stenberg, M.

II. Conceptualization and method: Söderström, M., Piikki, K., Wolters, S.; Data collection Söderström, M.; Data collection and preparation: Söderström, M., Wolters, S.; Data analysis and visualisation: Wolters, S.; First draft preparation: Wolters, S., Piikki, K., Söderström, M.; Review and editing: Wolters, S., Piikki, K,. Söderström, M., Börjesson, T., Petterson, C.

III. Conceptualization and method: Börjesson, T., Piikki, K., Wolters, S., Marzec-Schmidt, K.; Data collection Arvidsson, A. (Swedish board of agriculture), Swedish meteorological and hydrological Institute (SMHI), Börjesson, T.; Data preparation: Wolters, S.; Data analysis and visualisation: Wolters, S.; First draft preparation: Wolters, S., Marzec-Schmidt, K.; Review and editing: Wolters, S., Marzec-Schmidt, K., Börjesson, T., Piikki, K,. Söderström, M.

List of tables

Table 1. Local context references (not exhaustive) for research and development of variable rate crop fertilisation. 25

Table 2. Examples of fertilisation strategies. Zadoks stage in decimal scale (DC). Examples of potential N fertilisation strategies are reproduced from the advisory document by the Swedish board of agriculture (2020). 29

Table 3. Comparison of different spectral sensing platforms (UAV = Unmanned aerial vehicle) and their approximate capabilities in terms of resolution, field of view, payload mass and costs (labour and monetary cost). 34

Table 4. Sentinel-2 multispectral instrument (MSI) details. Spatial resolution (SR), Satellites 2A and B of Sentinel constellation (S2A, S2B), Wavelength (λ), bandwidth (W), near-infrared (NIR), Shortwave infrared (SWIR) 42

Table 5. Band central wavelength (λ), bandwidth (W) and name of the nine bands (B1-B9) of the MAIA multispectral instrument (MSI) and the corresponding band in the Sentinel-2 (S2) MSI (see Table 3). NIR is short for near-infrared. 44

Table 6. Vegetation indices in Paper with reference organised by type. Visible spectrum area (VIS). Red-edge chlorophyll index (CI), modified soil adjusted vegetation index (MSAVI2), optimized soil adjusted vegetation index (OSAVI), red-edge inflection point (REIP), transformed chlorophyll absorption reflectance index (TCARI), normalised-difference vegetation index (NDVI, with alternatively using bands 8 and 6 NDRE86), normalised

difference red-edge (with bands 7 and 5: NDRE75). A ratio index between TCARI and OSAVI (TCOS). 47

Table 7. Summary of data and methods used per Paper 50

Table 8. Findings from individual cultivar- or cultivar classes models 52

Table 9. Validation results leave-one-trial-out (LOTO) of crude protein content (CP) content in wheat for the eight modelling strategies, using two model types (linear (a), MARS (b)), and four different combinations of predictor variables (1-4). Indices used are shown in Table 3. Goodness of fit (R^2), mean absolute error (MAE). 54

Table 10. Summary statistics of the different univariate models and the multivariate model for different models. Multivariate model variables are: county, cultivar and accumulated daylength. Organised by modeltype with growing degree days (GDD). With root mean squared error (RMSE (in unrounded DC stages) and sample size (n). 63

List of figures

Figure 1. Simplified and schematic illustration of Zadoks scale stages for winter wheat. Reproduced and adjusted after Fowler (2018). 28

Figure 2. Schematic illustration of spectral data retrieval from image pixels. 30

Figure 3. Example of spectral signature for different nitrogen (N) rates (kg ha $^{-1}$) in winter wheat. Values derived and interpolated from multispectral dataset presented in Paper II. 30

Figure 4. Brantevik 18th of June 2020 trial (trial close-up view) in different band composits (a, b, c). Dataset presented in Paper II. 32

Figure 5. Simplified schematic top-down view of the data retrieval geometry (light grey areas) of different types of sensing instruments..... 35

Figure 6. Workflow for Papers. Linear model (LM), unmanned aerial vehicle (UAV), vegetation index (VI), multivariate regression splines (MARS), nitrogen (N), meteorological analysismodel (MESAN). Numbers refer to sections in this thesis. An oval represents a start, rectangled boxes in indicate a proces, a paralellogram indicates input/output and a diamond indicates a decision or result. 40

Figure 7. Unmanned aerial vehicle (UAV)-borne Maia multispectral instrument (MSI) with 9 filter inlets. 43

Figure 8. Indices (inter-)correlation (Spearman) vs. crude protein content (CP). Vegetation indices: red-edge chlorophyll index (CI), the ratio TCOS,

transformed chlorophyll absorption reflectance index (TCARI), ratio red-edge inflection point (REIP) and optimized soil adjusted vegetation index (OSAVI). 53

Figure 9. Band correlation between processing levels: level 1C and level 2A. 55

Figure 10. Comparison of model prediction of nitrogen uptake by winter wheat (of five indices) at two different Sentinel-2 data processing levels (L1C, L2A). For explanation of vegetation indices, see Table 6. 56

Figure 11. Range of within-field variation (2.5-97.5% inter-percentile) in nitrogen (N) uptake for winter wheat (kg ha^{-1}) at growth stage DC37 for field sizes (a) between 5 and 10 ha, (b) between 10 and 30 ha and (c) 30 ha or more, in southern Sweden..... 57

Figure 12. Average nitrogen (N)-uptake (kg ha^{-1}) in fields > 5 ha in the case study area decribed in Paper I. 58

Figure 13. Nitrogen (N) uptake (high to low in kg ha^{-1}) as (left) predicted by a model based on satellite data and red-edge chlorophyll index (CI) and (right) measured by a tractor-mounted N-sensor. 59

Figure 14. Observed and predicted crude protein content (CP %) in winter wheat on different sites (denoted with letters A-I), with 1:1 line. Results from transfer of a UAV-based univariate linear chlorophyll index (CI) model (a) before and (b) after model transfer to the satellite dataset. 60

Figure 15. Distribution of field observations 2010-2019 for growing degree days (GDD) with a base temperature of 5 ˚C and Zadoks development stage (DC) in winter wheat. 2D point density distribution is indicated by the color scale. 61

Figure 16. Prediction models, for 2019, for a growing degree days (GDD) calculation with T_{base} = 5 ˚C and Zadoks development stage (DC) with leave-one year-out (LOYO) validation for the univariate solutions. The Gompertz model result is shown in panel a and the splines model result in panel b. 62

Figure 17. RMSE for univariate models Gompertz (a) and regression splines (b) per DC class. Early, medium and late maturation stages (subsequently indicated as: E, M, L). Regions Skåne, Västra Götaland and Uppsala (subsequently indicated as: sk, vg, up). Growing degree days (0 degree base GDD0, 5 degree base GDD5). 64

Abbreviations

λ	Wavelength
BOA	Bottom of atmosphere
CI	Chlorophyll index
CP	Crude protein content
DC	Decimal code
DSS	Decision support system
E	Nash-Sutcliffe model error
GDD	Growing degree days
ISPA	International society of precision agriculture
L1C	Level 1C
L2A	Level 2A
LM	Linear model
LOTO	Leave one trial out cross validation
LOYO	Leave one year out cross validation
MAE	Mean absolute error
MARS	Multivariate adaptive regression splines
ME	Mean error
MESAN	Meteorological analysis model
MGRS	Military grid reference system

MSAVI2	Modified soil adjusted vegetation index
MSI	Multispectral instrument
N	Nitrogen
NDRE	Normalised difference red-edge index
NDVI	Nomalised difference vegetation index
NFTS	Nordic field trial system
NIR	Near-infrared
NIT	Near-infrared transmittance
OSAVI	Optimised soil adjusted vegetation index
R^2	Goodness of fit
REIP	Red-edge inflection point
RMSE	Root mean squared error
RS	Regression splines
SD	Standard deviation
SR	Spatial resolution
SWIR	Short-wave infrared
TCOS	Ratio normalised difference red-edge
TCARI	Transformed chlorophyll absorption reflectance index
TOA	Top of atmosphere
VI	Vegetation index
VRA	Variable rate application
UAV	Unmanned aerial vehicle
W	Bandwidth

1. Introduction

Fertiliser prices commonly fluctuate and at the time of writing of this thesis prices are skyrocketting, as a consequence of the political situation in Europe. This situation, together with global climate change and a growing world population make the considerate use of nitrogen (N) fertiliser a topical issue.

Nitrogen (N) fertilisation in winter wheat can be made more efficient by moving from uniform application to meeting variable crop requirements within fields. Taking within field variations in crop growth and timing in to account could result in a more strategic N fertilisation plan that more closely meets N-uptake (N in aboveground biomass) requirement. This facilitates stress-free crop growth and can help mitigate negative environmental effects caused by N losses to the natural environment (*e.g.* Carpenter *et al.* 1998; Sawyer 1994). Hence, variable rate application (VRA) of N fertiliser is proposed as good practise for more precise, and therefore more efficient use of N fertiliser. In production fields, methods using sensing technology make VRA feasible in practice, as sensing data can be used for retrieval of a vast amount of information on crops and soil status, with smaller effort compared to manual field observation (*e.g.* Zhang 2002; Stafford 2000).

Nitrogen fertilisation in winter wheat is commonly scheduled to occur on several occasions during the cropping season, with split doses timed carefully to achieve economically optimal fertilisation and a desirable protein content in the harvested grain (Diacono *et al.* 2013; Efretuei *et al.* 2016). Estimation of fertiliser requirements by predictive models generates information on variations in growing conditions within or between fields. Models using spatial and temporal information and/or resulting prescription maps (a map containing target rate information) are effortlessly communicated to users via information technology based systems, called

decision support systems (DSS). A comprehensive definition of decision support systems for agriculture was formulated in the research project called internet of food and farm:

> "DSS are software-based systems that gather and analyse data from a variety of sources. Their purpose is to smoothen the decision-making process for management, operations, planning, or optimal solution path recommendation." (IoF horizon 2019)

Web-based DSS can be expected to be powerful tools, bringing ease in communication with a large number of farm managers. Prediction models in DSS fascilitate description of relevant crop characterics and can result in advice aimed at fertilisation effort to be undertaken during cropping seasons. The expectation is that DSS can advance fertilisation of winter wheat in practice, by integrating both efficiency and product quality aspects.

1.1 Aim and objectives

The main aim of this thesis work was to develop and evaluate empirical models of crop properties pertinent to N fertilisation of winter wheat. The hypothesis is that various existing datasets can be used in this process, and that the results can be used to improve related functionality in decision support systems. Specific research objectives were as follows:

- To develop a prediction model for crop N-uptake, thereby providing potential upscaling of hand-held sensing measurements (Paper I).

- To explore possibilities of reflectance data by an unmanned aerial vehicle (UAV)-mounted instrument for protein estimation, as well as testing of possibilities for model transfer from UAV- to satellite data scale (Paper II).

- To investigate options for modelling winter wheat crop phenology for systematic timing of fertilisation using weather data (Paper III).

Research questions were as follows:

Q1. Can variable-rate N-uptake be estimated and upscaled using proximal- and remote sensing measurements in winter wheat?

Q2. Can pre-harvest protein content in winter wheat be determined using spectral reflectance data sensed by unmanned aerial vehicle (UAV)-mounted instruments, and can a resulting model be transferred to satellite scale?

Q3. Can winter wheat crop phenology from weather information be predicted for timing of in-season fertilisation?

2. Background

2.1 Precision agriculture

Precision agriculture is the technology driven practise of agricultural science that allows for spatial variability and timing to be taken into account, while planning farm operations (Pierce & Nowak 1999; Stafford 2000). Through co-operation, the international society of precision agriculture (ISPA) strengthens exchange of knowledge between scientists, industry and those working with precision agriculture in practice (ISPA 2022). The ISPA currently uses the following definition of precision agriculture:

"Precision Agriculture is a management strategy that gathers, processes and analyzes temporal, spatial and individual data and combines it with other information to support management decisions according to estimated variability for improved resource use efficiency, productivity, quality, profitability and sustainability of agricultural production." (ISPA 2019)

2.1.1 Spatial variation

Nitrogen content in crop biomass varies spatially within fields, mainly due to variations in soil conditions (Carr *et al.* 1991; Robert 1982). For example, differences in N mineralisation from soil during a growing season can vary over 100 kg N ha^{-1} within a single field (Delin & Lindén 2002). In order to manage the addition of N site specifically, the spatial variability in crop N content within fields needs to be measured and understood.

Fundamental research exploring whether reflectance data could be of value in assessing crop canopy characteristics for agricultural management was conducted in the late 1960s and 1970s, thereby laying the foundations

for precision agriculture (*e.g.* Knipling 1970; Myers & Allen 1968). Precision agriculture and VRA of N fertiliser have been the subject of active research since the mid-1980s (Diacono 2013; Pierce & Nowak 1999; Mulla 2013; Mulla & Khosla 2015; Robert 1982).

2.1.2 Challenges for precision agriculture

Positioning has been a major challenge in precision agriculture. Data in precision agriculture only make sense in relation to the spatial location to which they apply. In the early development of precision agriculture, accurate position recording was an important limiting factor for technical progress. Currently positioning is optimised to such extent that this is no longer considered of concern for precision agriculture research and implementation, when ground positioning equipment is used (Heege 2013; Pierce & Nowak 1999).

An ongoing major challenge is the adoption of precision agriculture. While adoption rates are generally not well known, apart from in a few countries (*e.g.* Australia, Denmark, United States of America), adoption in practice is frequently perceived as slower than initially expected by early advocates (Diacono *et al.* 2013; Lowenberg-DeBoer & Erickson 2019; McBratney *et al.* 2005; Nowak 2021; Robert 2002). Together with financial aspects, practical adoption of precision agriculture appears to be partly governed by developments on the implementation side (Fountas *et al.* 2005; McBratney *et al.* 2005; Thenkabail *et al.* 2019). Implementation of precision agriculture principles via DSS in on-farm practice could improve adoption rates of precision agriculture.

2.2 The importance of decision support systems

Producers, farm advisors and other public agencies continuously seek precise information for crop management. Before any knowledge from research can be applied, software-based tools are frequently required to make the translation of scientific knowledge to practice, and to do this efficiently during cropping seasons with local input-data. In Sweden, cereal cropping fields are often large and therefore much potential for optimisation is in the use of satellite-based DSS, which can support VRA on a larger scale, while using local information.

Local development context

In Sweden, for achieving improved production quantity as well as quality in a multitude of crops a web-based system, called CropSAT, was developed (CropSAT 2022). Scientific knowledge available through the winter wheat (*Triticum aestivum* L.) fertilisation module in CropSAT can help to improve harvest quantity and product quality (*e.g.* protein content in grain kernels) and offers farm managers the possibility to systematise and improve the timing of fertilisation during cropping seasons.

The objectives addressed in this thesis were part of this collaborative effort to improve web-based based decision support for precision agriculture in Sweden. In previous work, DSS for fertilisation were put in place and used to provide basic satellite-based decision support for VRA. Apart from the work described in this thesis, other (related) research has been conducted in Sweden in the past five years. Some examples of recent research aimed at facilitating practical implementation of VRA crop fertilisation in Sweden is summarised in Table 1.

Table 1. Local context references (not exhaustive) for research and development of variable rate crop fertilisation.

Primary work	Reference
Testing use of a combination of Sentinel-2 and DMC satellite data for N-uptake prediction in winter wheat.	Söderström *et al.* 2017
N-uptake modelling in winter wheat using hand-held N-sensor and Sentinel-2 satellite data.	Wolters *et al.* 2019
Satellite data-based modelling of protein content in barley.	Börjesson *et al.* 2019
Satellite data- and experiment -based predictions prediction of in-season N fertilisation rates in winter wheat.	Piikki *et al.* 2022

2.3 Fertilisation of winter wheat

Fertilisation of winter wheat impacts both grain quantity and quality of harvest (*e.g.* Triboi *et al.* 2000; Sinclair & Rufty 2012). In developed countries, nitrogen is therefore often applied in abundance. This can be

improved with VRA, as well as improved timing of application (*e.g.* Efretuei *et al.*; Raun *et al.* 2002).

2.3.1 Winter wheat

Winter wheat are cultivars of wheat sown in autumn which remain in a vegetative state in the winter to resume growth in early spring. In Sweden, in 2022, 472 200 ha (of total arable 2 551 500 ha) were cropped with winter wheat, which has increased substantially since 2010 when 400 000 ha (of total arable 2 633 500 ha) were cropped with winter wheat (SCB 2022). Wheat is densely sown, forming a relatively "closed" canopy, late season. In Sweden, winter wheat is usually sown in the period September-November and harvested in summer or early autumn the following year (July-August in southern Sweden).

2.3.2 Crop chlorophyll content and N-uptake

Sufficient availability of N in soil is crucial for crop biochemistry and therefore for grain quality, since N is an important element in proteins (*e.g.* the enzyme RuBisCo, a key enzyme in photosynthesis), nucleic acids and chlorophyll (Stern *et al.*, 2010; Triboi *et al.* 2000). Nitrogen uptake is estimated as the N-concentration in aboveground plant tissues multiplied by above-ground dry matter mass (Lemaire & Gastal 1997). Environmental stressors, such as a lack of water or nutrients result in non-optimal uptake, which often leads to reduced yields (Delin & Stenberg 2014; Schils *et al.* 2018). An excess of N, on the other hand, could increase risk for lodging (Gooding & Davies 1997). The rate of accumulation of N in plant tissues is highly variable during different stages of crop development and between sites and years, as N-uptake is strongly related to growth rate and biomass accumulation (Gastal & Lemaire 2002; Robertson & Vitousek 2009). Information on crop nutrient status at the time of fertilisation is important for management towards harvest quantities close to yield potential, which is the maximum possible yield for an area (Bingham 1967).

2.3.3 Determination of grain seed quality

Besides affecting yield quantity, N fertilisation has significant effects on grain quality of winter wheat. Physiological processes throughout winter wheat development affect the quality of the harvested grain, and thereby the

efficiency of production in industrial wheat grain processing (Gooding & Davies 1997; Salgó *et al.* 2005). The crude protein content (CP %) concentration in grain is an important quality indicator for the bread wheat industry and is generally taken as a good indicator of quality in a grain batch. Gliadin and glutenin together represent about 80% of total protein in wheat flour and are the most important indicators of the functional properties of wheat flour (Tatham & Shewry 1985; Uthayakumaran & Wrigley 2007). In Sweden, CP concentration is determined on all grain deliveries, using near-infrared transmittance (NIT) sensing. A wheat protein content exceeding 11.5% is desirable for providing bread dough with sufficient gluten to meet quality requirements. Nitrogen is bound in proteins and early-season N availability can have a strong influence on N partitioning at maturity (Curran 1989; Kokaly 2009; Triboi *et al.* 2000; Wuest & Cassman 1998). Overall winter wheat response to N varies between cultivars (Fowler, 2003; Triboi *et al.* 2000).

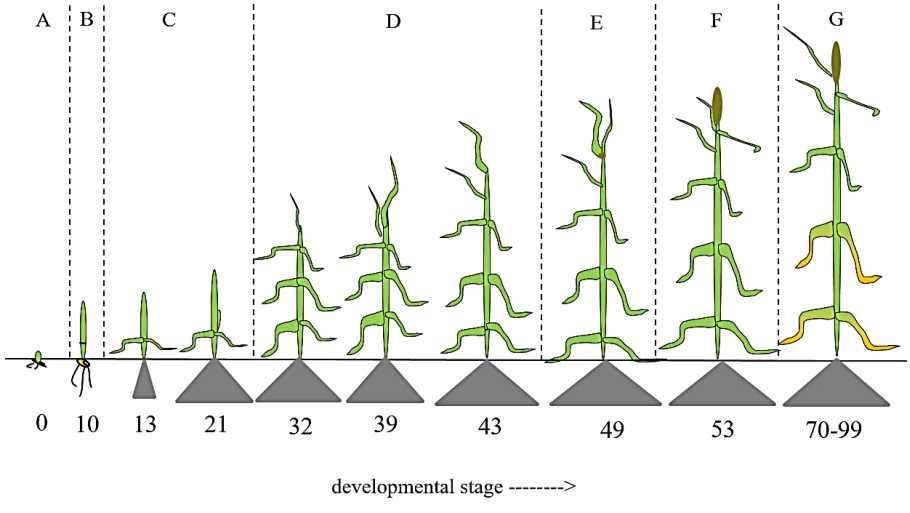
Pre-harvest protein prediction is potentially useful for late-season supplementary nitrogen (N) fertilisation to reach CP targets, or for field partitioning in harvest zones with different expected CP, even within the same season, for different grain qualities. Yield maps from monitors on combine harvesters and combine harvesters equipped with NIT sensors for CP mapping during harvest are available. With the possibility for CP estimation before harvest, there is the option to split fields into harvesting zones, with different expected CP, in the current season.

According to Freeman *et al.* (2003), the ability to use remotely sensed data to determine the N status of wheat and relate it to N accumulation in the grain opens up the possibility for indirect prediction of wheat grain protein content. This is supported by, Basnet *et al.* (2003), who demonstrated that several Landsat satellite bands could be related to crop quality indicators, such as protein content. Bastos *et al.* (2021) reviewed CP predictions, including remote and proximal sensing as well as on-combine sensors, stating that while on-combine sensors outperform remote-sensing methods, there is potential for CP modelling, even pre-harvest.

2.3.4 Phenology and timing of fertilisation

Crop variability has both a spatial and temporal component (Raun *et al.* 2002). Crop phenology, defined as the different physiological development stages of crop growth from planting to harvest, is an important factor

influencing fertiliser requirement in cereals. Nitrogen fertilisation could thus be targeted to specific cereal crop development stages in order to meet grain quantity and quality goals more precisely. For example, Petterson (2007) suggested adjustment of N application rate at a specific development stage to control grain CP content at harvest in malting barley and also winter wheat. Descriptive scales, *e.g.* those developed by Feekes (1941), Zadoks *et al.* (1974), Witzenberger (1989) and Lancashire (1991), are commonly used to describe crop development stages. Including crop phenology in a DSS helps schedule N fertilisation with increased precision, and in Sweden this is commonly done using the Zadoks winter wheat decimal code (DC) development scale (Zadoks *et al.* 1974). This numerical ordinal scale (ranging 0-100) is accompanied by qualitative description of observable crop developmental appearance (Figure 1). Fertilisation in Swedish winter wheat production can be timed using different strategies (examples in Table 2).



A. emergence, B. one shoot, C. tillering, D. stem extension, E. head visible F. heading, G. ripening

Figure 1. Simplified and schematic illustration of Zadoks scale stages for winter wheat. Reproduced and adjusted after Fowler (2018).

Table 2. Examples of fertilisation strategies. Zadoks stage in decimal scale (DC). Examples of potential N fertilisation strategies are reproduced from the advisory document by the Swedish board of agriculture (2020).

	Nitrogen dose % of total / Zadoks developmental stage (DC)		
Strategy	1	2	3
I	70% / 24	30% / 39	
II	60% / 24	20% / 39	20% / 39
III	25% / 22	50% / 30	25% / 39
IV	25% / 22	75% / 30	
V	100% / 24		

Temperature is a main factor for the general rate of phenological development of a plant (Crauford & Wheeler 2009; Went 1953). The photoperiod is another factor of relevance for modelling crop development, with increasing day length accelerating the rate of development (Garner & Allard 1920; Giese 2013; Jamieson 1995).

2.4 Sensing instruments, data and platforms

Data for VRA is gathered via a wide range of instruments, from which optical sensing instruments are common. Optical instruments van be mounted on different platforms and can gather reflectance data, which holds valuable information on crop health and vigour (Taylor *et al.* 2021).

2.4.1 Reflectance

Chlorophyll in crop leaves provides valuable information on the physiological status of plants and can be measured indirectly based on reflectance of light (Gitelson 2003). Crops absorb, reflect, or transmit incoming light (irradiance). Crop chlorophyll absorbs much light in the visible part of the electromagnetic spectrum. Canopy reflectance is expressed as the fraction of incident radiation reflected back, with 0 denoting total absorption and 1 denoting total reflectance (Gates *et al.* 1965). In an image, each pixel is constructed from different segments in the electromagnetic spectrum (Figure 2).

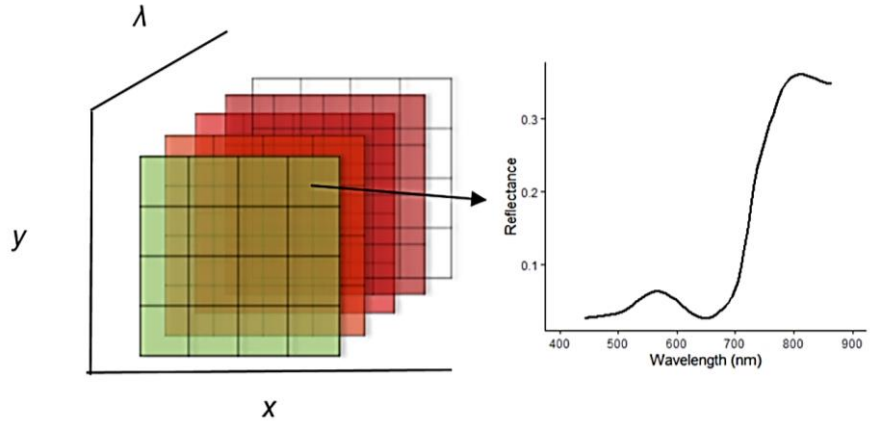


Figure 2. Schematic illustration of spectral data retrieval from image pixels.

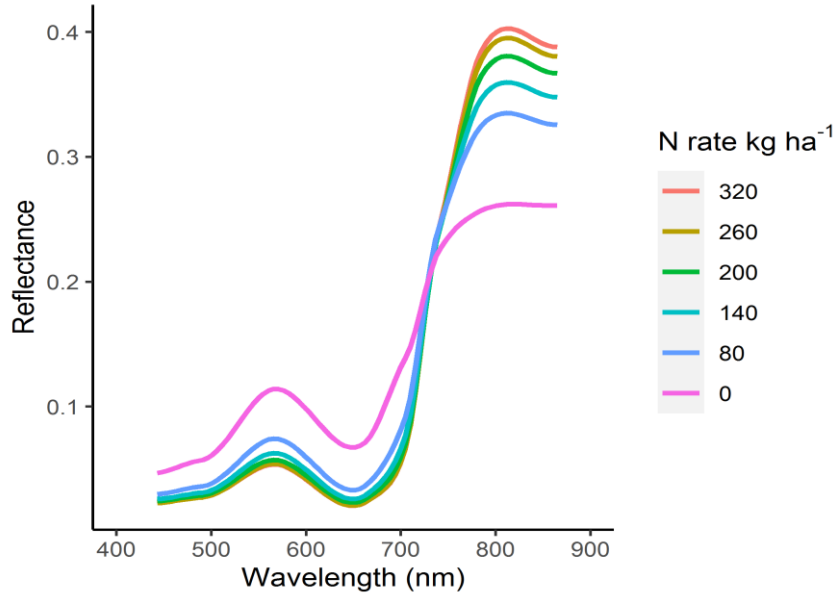


Figure 3. Example of spectral signature for different nitrogen (N) rates (kg ha $^{-1}$) in winter wheat. Values derived and interpolated from multispectral dataset presented in Paper II.

Light reflected from a crop is represented in its spectral signature. The spectral signature indicates health status and stresses in plant development, as can easily be observed by overlooking the signatures of winter wheat subjected to different fertiliser rates (Figure 3). Note that Figure 3 also can be interpreted to indicate the wavelengths in which plant responses are most different or similar.

2.4.2 Composites and vegetation indices

Data from different wavelengths can be used for calculations by taking data from wider (multispectral) or finer (hyperspectral) width (spectral range in section, *i.e.* bands) across the spectrum. Composites can be calculated by selecting bands from the electromagnetic spectrum, which can help to highlight or suppress crop characteristics for (pre-)analysis. Figure 4 shows an example of three different composites created from different data. Using visible light (a) different N-levels in a field-trial can be detected, however, using a false colour composite replacing red with near-infrared reflectance (b; NIR) will highlight vegetation from background (vegetation gets highlighted in a red color). Variations in field reflectance values are usually caused not only by crop characteristics, but also by atmospheric and soil effects (*e.g.* Gebbers *et al.* 2013). Including NIR as well as red-edge reflectance (c) can help highlight information on crop health and N-status (Guo *et al.* 2018; Hatfield *et al.* 2008).

Use of a vegetation index is another way of highlighting crop or canopy properties. Several studies have therefore attempted to calculate refined ratios in order to minimise the disturbances from such effects (Clevers 1986; Huete & Jackson 1988). Vegetation indices (VIs) are mathematical composits from surface reflectance, at two or more wavelengths, commonly in the areas of green, red and infrared.

The earliest investigations using multispectral satellite remote sensing data involved a ratio of Landsats bands 7 and 5 introduced by Rouse *et al.* (1973). Since then, more than 500 VIs have been reported in the scientific literature, although it is important to note that not all of these calculations have been extensively tested (Henrich & Brüser 2022). Apart from vegetation indices based on multispectral analysis, hyperspectral (narrow band) vegetation indices have also been developed (*e.g.* Haboudane *et al.* 2002).

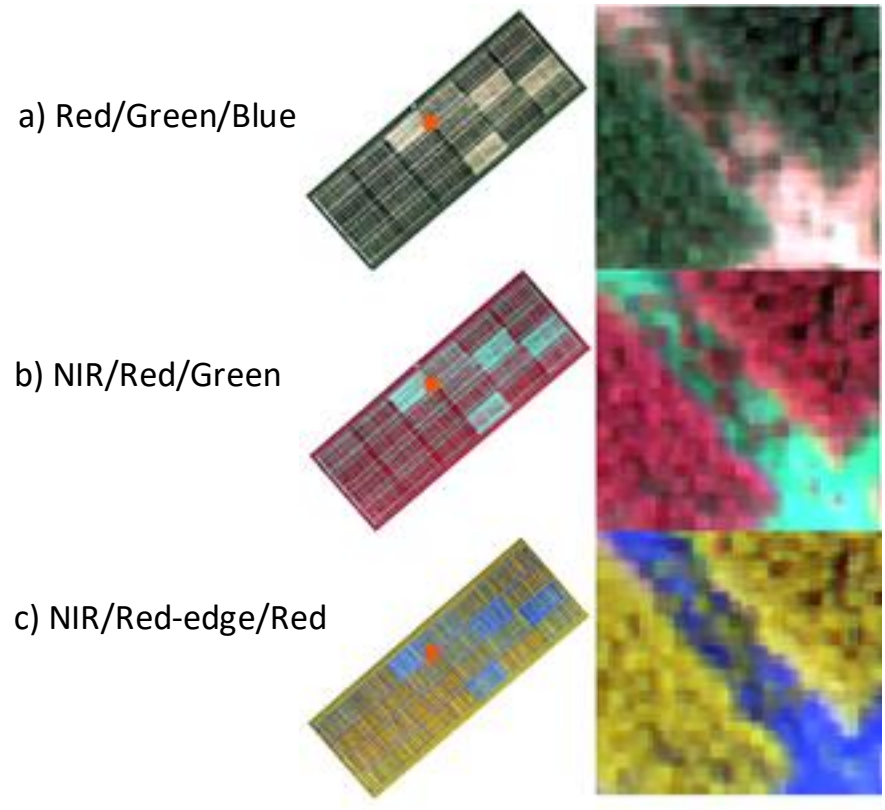


Figure 4. Brantevik 18th of June 2020 trial (trial close-up view) in different band composits (a, b, c). Dataset presented in Paper II.

2.4.3 Active and passive sensing

Remote sensing applications used in precision agriculture are primarily passive, *i.e*. based on sensing reflectance of the sun's visible and near-infrared light by soils or crops (Mulla & Khosla 2015; Thenkabail *et al.* 2019). Active sensors, such as versions of the Yara N-sensor (Yara Gmbh, Hanninghof, Germany) described by Reusch (2005, 1997), are another possibility. This type of sensor uses its own source of radiation to "illuminate" objects (ESA 2022).

2.4.4 Sensor mounts

When sensors are mounted on ground vehicles or hand-held rather than aerial platforms, the term proximal sensing is used. An example of a potential range

for resolution, field of view, payload mass and costs of different platforms is summarised in Table 3. In principle, when practically feasible, any sensor could be carried on any platform. Sensing instruments are central for the VRA approach taken, although the choice of platform influences how the instruments can be used. The platform that carries the instrument and/or its operator largely determine the availability and quality and quantity of spectral data.

A common challenge in the use of remote sensing data is timeline discontinuity in field datasets. For satellite platforms, this is usually the result of cloud cover and other atmospheric effects (Luca *et al.* 2022). Large-scale spatial coverage by satellite data generates potential for management, giving detailed within- and between field information on a large scale. Sensors mounted on a UAV platform are less affected by atmospheric effects and vehicle sensing can support on-the-go solutions, which in grain production can be relevant for providing information while fertilising or harvesting (*e.g.* Taylor *et al.* 2021).

Table 3. Comparison of different spectral sensing platforms (UAV = Unmanned aerial vehicle) and their approximate capabilities in terms of resolution, field of view, payload mass and costs (labour and monetary cost).

	Resolution	**Field of view**	**Payload mass**	**Costs**
Hand-held device	From mm	Up to several m^2 per observation	Can be limited	High in time, medium in purchase of instruments
Tractor (or comparable ground-vehicle)	0-1 m	50 m^2/ sec	Can be large	On-the-fly: low in time, medium in purchase of instruments
UAV	0.5–10 cm	50–5000 m^2	Can be limited	Low in time, medium in purchase of instruments
Satellite	1–25 m	10–300 km	Can be large	Very low in time, often free to use

Data from UAV-based models can be used as supporting information to satellite-based models, due to their finer spatial resolution and fewer atmospheric effects, compared with satellite sensing. Hand-held instruments are useful in the field, but retrieving data with hand-held instruments is generally considered labour-intensive. This method usually gives point data, which can be interpolated when spatial autocorrelation in the crop or soil-characteristic of interest can be assumed. Hand-held instrument data provide valuable *in situ* information that can be used to calibrate remote sensing data. It is worth mentioning that using hand-held instruments can bring flexibility over instruments mounted on vehicles, *e.g.* there are more options to set sensing instrument height and to control for moisture effects in the canopy.

2.4.5 Sensor geometry and view angle

Each sensing platform has pre-defined geometric opportunities and restrictions. A simplified schematic diagram of relevant geometry is given in Figure 5, where the areas in grey show the measurement area covered by hand-held, tractor-mounted sensing, UAV or satellite sensing.

It is important to note that the different types of sensing instruments differ not only in their geometry, but also in their viewing angle. Sensor sensing at nadir and oblique views can give very different results, which should be considered when comparing data from different sensing platforms (Lu *et al.*, 2019). Ideally, the spatial resolution of any instrument should at least match the spatial resolution of the machinery footprint used in the VRA application.

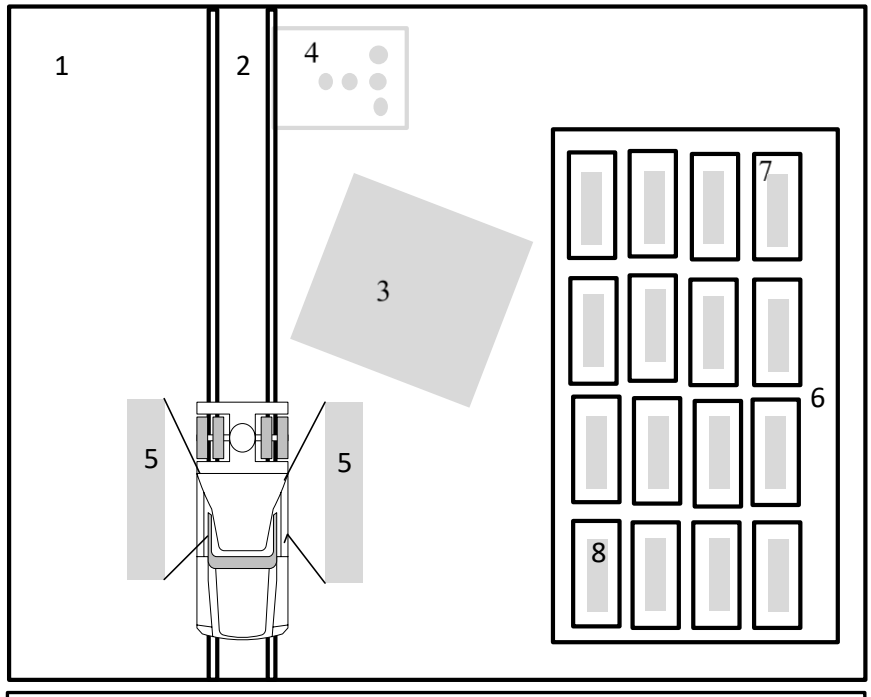


Figure 5. Simplified schematic top-down view of the data retrieval geometry (light grey areas) of different types of sensing instruments.

2.5 Options for collecting field information

2.5.1 Field sampling

Measured or sensed field-data is the source of *in-situ* information. Field data can also be used to calibrate remote sensing models. When carrying out field measurements using a sensor for reflectance, the field of view must be taken into account. Measuring close to noon is preferable for passive sensors (10:00-14:00 h). Wind can act as a disturbance when carrying out measurements. In finer spatial resolutions in particular, wind can be a cause of inconsistency in a canopy structure Clevers (1986), which gives inconsistency in the data.

In situ field sampling can be eased with *in-situ* threshing instruments. Data retrieval for lab-analysis can be done with a field threshing system (*e.g.* Minibatt+, Godé, Le catelet, France). This reduces the volume of samples to carry out from the field significantly.

2.5.2 Field trials

Over the years, much experience has been gained from preparing and analysing field trials or experiments. Field trials are valuable for testing new crop varieties, new fertilisers or other crop treatments. With their experimental form, they can help to standardise crop treatment results by excluding influences other than those experimentally applied in the study. Trials therefore make it possible to study the effects of crop treatments in a rather controlled manner. Additionally, field trials can be randomised to reduce potential unwanted effects.

2.6 Multi-source data

Recent research reported on options for using multi-source data (Reyniers *et al.* 2006; Sarvia *et al.* 2021; Wang *et al.* 2020). While some efforts has been taken to explore the potential of using multi-source data, finding the most effective combination of multi-source data requires more research (Wang *et al.* 2020).

Data from a ground optical sensor can be compared with aerial optical imagery to get a better understanding of sensing precision, when using data sensed from different platforms in continuous measurements (Reyniers *et al.* 2006). Data acquired via sensing instruments from multiple aerial platforms can be compared to study potential instrument-related effects on any crop information derived (Sarvia *et al.* 2021).

If multi-source information via a DSS is to be used for a particular site throughout a cropping season, a better understanding is needed of data interchangeability from similar instruments sensed via different platforms. This can be done by studying the effects of technical instrument properties. Not by solely comparing instruments, but also by deducing to specific crop traits, such as N-uptake.

3. Materials and methods

3.1 Workflow

A summary of the workflows for the embedded Papers is presented in Figure 6.

In Paper I, hand-held derived N-uptake approximations and Sentinel-2 satellite data were combined for predicting N-uptake. Five vegetation indices were calculated for satellite data of two satellite data processing levels. Upscaling was done via linear regression (LM) and a fitted power function. From the results for five different vegetation indices, one model was selected to be applied on a study area with the geometry of two merged satellite tiles in a wheat-dense production region. An upscaled result was compared with tractor N-sensing measurements. The model result was implemented in CropSAT.

In Paper II, prediction models for crude protein content (CP %) in winter wheat (*Triticum aestivum L.*) based on multispectral reflectance data from field trials were developed and evaluated for independent trial sites. Two different prediction methods were tested: LM and multivariate adaptive regression splines (MARS) for seven vegetation indices. Unfertilised blocks (representing zero-plots) and blocks fertilised with a very large amount of N (representing max-plots) were used to test possibility for predicting with zero- and/or max plots. Unmanned aerial vehicle (UAV)-borne camera data in nine spectral bands had a similar specification to nine bands of Sentinel-2 satellite data. This gave the possibility to test model transfer from a UAV-based model to a small satellite dataset on independent sites.

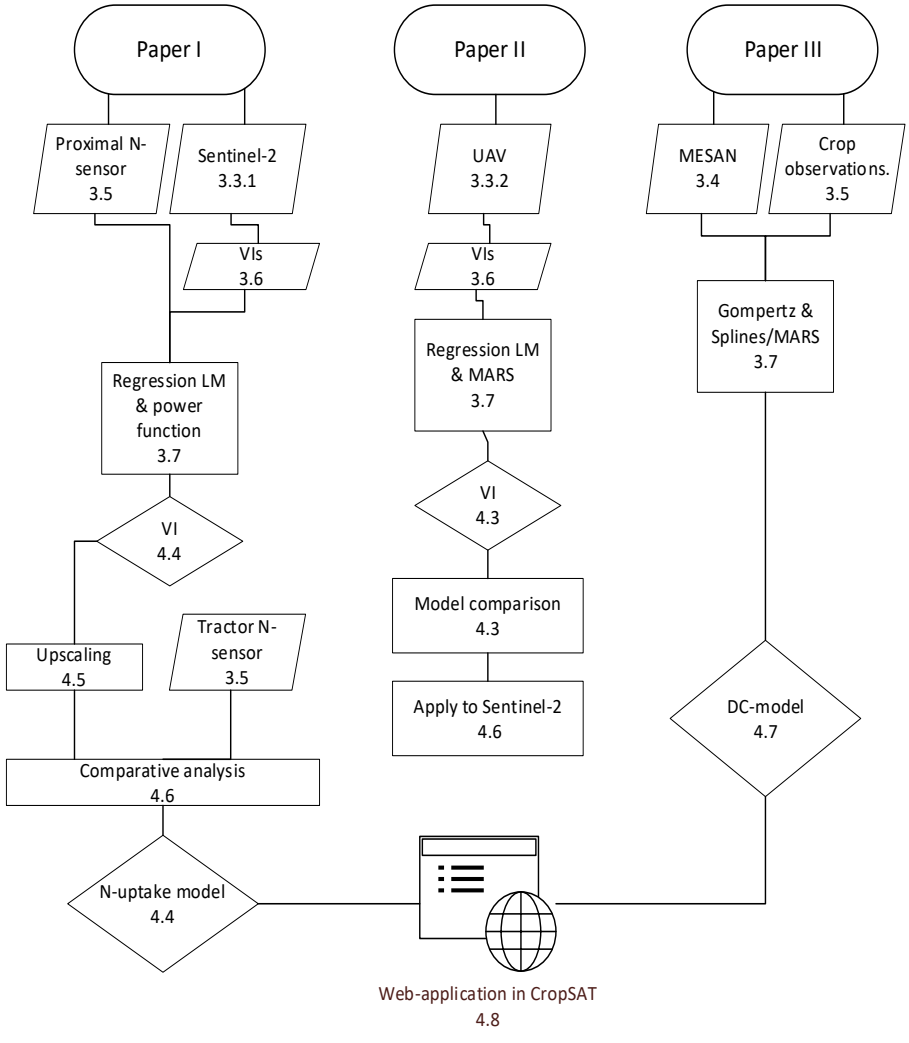


Figure 6. Workflow for Papers. Linear model (LM), unmanned aerial vehicle (UAV), vegetation index (VI), multivariate regression splines (MARS), nitrogen (N), meteorological analysismodel (MESAN). Numbers refer to sections in this thesis. An oval represents a start, rectangled boxes in indicate a proces, a paralellogram indicates input/output and a diamond indicates a decision or result.

In Paper III, thermal-time-based models were developed and evaluated for use in a DSS and for getting a better understanding of implementation and accuracy in different crop developmental stages. Developmental stages expressed in DC (Zadoks et al. 1974) of winter wheat from field observations

in monitoring plots were compiled with weather data. Models predicting growth stages were developed based on weather variables, calculated to thermal-time units. Three different prediction methods Gompertz fit, regression splines (RS) and MARS were used to find the best-fitting procedure for growth stage prediction. Testing was done for different counties and cultivars were classified for testing with differences in speed of maturation.

3.2 Study area

The area most suitable for winter wheat cropping in Sweden is found in the southern counties and generally is located between 55° to 61° N and from 10° to 19° E. Paper I and III have the entire southern Sweden as a study area and Paper II focussed on the southernmost region of Sweden, Scania (Skåne).

3.3 Remote sensing

3.3.1 Sentinel-2

The satellite data used were taken from the Copernicus Sentinel-2 mission, a constellation comprising two polar-orbiting satellites phased at 180° to each other. The 13-band multispectral instrument (MSI) that is the payload on the Sentinel-2 satellites uses a pushbroom concept, meaning that the collected data consist of large arrays in an orbital swath (width 290 km), pushed by the forward motion of the spacecraft (ESA 2022). This results in a revisiting time of 5 days at the equator and under cloud-free conditions this gives data every 2-3 days at mid-latitudes like Sweden. Once data are transferred from the ground segment operations via collaborative interfaces, they are available for free. The data are organised via military reference grid system (MRGS) tiling, while containing within-tile overlapping of the swath. Instrument properties of the 13-band MSI carried by the Sentinel-2 satellites are described in Table 4.

Some examples of previous satellites of interest for agriculture were Landsat 1-7 (at most seven bands) and SPOT (four bands). A large improvement in the Sentinel-2 constellation compared with these satellites is the revisiting time, which dropped significantly (by 7 and 11 days, respectively). Date products from this mission are available in different

processing levels (Level-1B, Level-1C and Level-2A) (Drusch *et al.* 2012; ESA 2022). L1C products are 100 x 100 km2 gridtiles with only radiometric and geometric corrections, including orthorectification and spatial registration (ESA 2022). Level-2A (L2A) correction algorithms are supplemented with a series of threshold tests. For these products, top of atmosphere (TOA) reflectance data are modified using auxiliary information to create data resembling bottom of atmosphere (BOA).

Table 4. Sentinel-2 multispectral instrument (MSI) details. Spatial resolution (SR), Satellites 2A and B of Sentinel constellation (S2A, S2B), Wavelength (λ), bandwidth (W), near-infrared (NIR), Shortwave infrared (SWIR)

Band	**SR**	**S2A**		**S2B**	
	(m)	**λ (nm)**	**W (nm)**	**λ (nm)**	**W (nm)**
1 – Coastal aerosol	60	443.9	27	442.3	45
2 – Blue	10	496.6	98	492.1	98
3 – Green	10	560.0	45	559.0	46
4 – Red	10	664.5	38	665.0	39
5 – Vegetation red-edge	20	703.9	19	703.8	20
6 – Vegetation red-edge	20	740.2	18	739.1	18
7 – Vegetation red-edge	20	782.5	28	779.7	28
8 – NIR	10	835.1	145	833.0	133
8A – Narrow NIR	20	864.8	33	864.0	32
9 – Water vapour	60	945.0	26	943.2	27
10 – SWIR – Cirrus	60	1373.5	75	1376.9	76
11 – SWIR	20	1613.7	143	1610.4	141
12 – SWIR	10	2202.4	242	2185.7	238

3.3.2 Unmanned aerial vehicle

Unmanned aerial vehicle (UAV)-borne MAIA MSI (MAIA S2, SAL Engineering, Eoptis and Fondazione Bruno Kessler, Italy) with nine spectral bands using similar specification to nine visible (VIS) and near infrared (NIR) bands of Sentinel-2 satellite data was used in Paper II (Figure 7). The UAV MSI spatial resolution, unlike the Sentinel-2 (10-60 meters resolution), the MAIA MSI is not different in spatial resolution for bands. Data were pre-processed with MultiCam Stitcher Pro software, provided with the instrument. The output was a nine-band tiff for each acquisition. Afterwards, orthomosaics were generated using the web application solvi.ag (Solvi AB, Gothenburg, Sweden).

An overview of the spectral properties of the MAIA MSI is given in Table 5, where a comparison with Sentinel-2 satellite data is made in the right-hand column.

Figure 7. Unmanned aerial vehicle (UAV)-borne Maia multispectral instrument (MSI) with 9 filter inlets.

Table 5. Band central wavelength (λ), bandwidth (W) and name of the nine bands (B1-B9) of the MAIA multispectral instrument (MSI) and the corresponding band in the Sentinel-2 (S2) MSI (see Table 3). NIR is short for near-infrared.

MAIA-S2 sensor	λ (nm)	W(nm)	Band name	Corresponding band in S2 MSIs
S1	443	20	Violet	1
S2	490	65	Blue	2
S3	560	35	Green	3
S4	665	30	Red	4
S5	705	15	Red-Edge 1	5
S6	740	15	Red-Edge 2	6
S7	783	20	NIR 1	7
S8	842	115	NIR 2	8
S9	865	20	NIR 3	8A

Calibration panels

Panels with known reflectance can serve as field calibration for remote sensing reflectance data. Reflectance panels have a known reflectance, so raw pixel values for the panel can be converted to reflectance values and used to calibrate other images with the same lighting conditions in the field. In Paper II, panels with near-lambertian reflectance characteristics (2%, 9%, 23%, 44% and 75%), within the 400-900 nm range of the electromagnetic spectrum were used.

3.4 Mesan grid

The Mesan weather grid system is a grid-based meteorological analysis model that can be conveniently accessed for use in DSS in Sweden (Häggmark *et al.* 1997). Using weather data from the Mesan grid is supplementary to using data from individual weather stations, because of the limited number of weather stations available (www.smhi.se). Information

from individual weather stations is interpolated to grid format using meteorological interpolation models. The grid cell size is 11 km x 11 km. Collection of daily temperature data for the study area in this thesis resulted in around 40,000 records per year for a period of 10 years. In Paper III, these data were connected to growth stage (Zadoks DC observations).

3.5 Field observations

In Paper I, hand-held data were retrieved using a Yara N Sensor® (Yara Gmbh, Hanninghof, Germany) for proximal measurements during two seasons (2017 and 2018). This allowed for a large dataset (> 700) observations) for estimation of N-uptake to be generated. N-uptake data were also gathered by tractor scanning in 13 fields (ranging in size from 2 to 30 ha) in 2017 in the area around the village of Ardala, in Västra Götaland County in southwestern Sweden. A passive Yara N-Sensor® was used, and the tractor was driven in a regular pattern with 24 m between tram lines over the fields, registering N-uptake every second (with approximately 3 m between registrations). Field data on harvested protein content (CP in % of dry matter) at the end of anthesis until milk maturity were extracted in five field trials in 2019 and 2020 for Paper II, where for each plot in the Nordic Field Trials System (NFTS; <https://nfts.dlbr.dk>; Danish Technological Institute and SEGES, Aarhus, Denmark) CP was determined by FOSS Infratec1241 NIT equipment (FOSS, Hillerød, Denmark). Weekly observations of the winter wheat developmental stage in Paper III were done on a fixed day of the week for years from 2010 to 2017, and multiple observations per week per site for the latest years in the dataset (2018-2019), according to instructions described in Eriksson (2022). GDD was calculated from January, each year, using Equation 1, using either 0°C or 5°C for base temperature (GDD_{base}). The subscript *i* denotes the day number. The unit for GDD is °C days. T_{max} is the daily maximum temperature and T_{min} is the daily minimum temperature.

(1)

$$\text{Growing degree days } (GDD) = \sum_{i=0}^{n} \frac{\text{Tmax}_i + \text{Tmin}_i}{2} - T_{base}$$

The multivariate adaptive regression splines model used photoperiod, county and cultivar as additional variables. Photoperiod expressed in accumulated photoperiod (DL) was calculated from January according to Forsythe *et al.* (1995). Accumulated DL was a sum of daylight hours per day (dl_i) up to the date of each field observation (Equation 2). Also in this equation the subscript *i* denotes day number. The unit for DL is daylight hours.

(2)

$$\text{Accumulated photoperiod } (DL) = \sum_{i=0}^{n} dl_i$$

3.6 Vegetation indices

Vegetation indices were used in Paper I and II. Formula's for the index calculation of respective Papers respective are summarised in Appendix A1. A more detailed justification for the choice of each equation individual index is described in the respective Papers. The indices used were categorised based on common characteristics (Table 6).

Table 6. Vegetation indices in Paper with reference organised by type. Visible spectrum area (VIS). Red-edge chlorophyll index (CI), modified soil adjusted vegetation index (MSAVI2), optimized soil adjusted vegetation index (OSAVI), red-edge inflection point (REIP), transformed chlorophyll absorption reflectance index (TCARI), normalised-difference vegetation index (NDVI, with alternatively using bands 8 and 6 NDRE86), normalised difference red-edge (with bands 7 and 5: NDRE75). A ratio index between TCARI and OSAVI (TCOS).

	Platform	Green	Red	Red-edge	Red-edge	Near-infrared	Near-infrared		
	Sentinel-2 MAIA	3	4	5	6	7	8		
Index type	Index							Reference(s)	Paper(s)
Chlorophyll ratio	CI				■	■		Gitelson *et al.* 2003	I, II
	TCARI	■	■	■				Kim *et al.* 1994; Haboudane *et al.*	II
Red-red-edge	REIP		■	■	■	■		Guyot *et al.* 1988	II
Soil- adjusted	MSAVI2		■				■	Qi *et al.* 1994	I, II
	OSAVI		■			■		Huete 1988; Rondeaux *et al.*	II
Normalised difference	NDVI		■				■	Rouse *et al.*, 1974	I, II
	NDRE85			■			■	Derived from Barnes *et al.* 2000	I, II
	NDRE75			■		■		Barnes *et al.* 2000	II
	NDRE86				■		■	Derived from Barnes *et al.* 2000	I, II
Combination of indices	TCOS		■	■		■		Haboudane 2002	II

3.7 Statistics

Both parametric methods and non-parametric methods were tested. Modelling was done using LM, RS or MARS. In terms of model performance, experience has shown that flexible models work well on both linear and non-linear datasets, while linear regression approaches often struggle with non-linearity (Hastie *et al.* 2009). Models that are intrinsically linear in their mathematical form, can be adapted to non-linear patterns in the data, by manually adding non-linear model terms such as squared terms or interaction effects, or other transformations of the original elements used in the calculation (Boehmke & Greenwell; Friedman 1991). In this way, regression splines estimates hinge functions with knots, on critical model areas, where slope requires change. This method can handle both continuous and categorical data and the method works well with multivariate datasets.

The Gompertz function is a mathematical model used for time series (Equation 3). Where k, and lag are tuning parameters and x is the numeric vector of values at which to evaluate the model. It is a sigmoid function commonly used to describe biometric measurements as being slowest at the start and at the end of a given period (Gompertz 1825; Tjørve & Tjørve 2017). The limits of the curve are set by Y_0 and Y_{max}.

(3)

$$DC = \mathrm{Y}_0 + (\mathrm{Y}_{\max} - \mathrm{Y}_0)^{-\frac{(k*(lag-x))}{Y_{max}-Y_0}+1}$$

Validation

Validation on independent data is used for assessing prediction model performance. This can be done using data on independent sites, such as independent trial sites (called leave-one-trial-out cross-validation, LOTO), or on data that differ temporally, such as different cropping years (leave-one-year-out cross-validation, LOYO). Spearman's rank correlation (Spearman 1904) was used in Paper I and II.

Model error quantification

The accuracy indicators used were goodness of fit (R^2), model error (E; Nash & Sutcliffe 1970) mean error (ME), mean absolute error (MAE) and root mean squared error (RMSE).

Spatial statistics

Spatial statistics were used for geoprocessing of data in all Papers. Ordinary block kriging (as explained in Burrough & McDonnell, 1989) of tractor-mounted sensor scanning data was used to convert the data (Yara Gmbh, Hanninghof, Germany) to raster format for N-uptake calculation in Paper I. Zonal statistics were used in Paper II for calculation of median reflectance in trial plots. A spatial join was used to connect crop development observations to MESAN grid data in Paper III.

3.8 Software

The main analysis software used were SQL server, accessed via SQL server management studio, R accessed via RStudio, Python accessed via Visual studio code and ArcGIS 10 (Microsoft, Redmond, Washington, USA; R Core Team; RStudio, PBC, Massachusetts, USA; Python Software Foundation, Delaware, USA; ESRI, California USA) and Microsoft Office professional was used (Microsoft corporation, Redmond, Washington, USA).

3.9 Data and method summary

An overview of main data and methods for each Paper is given in Table 7.

Table 7. Summary of data and methods used per Paper

	Paper I	Paper II	Paper III
Main Source	Satellite/Yara-hand-held and tractor mounted N-sensor	UAV/ Tissue analysis	MESAN/field observations
Model n (application set)	251	96(34)	26729
Validation design	LOYO	LOTO	LOYO
Result	Linear regression	Linear regression/Regression splines/MARS	Gompertz/Regression splines/MARS
Study area	Southern Sweden	Skåne county	Southern Sweden

4. Results

4.1 Cultivar specific models (Paper I, II & III)

Results were implemented using a general model or multiple cultivar specific models. In Wolters *et al.* (2019) and in Paper II cultivar specific models were tested. In Paper III early, medium and late maturation classes of cultivars were tested. In all of the aforementioned studies cultivar specific models gave a different result from general models. In the results of Paper I and II, cultivar specific models show substantial difference in performance, this indicates that there can be a benefit in implementing models for specific cultivars.

Table 8 summarises the findings from the cultivar specific models tested either with individual cultivars, with cultivar classes, or classed by early, medium or late maturation rates. The amount of cultivars present in studies varied between the studies and not all cultivars were used to build prediction models. The dataset used in Paper I showed in Wolters *et al.* (2019) that selected cultivars had a better goodness of fit (R^2) in a cultivar specific N-uptake prediction model ($R^2 = 0.84$ vs. $R^2 = 0.79$). No further cultivar specific testing was performed in Paper I. Testing of cultivar-specific models in Paper II gave large variation in results, both between cultivars and between a cultivar specific and a general model, with R^2 ranging between 0.33 to 0.86 and MAE ranging between 0.40 to 1 % CP.

Maturation rates by classes of cultivars are cultivars grouped by the expected properties of that cultivar in terms of maturation speed. The different classes in Paper III were not very different in performance (RMSE 4.54-4.66 DC). The results in Paper III showed no clear benefit in distinguishing between maturation rate classes.

Table 8. Findings from individual cultivar- or cultivar classes models

Paper I	Paper II	Paper III
Types of model tested (individual, cultivar specific)		
No cultivar specific models were tested in Paper I. Cultivar specific models were tested for this dataset in Wolters *et al.* (2019).	Cultivar specific models were tested.	Early, medium and late maturation rates of classes of cultivars were tested.
Number of different cultivars represented in the dataset		
14 different cultivars.	10 different cultivars per year, 20 unique cultivars.	45 different cultivars.
Benefit from using cultivar specific models		
Yes, A cultivar named 'Julius' had lower error than other cultivars.	Yes. Large differences in performance between cultivars.	No. Differences in performance of classed cultivars were small.

4.2 Performance by growing stage (Paper I & III)

Developmental limits could be of relevance for evaluating model fit in developmental stages most crucial for fertilisation management decision making. The very beginning of crop development and the latest stages in maturation could in some cases be more difficult to predict from field data and these are also expected to be less crucial for fertilisation management decisions. In Wolters *et al.* (2019), by subsetting the data into certain growth stages (DC 31-45) predictions were not improved. A minor difference in performance was seen when models were restricted to DC 22-53 (Paper I, R^2 in Figure 5 & Table 2). Performance was slighly improved with a developmental limit. Moreover, in Paper III, prediction models showed different performance (RMSE) over different DC-classes. This result indicates that it is important to evaluate model performance in different stages of crop development. Stage-wise evaluation of models can improve

understanding of error in different periods during crop development (more details are given in section 4.7.1).

4.3 UAV-based CP predictions (Paper II)

Vegetation indices intercorrelated, as shown in the UAV reflectance vs. protein dataset (Paper II). Some indices were highly correlated (-0.9 < r > 0.9), such as CI-NDRE75/TCOS/REIP and TCOS-TCARI (Figure 8).

Vegetation indices of a similar type (see Table 5) can be expected to intercorrelate more strongly in a reflectance dataset. The indices that correlated most strongly (positive) with CP were CI and NDRE75, while the smallest correlation (positive) with grain protein was found for OSAVI and REIP (Figure 8). A moderate negative correlation was found between TCARI, TCOS and CP.

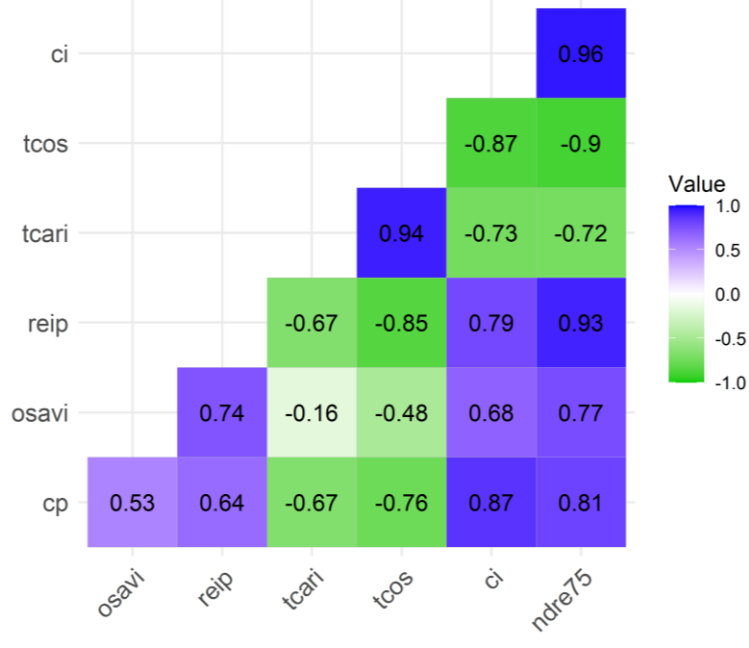


Figure 8. Indices (inter-)correlation (Spearman) vs. crude protein content (CP). Vegetation indices: red-edge chlorophyll index (CI), the ratio TCOS, transformed chlorophyll absorption reflectance index (TCARI), ratio red-edge inflection point (REIP) and optimized soil adjusted vegetation index (OSAVI).

The models tested showed some potential for predicting winter wheat grain CP from reflectance data. However, the models did not give consistent results for all sites (for interpretation of model performance in specific sites see Paper II, Figure 4). Linear methods performed well and showed consistent results between different sites. MARS modeling gave unstable results and did not optimize on a dataset this small in size.

Validation results of uni- and multivariate LOTO are shown in Table 9. Using the linear regression method and the best-performing index (CI), gave $R^2 = 0.71$ and MAE 0.64% CP. Inclusion of zero-plots reduced the accuracy of this model slightly ($R^2 = 0.60$, MAE 0.71% CP). The max-plot CI value (CI-max) did not influence the prediction outcome, $R^2 = 0.71$, MAE = 0.64% CP. A model with both the zero-plots and max plot gave $R^2 = 0.60$ and MAE = 0.71% CP for the linear method. The MARS results were unstable, however, showing better prediction with zero-plots included than with max-plots included and generally showing large variation (R^2 0.36-0.70, MAE 0.64-1.19 % CP).

Table 9. Validation results leave-one-trial-out (LOTO) of crude protein content (CP) content in wheat for the eight modelling strategies, using two model types (linear (a), MARS (b)), and four different combinations of predictor variables (1-4). Indices used are shown in Table 3. Goodness of fit (R^2), mean absolute error (MAE).

Strategy	**R^2**	**MAE (% CP)**	**Predictors in final model**
1a	0.714	0.64	CI
2a	0.602	0.71	CI, CI-zero
3a	0.710	0.64	CI, CI-max
4a	0.601	0.71	CI, CI-zero, CI-max
1b	0.504	0.90	CI, OSAVI, TCOS, REIP
2b	0.630	0.70	CI, CI-zero, OSAVI, TCOS
3b	0.363	1.19	CI, CI-max, OSAVI, TCOS, REIP
4b	0.703	0.60	CI, CI-zero, CI-max, NDRE75, REIP, TCARI, OSAVI

4.4 Sentinel-2 data processing levels (Paper I)

Atmospheric processing of Sentinel-2 data in critical bands for vegetation modelling gave a systematic trend of deviation from the 1:1 line, when correlated (Figure 9). As should be expected, Spearman correlation was high ($r > 0.98$) between the L2A and L1C Sentinel-2 data processing levels. In bands 3 and 4, the L1C values were larger than the L2A values. In bands 6 and 8, the opposite pattern was seen, with L2A values being lower than L1C values. Atmospheric effects in the dataset were thus corrected for, as can be seen in band 3-4 and band 6-8 reflectance. This shows the relationship between data from two different processing levels.

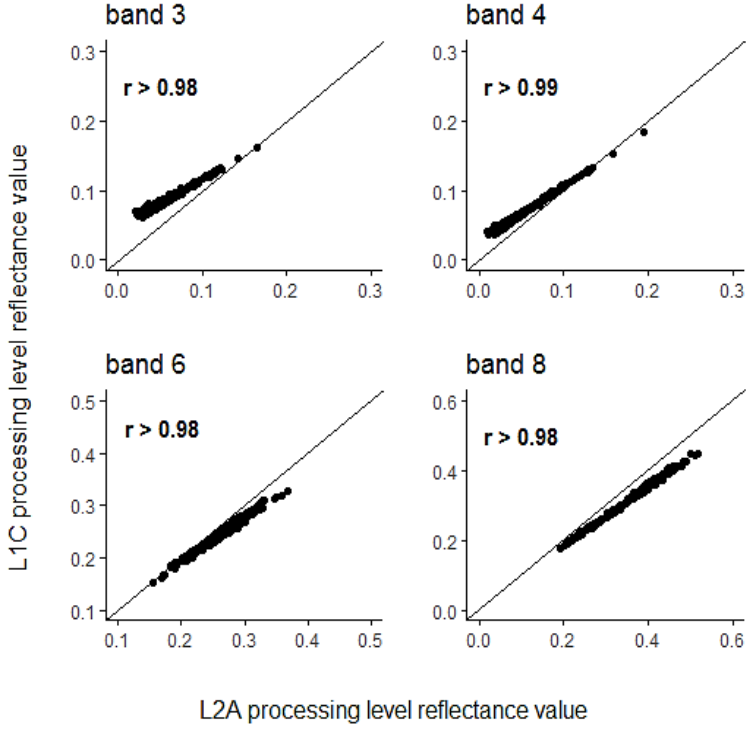


Figure 9. Band correlation between processing levels: level 1C and level 2A.

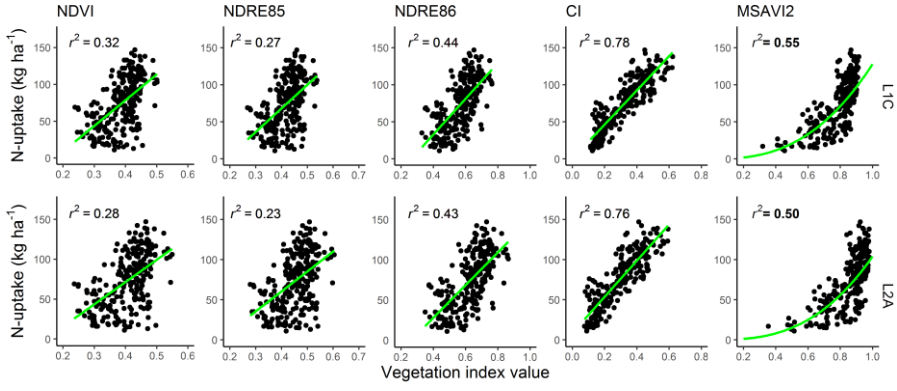


Figure 10. Comparison of model prediction of nitrogen uptake by winter wheat (of five indices) at two different Sentinel-2 data processing levels (L1C, L2A). For explanation of vegetation indices, see Table 6.

For the same dataset (described in more detail in Paper I), models that were calculated using data in different processing levels were similar. In all index-based models, L2A underperformed sligthly L1C processing levels in DC 22-53 when using goodness of fit (R^2). Visual interpretation suggests most difference in NDVI and NDRE85. This is not clearly related to a specific band or bands.

Five vegetation indices were modelled to predict N-uptake (Figure 10). The CI-index performed well in comparison to other vegetation indices, that were also calculated in Paper I. A linear prediction model based on the red-edge chlorophyll index (CI) was able to predict N-uptake (L1C data: R^2 = 0.74, mean absolute error; MAE = 14 kg ha^{-1}). Although the normalised difference indices tended to non-linearity, this tendency was unclear, therefore these data were also approached with a linear model.

4.5 Upscaling a prediction model (Paper I)

The availability of Sentinel-2 data, made it possible to update N-uptake models for implementation in CropSAT, which was previously not functioning for Sentinel-2 data. Upscaling of results shows that it is possible to pinpoint fields where VRA of N will likely be most useful. Model application was examined in a study area (the size of two satellite tiles in a wheat-dense area) other than the area in which models were calibrated. This gave the opportunity to showcase small regional differences in N-uptake as predicted by this model (Figure 11). The estimated field variation increased slightly (mean 30-41 kg ha^{-1}), when subsetting for field size within the inter-percentile range of 2.5-97.5 %. Standard deviation (SD) differed slightly, from 12 to 13 kg ha^{-1}.

As shown in Figure 11, within-field variation could be high in fields larger than 5 ha in the selected case study area at growth stage DC37. This implies that information for fields where VRA has most potential could be selected by upscaling the N-uptake model. This information can be used in DSS, as guiding information in farm management of VRA. In the 4169 fields studied in Paper I, the average variation in N-uptake was 90 kg ha^{-1} (SD = 20 kg ha^{-1}) (Figure 12).

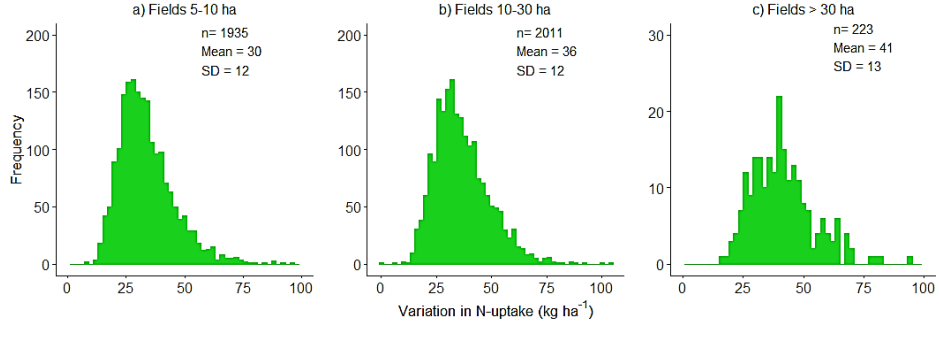


Figure 11. Range of within-field variation (2.5-97.5% inter-percentile) in nitrogen (N) uptake for winter wheat (kg ha^{-1}) at growth stage DC37 for field sizes (a) between 5 and 10 ha, (b) between 10 and 30 ha and (c) 30 ha or more, in southern Sweden.

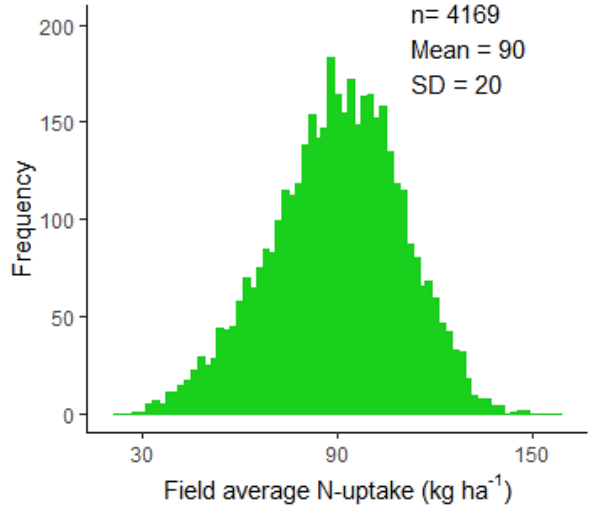


Figure 12. Average nitrogen (N)-uptake (kg ha^{-1}) in fields > 5 ha in the case study area decribed in Paper I.

4.6 Multi-source data (Paper I & II)

4.6.1 Index model VS. vehicle measured N-uptake (Paper I)

Nitrogen uptake based on tractor scanning with a N-sensor showed much similarity with N-uptake predicted by the CI-based model from satellite data at DC37. Within-field variation for a selection of fields in the study area used in Paper I showed similarities, which could be easily identified by visual comparison (Figure 13). Although the maps produced by CI-based modelling show higher N-uptake, within-field high and low values were in similar location.

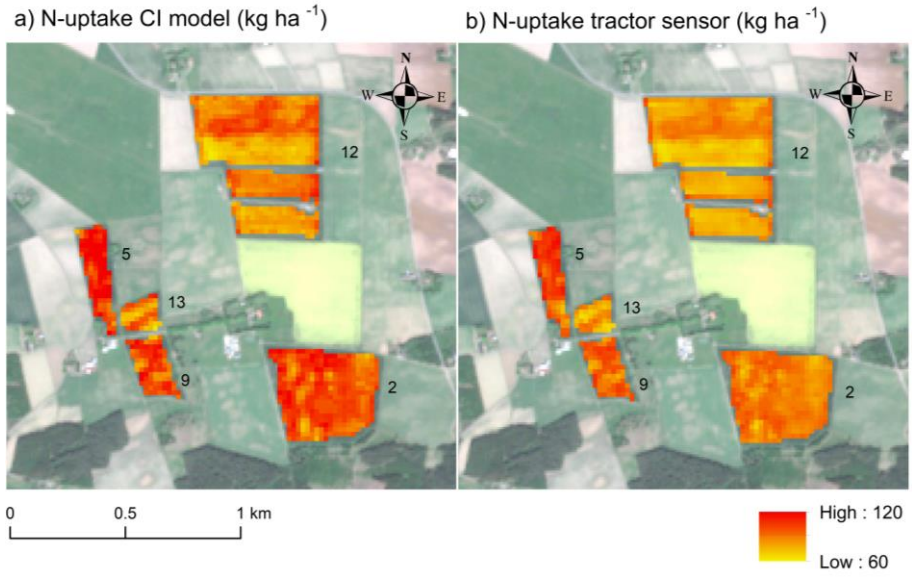


Figure 13. Nitrogen (N) uptake (high to low in kg ha^{-1}) as (left) predicted by a model based on satellite data and red-edge chlorophyll index (CI) and (right) measured by a tractor-mounted N-sensor.

4.6.2 UAV model transfer (Paper II)

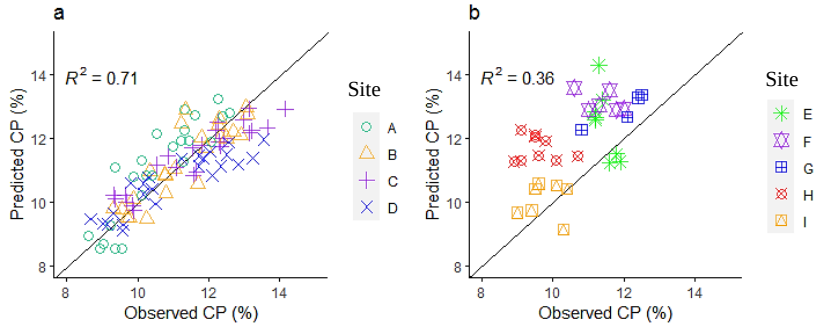


Figure 14. Observed and predicted crude protein content (CP %) in winter wheat on different sites (denoted with letters A-I), with 1:1 line. Results from transfer of a UAV-based univariate linear chlorophyll index (CI) model (a) before and (b) after model transfer to the satellite dataset.

In an attempt to transfer a UAV model to satellite scale, a field trial-based protein prediction model was used to test transfer of two models, a univariate regression model (Figure 14) and a MARS model using multiple indices. Transfer from the two types of UAV-based models to a small satellite dataset for the same region was not successful (univariate model transfer is shown in Figure 14). It appeared that the linear regression method overpredicted CP in the case of univariate model transfer (data above 1:1 line in Figure 14b). Transfer of the multivariate model was less successful than transfer of the univariate option with CI (for results from MARS modelling, see Paper II).

4.7 Predicting crop phenology (Paper III)

Historical field observations of crop development (for the period 2010-2019) are shown in Figure 15. The data distribution showed a high point density in the range between Zadoks stage DC 30-32, as highlighted by adding colour to the graph for point density, in Figure 15. Stages DC 30-32 were more frequently noted by the observers in the field, reasons for this could be an observation inaccuracy or stages in the range 30-32 take a longer time to complete. This difference in distribution possibly had implications for the results.

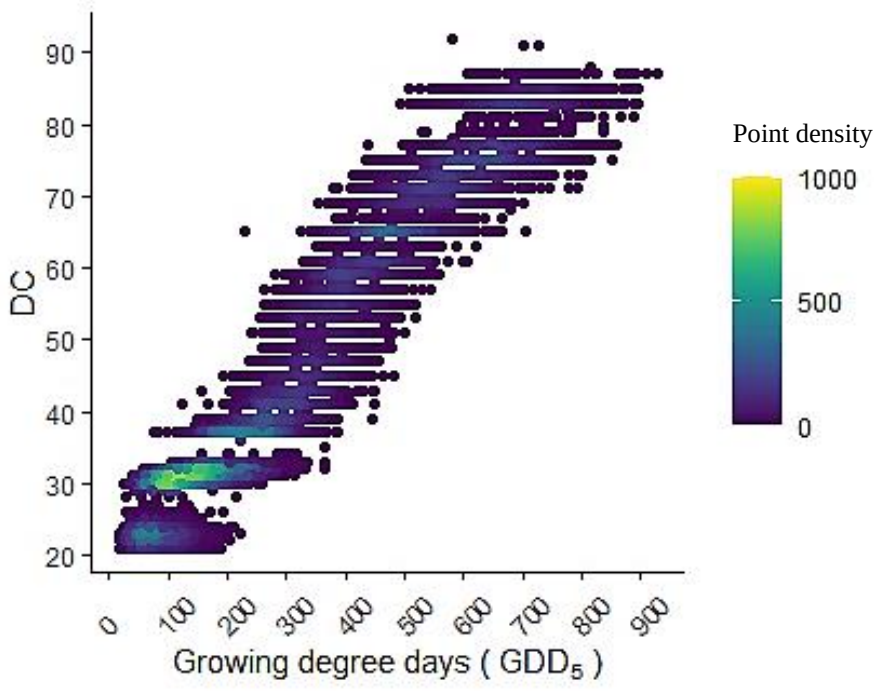


Figure 15. Distribution of field observations 2010-2019 for growing degree days (GDD) with a base temperature of 5 ˚C and Zadoks development stage (DC) in winter wheat. 2D point density distribution is indicated by the color scale.

Criteria for evaluating the possibility of crop phenology prediction were straightforward and statistical quantification was done in model error. The results showed that the large MESAN dataset could be used to predict Zadoks DC-growing stages from GDD (GDD_5 Gompertz model: RMSE = 4.08 DC, GDD_5 splines model RMSE = 4.64 DC) (Figure 15 & 16). Differences in model performance were not large (Table 10). The multivariate model, using DL, county and cultivar as additional variables showed performance was comparable to GDD_5 models, although this showed a lower RMSE than the splines univariate model (RMSE = 4.17 DC).

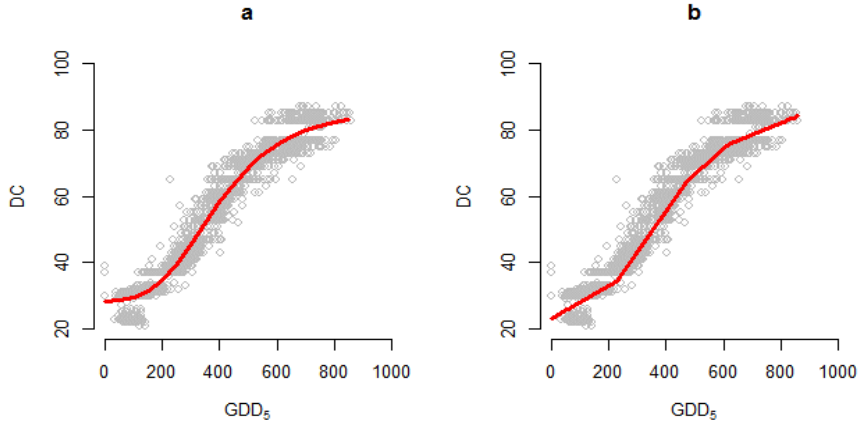


Figure 16. Prediction models, for 2019, for a growing degree days (GDD) calculation with T_{base} = 5 ˚C and Zadoks development stage (DC) with leave-one year-out (LOYO) validation for the univariate solutions. The Gompertz model result is shown in panel a and the splines model result in panel b.

In an effort to explore options for improving estimation of winter wheat phenology, initial testing of models was performed with a multitude of possible predictors, including photoperiod and county location. Small differences were found between thermal time models for two different base temperatures, the counties, the maturation rates and the multivariate model (MARS), as shown by differences in model RMSE (3.47-5.17 DC; Table 10). Differences between counties DC observations in raw data were also small (see Appendix A2).

Table 10. Summary statistics of the different univariate models and the multivariate model for different models. Multivariate model variables are: county, cultivar and accumulated daylength. Organised by modeltype with growing degree days (GDD). With root mean squared error (RMSE (in unrounded DC stages) and sample size (n).

	Gompertz		**Splines (MARS)**	
	RMSE	**n**	**RMSE**	**n**
Thermal Time				
GDD_0	5.17	26739	5.15	26739
GDD_5	4.08	26739	4.64	26739
County				
GDD_5 Skåne	3.96	7300	4.48	7300
GDD_5 Västra Götaland	3.47	3435	4.71	3435
GDD_5 Uppsala	3.58	2473	4.26	2473
Maturation rate				
GDD_5 Early	4.64	5381	4.65	5381
GDD_5 Medium	4.62	14992	4.57	14992
GDD_5 Late	4.60	4194	4.54	4194
Multivariate				
GDD_5 Multivar	-		4.17	26739

4.7.1 Modelling error of growth stage classes (Paper III)

Model error (RMSE) values for different expert-assigned DC classes are shown in Figure 17. For most models RMSE is small and similar. GDD_0 was predicted with higher RMSE than GDD_5 for both Gompertz and splines models (see black arrows). Some classes (indicated by solid-lined boxes in Figure 17) showed a trend for lower RMSE than the other classes. One of these, the DC class in the range DC 30-32, had a larger number of observations, which likely explains the lower RMSE in these classes. Another class after flowering (milk development) (DC 71-75) also had a slightly smaller RMSE result than surrounding classes. This may mean that the models performed well, with low (< 3) RMSE, for the most common period for supplementary N fertilisation in winter wheat in Sweden (around DC 30-39), provided that the models were not biased by the larger amount of observations around DC in the range 30-32 DC.

Regional Gompertz models resulted in a lower RMSE (indicated by the dash-lined box). This trend was not seen with splines modelling and is possibly due to tuning in the Gompertz fitting procedure.

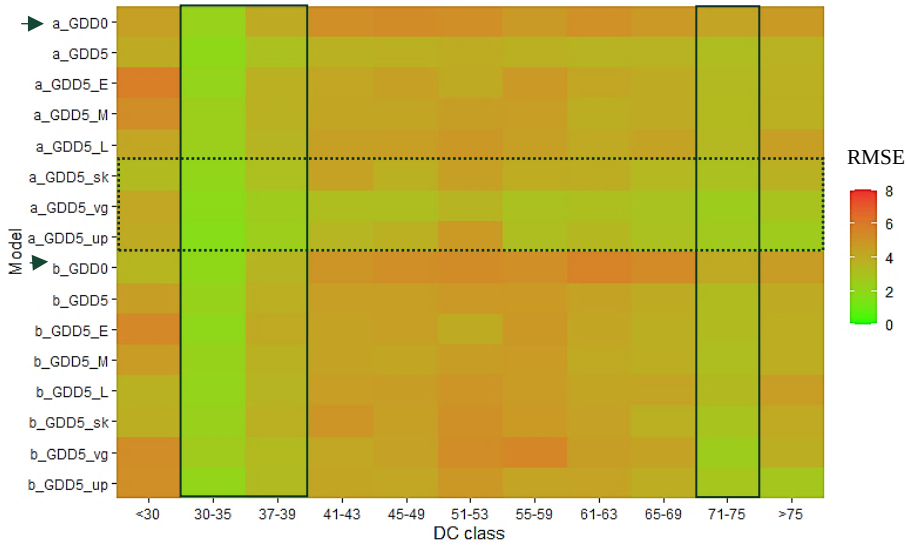


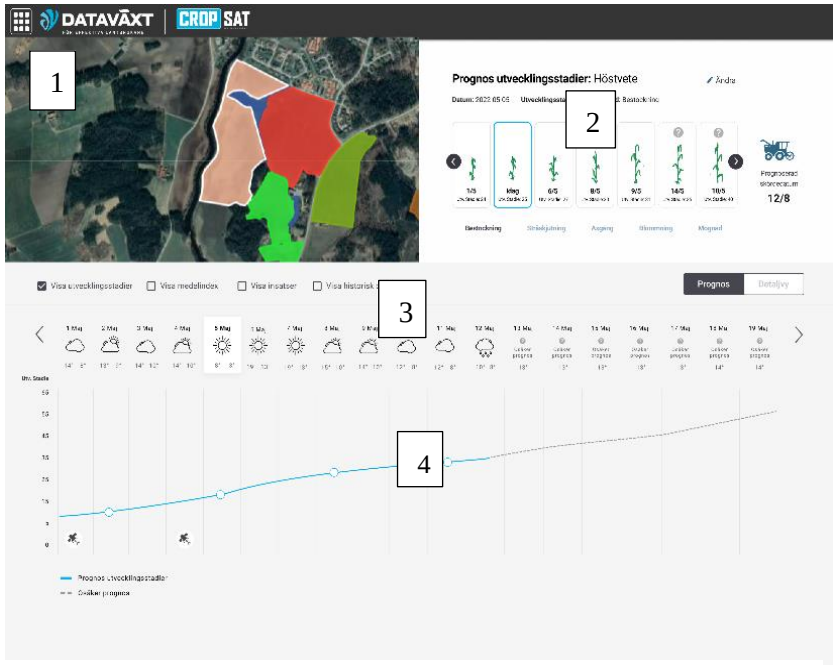
Figure 17. RMSE for univariate models Gompertz (a) and regression splines (b) per DC class. Early, medium and late maturation stages (subsequently indicated as: E, M, L). Regions Skåne, Västra Götaland and Uppsala (subsequently indicated as: sk, vg, up). Growing degree days (0 degree base GDD0, 5 degree base GDD5).

4.8 Implementation

The N-uptake result from Paper I as well as phenology model from Paper III have been, or will be, implemented in the CropSAT DSS.

Paper I resulted in a proximal vs. remote sensing algorithm that can be used for upscaling N-uptake estimation. The resulting model for N-uptake can be used to generate VRA-maps using Sentinel-2 data. Application users in CropSAT are prompted for a total N amount planned for application, which will be distributed per field, according the VRA model.

In addition, however, not yet implemented at this stage is the option to use the the upscaled result for quantification of VRA potential. The potential for VRA implementation can optionally be provided in CropSAT, giving insight in benefits of VRA beyond single fields. Local business developers working on the CropSAT DSS have started work on implementation of a Gompertz-based phenology model (based on the Paper III results). The crop phenology addition is not yet implemented in the N fertilisation module, however, a design has been proposed for an initial pilot implementation (Figure 18). After deliniation of fields by users (1), crop development (2), the calendar (3) and timeline (4) can be used to indicate timing of fertilisation. Even with recalibration of the curve using user input, the amount of input data that is required from users is low. The result will be connected to user fertilisation strategies in the application in the future.



1. Field display showing selected fields
2. A carousel of developmental stages
3. Calendar
4. Timeline with developmental curve

Figure 18. Design for upcoming implementation of a winter wheat phenology application in CropSAT. Courtesy of Dataväxt AB, Grästorp, Sweden.

5. Discussion

5.1 Progress in the N fertilisation complex

In preparation of index-based information for implementation, decisions need to be made on the type of models to implement and how to implement them, *e.g.* regional implementation or at larger area, for specific cultivars or for multiple cultivars, in specific growing stages or over a wider period, or in stressed or stress-free conditions (Figure 19). A next step in decision support for N fertilisation is conversion of N-uptake rates to N dosage recommendations. Making recommendations on N fertilisation is not straightforward. For example, when crop properties suggest potential N deficiency, this commonly means that the application rate can be increased in early growing stages, whereas when indices indicate a rising red-edge ratio, fertilisation rates should preferably go down (Heege 2013). Besides the steps taken in this thesis, continuation from N-uptake toward crop recommendations will in addition require a connection with economical optimal nitrogen rates (Basso *et al.* 2011; Piikki *et al.* (2022). VRA can require targetting more than one economical optimum and there remains a challenge. Current outcomes can be tested for different fertilisation strategies in future research. The VRA algorithm (Paper I), created from proximal- and remote sensing measurements, could improve existing decision support for N fertilisation in CropSAT. Implementation can be adjusted for specific cultivars, since small differences between cultivars were found for the data. Pre-harvest CP estimation UAV-based models (Paper II) are a possibility. However, since models were tested on a small dataset and transferability of pre-harvest CP UAV-based models to a satellite scale is currently not well tested, no implementation is planned. Had the outcome been more

successful, then pre-harvest predictions could be included in VRA for N fertilisation. Modelling of crop phenology from historical data was possible by different methods, which can potentially be implemented to follow strategic timing of VRA in DSS implementations. Besides the implementation of a crop phenology implementation, it is required to link these models in future work to fertilisation strategies (Table II), and test the outcome of such an implementation.

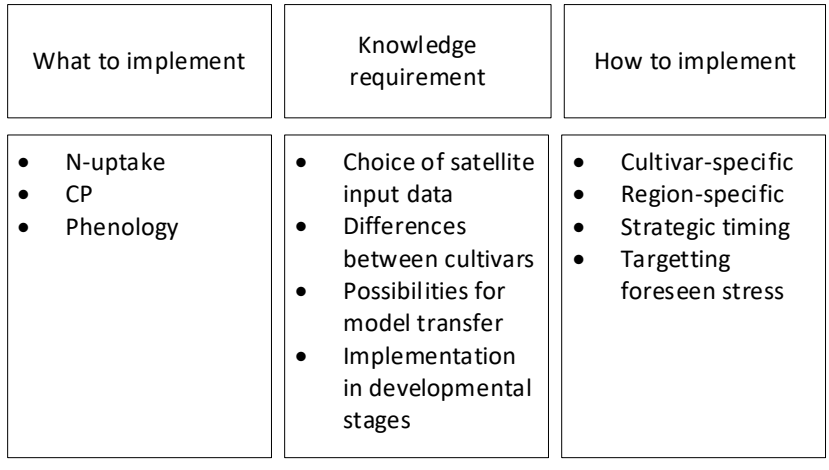


Figure 19. Example of considerations for model implementation.

5.2 Calibration for N-uptake

Hand-held sensor measurements are used throughout Sweden and reported on a weekly basis (Swedish board of agriculture 2022). In Paper I, these field data were used for upscaling to remote sensing scale, which is similar in approach to work presented in Söderström *et al.* (2017). The implicit assumption was made that hand-held sensor spectral data are sufficiently accurate in estimating within-field variation in N-uptake to serve as a proxy for field N-uptake. *In-situ* N-uptake sampling, analysed with laboratory plant tissue analysis gives, however, gives more accurate information on N-uptake (Prananto *et al.* 2020). Correlations between sensing instruments have to be strong when making predictions using multiple instruments. Considering accuracy in approximation of field N-uptake, *in-situ* tissue analysis would

have been a more appropriate choice. Using lab-analysis of field data is, however, costly and time-consuming. *In-situ* measurements were not available on the scale and in the amount as hand-held sensor measurements were available. Over 700 hand-held sensing measurements, distributed over the main winter wheat cropping areas in southern Sweden, were used. The scaled result from Paper I gave the possibility to pinpoint specific locations for VRA, thereby showcasing application potential. This information could be used in a DSS application for quantification of expected gain from VRA, resulting potentially in improved adoption rates.

5.3 Field trials for CP prediction

A multivariate approach testing vegetation indices and information of zero- and/or maximum N applications was used in Paper II. Two methods were tested, linear models and MARS modelling. The dataset was small and this limited possibilities for analysis. In Paper II, this resulted in MARS models not optimising, this could for example be seen from the inconsistency of models for different sites (optionally: see Figure 4 in Paper II). Study design could have been improved with access to a more extensive dataset. Vegetation indices derived from field trial data, with rising fertilisation levels (0-320 kg ha^{-1}) applied to the trial blocks, were used. Using field trial data with distinct and wide-ranging N rates for wheat protein estimation resulted in N fertiliser levels having a larger range than is usually the case in a field situation. This means that protein values based on such data are unrepresentative of the true range in N concentrations.

Another complication with using field trial data for pre-harvest CP predictions, is that trials are commonly placed where soil is homogeneous. Protein variation, however, could be better understood when soil variation is taken into account (*e.g.* Petterson 2007). Although several of the vegetation indices tested showed a relationship with winter wheat CP, auxiliary information on soil moisture content would have been useful (or possibly essential) to understand current findings in more detail.

Further exploration of the relationships between individual indices and protein would likely have resulted in a better understanding of potential use of individual indices in a DSS to predict protein variation (see Appendix A3). The two normalised difference indices (NDVI and NDRE86) showed a very different relationship with protein, indicating the importance of wavelengths

around the red-edge area of the electromagnetic spectrum (*e.g.* band 6 in the MAIA MSI). A mainly non-linear relationship between several vegetation indices and protein can be seen.

Alternative data

Assessment of N and protein via NIR based indices is useful, however, also short-wave infrared (SWIR)-based indices have proven to be succesful in assessment of these properties (Berger *et al.* 2020; Ferwerda *et al.* 2005; Herrmann *et al.* 2010). The SWIR wavelengths have been of value in *e.g.* Börjesson *et al.* (2019), confirming that water content of the crops is important for protein predictions. Sentinel-2 data would have been an opportunity, as its payload produces data in three SWIR bands and two of these (11&12) can potentially be useful together with *in-situ* sampling (ESA 2022). A comparison could have been made for improving our understanding of use of UAV-based models (Paper II).

According to Basnet *et al.* (2020), there are similarities between barley and wheat protein in terms of spectral response to N fertilisation. In addition, Petterson (2007) showed that vegetation indices can be used for predicting CP in barley and proposed a similar method for winter wheat. Comparison of results with results for other crops, such as malting barley, could have improved understanding of differences in canopy structure when using reflectance data. Unfortunately, this was outside of the scope of this thesis. It could, in addition, have been interesting to focus on reflectance in combination with potential physio-developmental traits, such as crop height (considering a potential relationship between protein and biomass/yield). The relationship between grain yield and grain protein is bi-directional, depending on available moisture. When water is limiting, grain yield decreases and grain protein content increases with N fertilisation, to produce a negative relationship (Terman 1979). In contrast, when N is limiting, protein and yield both increase with applied N, to produce a positive relationship. In a recent study, Zhou *et al.* (2021) demonstrated that crop height can be estimated using structure from motion, which means estimating three-dimensional structures from two-dimensional image sequences (Holman *et al.* 2016) and by applying machine learning methods afterwards prediction of protein can improve, thereby increasing understanding on the relationship between yield and protein.

5.4 Transfer from UAV- to satellite data scale

Testing of the UAV-based models to fill satellite timeline information gaps (induced by cloud cover, for example) on one site is potentially interesting. Latini *et al.* (2021) showed an encouraging result when spectral model transfer was conducted on the study site, results from two instruments were comparable. In the model application (Paper II), when tested on a new site it was seen that direct model transfer was more complicated than comparison on the same site. However, the data did not allow the option to control DC (nor the cultivar) in the fields where model transfer was tested. The comparability of MAIA-S2 and Sentinel-2 MSIs for possible model transfer therefore requires further testing. After an extensive literature review, Berger *et al.* (2020) rightfully note, that transfer of indices-based models to different sites can give unforeseen results. Yet, there was little evendence available at the time of the beginning of this work to support or reject this claim. The tests that were done in Paper II could therefore contribute to the understanding of the difficulties in simple model transfer. With the benefit of hindsight, model transfer could have been reported for more univariate linear regression models using a single vegetation index, such as TCOS or NDRE75 with absolute correlation $r > 0.76$ (Appendix A3), although results are not substantially different from current findings for this dataset.

5.5 Approaching phenology

Paper III demonstrates a GDD approach to crop phenology estimation using weather data and two different modelling approaches. Jamieson *et al.* (1995) have shown how during stem elongation in stress-free conditions developmental rate is approached with solely temperature data and independent of daylength. In the case of crop stress, which gives a sparser and warmer canopy, temperature is likely less good in predicting development. Near surface temperature, rather than air temperature was found more strongly related to phenology in earlier research (Jamieson 1995). There is an ongoing research debate on how to model the relationship between observed phenology, temperature and environmental variables, but developing a larger framework of different approaches could be a way forward (Boote 2019).

For DSS implementation near soil temperature data were not widely available. The concept of GDD during history has been expanded multiple

times in order to better capture crop phenology and recently in Butler and Huybers (2015) even a concept called killing degree days was introduced to describe the damaging effects of extreme heat stress on crops. While this concept may seem less relevant for Sweden now, extreme heat in spring in summer has occurred in Sweden in, for example, the year 2018. In such future scenario's a model using only GDD will not suffice.

Using satellite data for crop phenology estimation can be an alternative option for satellite-based DSS. With direct practical implementation in mind, the satellite-based approach does bring practical challenges. Gao & Zhang (2021), for example, describe how while in principle trend/curve based approaches to satellite data do the job, advancing to real-time estimation is, difficult, due to remote sensing data not always being sensitive to crop phenology. Detection using satellite data depends on the frequency and availability of cloud-free products, an approach using more easily available weather data was preferred with implementation in a DSS in mind. A satellite approach was planned for a future study and considered beyond the scope of this thesis. It would be interesting to compare a satellite-based approach with current weather-based options in future work.

While it can be a disadvantage to rely on user input in DSS applications, user input could be an additional luxury to a phenology implementation. Using simple models that rely on a limited amount of input data gives the opportunity to recalibrate our model implementation with user input, as demonstrated in the carousel in Figure 18 (panel 3). The effectiveness of this method is requires testing in practice after implementation is fully finalized.

5.6 Research for direct DSS implementation

In the previous sections, a number of situations were described where direct implementation has played a role in the conceptualization and choice of method. Multispectral datasets were the main datatype for DSS development in CropSAT. Common (and likely more common) in precision agriculture today is the use of data with higher spectral resolution and a larger spectral range (Mulla 2013). Higher-spectral resolution data require a good data infrastructure and flexible and data-driven methods for analysis, which are now widely available (Lu *et al.* 2019; Verrelst *et al.* 2015). The advent of space-borne hyperspectral sensors provides new options for improving understanding N fertilisation, for potential possibilties in future DSS

implementations. In an attempt to find possibilities for additional testing with data in higher spectral resolution via hyperspectral data I found that the available field sampling points had a very near miss with Prisma data (PRecursore IperSpettrale della Missione Applicativa, Agenzia Spaziale Italiana). As flexible statistical methods place higher demands on data-storage and transfer, preferably via open data in standardized form via *e.g.* datacubes (Tagliabue *et al.* 2022), this is currently not considered feasible for real-time application. Data-driven approaches, on the other hand, could have increased understanding for models that are currently implemented in CropSAT.

5.7 Using fast vs. elaborate algorithms

Fast crop production functions derived via statistical analysis, which simplify the biological or physical principles involved, may be inferior to elaborate models (Jame & Cutforth 1996). Research intended for DSS implementation requires balancing between fast- and more elaborate algorithms. While some approaches for current implementations are aiming for comprehension, their simplicity can make them biased. Elaborate approaches can be more information-dense and accurate, however, difficult to implement in practice. This can be a barrier to adoption of precision agriculture. In DSS for precision agriculture there is benefit in working towards reaching most potential out of input data, using few different types of input data and making choices for data as input that require little or no user input. This makes DSS applications easy to use and therefore could have a positive effect on adoption rates in precision agriculture.

At last, an important issue, but one that is scarcely reported, is the opportunity for greater adoption of VRA approaches by fast algorithms. Fast implementation of models could become a source of calibration data (when ethically feasible), as shown by Silva *et al.* (2020) for data collected via DSS, in an interesting example of possibilities for extending data generation opportunities in DSS development. The interaction between research and practice can become a cyclic process, where each benefits the other, resulting in faster advances from data-driven development approaches.

6. Conclusions and future outlook

General conclusions

- Recommendations in this thesis could contribute to advances in DSS development. Using prediction models in DSS applications is a way to direct farm management towards future efficiency, although more research is required to reach sufficient knowledge for advancing use of multi-source data.
- A close relationship between research for DSS development and parallel technical advances (mainly in the area of remote sensing) is required to keep up with, very rapidly emerging, new and updated technologies in this research field. Advancements in development of DSS could have a positive effect on this development, via practical testing.

Upscaling an N-uptake model

- Differences between the Sentinel-2 L1C (TOA) and L2A (BOA) datasets are currently small and in this thesis no improvement in N estimation was found for corrected products.
- Upscaling of field sensor data to satellite-data scale gave insight in VRA potential beyond single fields.
- It is recommended to include *in situ* data in continued N-uptake modelling.

Protein modelling

- Multispectral data can be used for predicting pre-harvest CP in winter wheat, although more data are required for a more complete understanding of the possibilities for prediction of pre-harvest CP for implementation.

Multi-source data

- Use of multi-source data is currently difficult to achieve in practice, but investigation of options should continue. Direct transfer of prediction models from UAV- to satellite scale requires further testing.

Timing of field operations

- Two different approaches were able to predict phenology in winter wheat. Future comparison of the tested GDD method with existing crop growth models and/or satellite data could be of interest in future research.
- Further translation of crop phenology prediction results for management of fertilisation timing is required and possibly a division between stressed or stress-free conditions could be included.

Applicability

- Models for N-uptake and phenology developed in this thesis are applicable in practice. There are potentially more future implementation options using results from this thesis, such as implementation of VRA beyond fields and connection of crop phenology predictions to the N fertilisation VRA model.

References

Barnes, E.M., Clarke, T.R., Richards, S.E., Colaizzi, P.D., Haberland, J., Kostrzewski, M. *et al.* (2000). Coincident detection of crop water stress, nitrogen status and canopy density using ground based multispectral data. In Robert, P.C., Rust, R. H., Larson, W.E., Proceedings of the 5th international conference on precision agriculture, Precision Agriculture Center, University of Minnesota, 16-19.

Basso, B., Ritchie, J., Cammarano, D., Sartori, L. (2011). A strategic and tactical management approach to select optimal N fertilizer rates for wheat in a spatially variable field. *European journal of agronomy*, 35, 215-222. doi:10.1016/j.eja.2011.06.004

Bastos, M., Froes de Borja Reis, A., Sharda, A., Wright, Y., Ciampitti, I.A. (2021). Current status and future opportunities for grain protein prediction using on- and off-combine sensors: A synthesis-analysis of the literature. *Remote sensing*, 13, 5027. doi:10.3390/rs13245027

Berger, K., Verrelst, J., Féret, J., Wang, Z., Wocher, M., Strathmann, M., Danner, M., Mauser, W., Hank, T. (2020). Crop nitrogen monitoring: recent progress and principal developments in the context of imaging spectroscopy missions. *Remote sensing of environment*, 242. doi:10.1016/j.rse.2020.111758

Bingham, J. (1967). Breeding cereals for improved yielding capacity. *Annals of applied biology*, 59, 312-315. doi:10.1111/j.1744-7348.1967

Bongiovanni, R., Lowenberg-DeBoer, J. (2004). Precision agriculture and sustainability, *Precision agriculture*, 5, Kluwer Academic Publishers, 359-387.

Boehmke, B., Greenwell, B. (2020). Hands-on machine learning with R, Retrieved from: www.bradleyboehmke.github.io

Boote, K. (2019). Advances in crop modelling for a sustainable agriculture. Taylor & Francis, London, UK. doi: 10.1201/9780429266591

Burrough P.A., McDonnell R.A. (1989). Principles of geographical information systems. Oxford University Press, Oxford, UK.

Butler, E.E., Huybers, P. (2015). Variations in the sensitivity of US maize yield to extreme temperatures by region and growth phase. *Environmental research letters*. doi:10.1088/1748-9326/10/3/034009

Börjesson, T., Wolters, S., Söderström, M. (2019). Satellite-based modelling of protein content in winter wheat and malting barley. In Stafford, J., *Precision agriculture*, Proceedings of the 12th European conference on precision

agriculture, Wageningen Academic Publishers, Wageningen, the Netherlands, 581-587.

Carpenter, S.R., Caraco, N.F., Correll, D.F., Howarth, R.W., Sharpley, A.N., Smith, V.H. (1998). Nonpoint pollution of surface waters with phosphorus and nitrogen, *Ecological applications*, 8, 559-568.

Carr, P.M., Carlson, G.R. Jacobsen, J.S. Nielsen, G.A., Skogley. E.O. (1991). Farming soils, not fields: a strategy for increasing fertilizer profitability. *Journal of production agriculture*, 4, 57-61.

Clevers, J.G.P.W. (1986). Application of remote sensing to agricultural field trials. Proefschrift, Wageningen University.

Crauford P.Q., Wheeler T.R. (2009). Climate change and the flowering time of annual crops. *Journal of experimental botany*, 60, 2529-2539.

CropSAT (2022). Retrieved [09-05-2022] from: www.cropsat.se

Delin, S., Lindén, B. (2002). Relations between net nitrogen mineralization and soil characeteristics within an arable field. *Acta agriculturae scandinavica, section B: Soil and plant science*. 52, 78-85.

Delin, S., Stenberg, M. (2014). Effect of nitrogen fertilization on nitrate leaching in relation to grain yield response on loamy sand in Sweden. *European journal of agronomy*, 52, 291-296.

Diacono, M., Rubino, P., Montemurro, F. (2013). Precision nitrogen management of wheat: a review. *Agronomy for sustainable deveevelopment*, 33, 219-241. doi:10.1007/s13593-012-0111-z

Drusch, M., Del Bello, U., Carlier, S., Colin, O., Fernandez, V., Gascon, F., Hoersch, B., Isola, C., Laberinti, P., Martimort, P., Meygret, A., Spoto, F., Sy, O., Marchese, F., Bargellini, P. (2012). Sentinel-2: ESA's optical high-resolution mission for GMES operational Services. *Remote sensing of environment*, 120, 25-36. doi: 10.1016/j.rse.2011.11.026

Efretuei, A., Gooding, M., White, E., Spink, J., Hackett, R. (2016). Effect of nitrogen fertilizer application timing on nitrogen use efficiency and grain yield of winter wheat in Ireland. *Irish journal of agricultural and food research*, 55, 63-73.

ESA (European Space Agency) (2022). Retrieved [in the period between 01-01-2022 to 16-06-2022] from: www.esa.int

Erikson, L. (2022). Bekämpningsrekommendationer: Svampar och insekter 2022. Retrieved from: www.jordbruksverket.se/bekampningsrek (in Swedish)

Feekes, W. (1941). De tarwe en haar milieu: Wheat and its environment. Verslagen van de technische tarwe commissie. 17, 523-888. (in Dutch and English)

Ferwerda, J.G., Skidmore, A.K., Mutanga, O. (2005). Nitrogen detection with hyperspectral normalized ratio indices across multiple plant species. *International journal of remote sensing*, 26, 4083-4095.

Fowler (2018). Winter wheat production manual, Ducks unlimited Canada and conservation production systems.

Freeman, K.W., Raun, W.R., Johnson, GV., Mullen, R.W., Stone, M.L., Solie J.B. (2003). Late-season prediction of wheat grain yield and grain protein. *Communications in soil science and plant analysis*, 34, 1837-1852. doi:10.1081/CSS-120023219

Gao, F., Zhang, X. (2021). Mapping crop phenology in near real-time using satellite remote sensing: challenges and opportunities. *Journal of remote sensing*. doi:10.34133/2021/8379391

Garner, W.W., Allard, H.A. (1920). Effect of the relative length of day and night and other factors of the environment on growth and reproduction in plants. *Monthly weather review*, 48, 415-415.

Gastal, F., Lemaire, G. (2002). N uptake and distribution in crops: an agronomical and ecophysiological perspective. *Journal of experimental botany*, 53, 789-799. doi:10.1093/jexbot/53.370.789

Gates, D.M., Keegan, H.J., Schleter, J.C., Weidner, V.R. (1965). Spectral properties of plants. *Applied optics*, Georgetown institute of technology press, Washington, D.C., USA, 11-20.

Gebbers, R., Tavakoli, H., Herbst, R. (2013). Crop sensor readings in winter wheat as affected by nitrogen and water supply. Precision agriculture '13, 9th European conference on precision agriculture, Wageningen academic publishers, Wageningen, the Netherlands.

Giese, A.C. (Ed.) (2013). Photophysiology: general principles. Action of light on plants, Academic press, New-York, USA.

Gitelson, A., Gritz, Y., Merzlyak, M. (2003). Relationships between leaf chlorophyll content and spectral reflectance and algorithms for non-destructive chlorophyll assessment in higher plant leaves. *Journal of plant physiology*, 160, 271-282. doi:10.1078/0176-1617-00887

Gooding, M.J., Davies, W.P. (1997). Wheat production and utilization. *Systems, quality and the environment*, Cambridge, UK.

Guo, B.B., Zhu, Y.J., Feng, W., He, L., Wu, Y.P., Zhou, Y., Ren, X.X., Ma, Y. (2018). Remotely estimating aerial N uptake in winter wheat using red-edge area index from multi-angular hyperspectral data. *Frontiers in plant science*, 9, 675. doi:10.3389/fpls.2018.00675

Haboudane, D., Miller, J., Tremblay, N., Zarco-Tejada, P., Dextraze, L. (2002). Integrated narrow-band vegetation indices for prediction of crop chlorophyll content for application to precision agriculture. *Remote Sensing of environment*, 81, 416-426. doi:10.1016/S0034-4257(02)00018-4

Hastie, T., Tibshirani, R., Friedman, J.H. (2009). The elements of statistical learning: data mining, inference, and prediction. 2nd ed., Springer, New York

Hatfield, J.L., Gitelson, A.A., Schepers, J.S., Walthall, C.L. (2008). Application of spectral remote sensing for agronomic decisions. *Agronomy Journal*, 100, 117-131. doi:10.2134/agronj2006.0370c

Henrich, V., Brüser, K. (2022). Retrieved [18-05-2022] from: www.indexdatabase.de

Heege, H.J. (ed.) (2013). Precision in crop farming: site specific concepts and sensing. Springer, Dordrecht, the Netherlands. doi:10.1007/978-94-007-6760-7_14

Herrmann, I., Karnieli, A., Bonfil, D.J., Cohen, Y., Alchanatis, V. (2010). SWIR-based spectral indices for assessing nitrogen content in potato fields. *International journal of remote sensing*, 31, 5127-5143.

Holman, F., Riche, A., Michalski, A., Castle, M., Wooster, M., Hawkesford, M. (2016). High throughput field phenotyping of wheat plant height and growth rate in field plot trials using UAV based remote sensing. *Remote Sensing*, 8, 1031. doi:10.3390/rs8121031

Huete, A.R., Jackson, R.D. (1988). Soil and atmosphere influences on the spectra of partial canopies. *Remote Sensing of Environment*, 25, 89-105. doi:10.1016/0034-4257(88)90043-0

Hāggmark, L., Ivarsson, K., Gollvik, S., Olofsson, P. (2003). MESAN, an operational mesoscale analysis system. *Tellus A.*, 52, 2-20. doi:10.1034/j.1600-0870.2000.520102.x

Jamieson, P., Brooking, I., Porter, J., Wilson, D. (1995). Prediction of leaf appearance in wheat: a question of temperature. *Field crops research*, 41, 35-44. doi:10.1016/0378-4290(94)00102-I.

Fountas, S., Søren, M., Pedersen, S.M. Blackmore, S. (2004). ICT in precision agriculture: diffusion of technology an overview of precision agriculture. Sciences, New York, 1-15. doi:10.13140/2.1.1586.5606.

Friedman, J. (1991). Multivariate adaptive regression splines, *Annals of Statistics*, 19, 1, 1-67.

IoF (Internet of food and farm) (2019). Horizon 2020. Retrieved [09-05-2022] from: www.iof2020.eu

International society of precision agriculture (ISPA) (2022). www.ispag.org

Jame, Y. Cutforth, H. (1996). Crop growth models for decision support systems. *Canadian journal of plant science*, 76. doi:10.4141/cjps96-003.

Knipling, E.B. (1970). Physical and physiological basis for the reflectance of visible and near-infrared radiation from vegetation. *Remote sensing of environment*, 155-159.

Latini, D., Petracca, I., Schiavon, G., Niro, F., Casadio, S., Del Frate, F. (2021). UAV-Based observations for surface BRDF characterization, IEEE International geoscience and remote sensing symposium IGARSS, 8193-8196. doi:10.1109/IGARSS47720.2021.9554496

Lancashire, P.D., Bleiholder, H., Langeluddecke, P., Stauss, R., van den Boom, T., Weber, E., Witzen-Berger A. (1991). A uniform decimal code for growth stages of crops and weeds. *Annals of applied biololgy*, 119, 561–601. doi:10.1111/j.1744-7348.1991.tb04895

Lemaire, G., Gastal, F. (1997). N uptake and distribution in plant canopies. In Lemaire, G. (eds.) Diagnosis of the nitrogen status in crops, Springer, Berlin, Heidelberg. doi:10.1007/978-3-642-60684-7_1

Lowenberg-DeBoer, J., Erickson, B. (2019). How does European adoption of precision agriculture compare to worldwide trends? In Precision agriculture ’19 Wageningen Academic Publishers, Wageningen, the Netherlands, 7-20.

Lu, B., He Y., Dao, P.D. (2019). Comparing the performance of multispectral and hyperspectral images for estimating vegetation properties, *Journal of selected topics in applied earth observations and remote sensing*, 12, 1784-1797. doi:10.1109/JSTARS.2019.2910558

Lu, N., Wang, W., Zhang, Q., Li, D., Yao, X., Tian, Y., Zhu, Y., Cao, W., Baret, F., Liu, S., Cheng, T. (2019). Estimation of nitrogen nutrition status in winter wheat from unmanned aerial vehicle based multi-angular multispectral imagery, *Frontiers in plant science*, 10. doi:10.3389/fpls.2019.01601

Mcbratney, A., Whelan, B., Ancev, T., Bouma, J. (2005). Future directions of precision agriculture, *Precision agriculture*, 6. doi:10.1007/s11119-005-0681-8

Myers, V.I., Allen, W.A. (1968). Electrooptical remote sensing methods as nondestructive testing and measuring techniques in agriculture. 7, 1819-1838. doi:10.1364/AO.7.001819

Mulla, D. (2013). Twenty five years of remote sensing in precision agriculture: key advances and remaining knowledge gaps. *Biosystems engineering*, 114, 358-371. doi:10.1016/j.biosystemseng.2012.08.009

Mulla, D,, Khosla. R. (2015). Historical evolution and recent advances in precision farming. Chapter 1. In (Lal, R. and Stewart, B.A. eds.) Soil-specific farming: precision agriculture, Advances in soil science, CRC Press, Florida, USA.

Nash, J.E., Sutcliffe, J.V. (1970). River flow forecasting through conceptual models part I: a discussion of principles. *Journal of hydrology*, 10, 282-290.

Nowak, B. (2021). Precision agriculture: where do we stand? A review of the adoption of precision agriculture technologies on field crops farms in developed countries. *Agricultural research*, 10, 515-522. doi:10.1007/s40003-021-00539-x

Petterson, C.G. (2007). Predicting malting barley protein concentration, based on canopy reflectance and site characteristics. Doctoral thesis no: 2007:56, Sveriges lantbruksuniversitet

Pierce, F.J., Nowak, P. (1999). Aspects of precision agriculture. *Advances in agronomy*, 67, 1-85. doi:10.1016/S0065-2113(08)60513-1

Piikki, K., Söderström, M., Stadig, H., Wolters, S. (2022). Remote sensing and on-farm experiments for determining in-season nitrogen rates in winter wheat: options for implementation, model accuracy and remaining challenges (*in review*)

Prananto, J., Minasny, B., Weaver, T. (2020). Chapter 1: near infrared (NIR) spectroscopy as a rapid and cost-effective method for nutrient analysis of plant leaf tissues. *Advances in agronomy*, Academic press, 164, 1-49. doi:10.1016/bs.agron.2020.06.001

Raun, W., Solie, J., Johnson, G., Stone, M., Mullen, R., Freeman, K., Thomason, W., Lukina, E. (2002). Improving nitrogen use efficiency in cereal grain production with optical sensing and variable rate application. *Agronomy journal*, 94. doi:10.2134/agronj2002.0815

Reyniers, M., Vrindts, E., de Baerdemaeker, J. (2006). Comparison of an aerial-based system and an on the ground continuous measuring device to predict yield of winter wheat. *European journal of agronomy*, 24, 87-94. doi:10.1016/j.eja.2005.05.002

Reusch, S. (1997). Entwicklung eines reflexionsoptischen sensors zur erfassung der stickstoffsversorgung landwirtsschaftlicher kulturpflantzen. forschungsbericht agrartechnik (VDI-MEG) 303, Dissertation, Kiel, Germany, 157. (in German)

Reusch S. (2005). Optimum waveband selection for determining the nitrogen uptake in winter wheat by active remote sensing. In Stafford, J., Werner, A. (eds.). Precision agriculture '05, Wageningen Academic publishers, Wageningen, the Netherlands, 261-266.

Robert, P.C. (1982). Evaluation of some remote sensing techniques for soil and crop management. PhD dissertation, University of Minnesota

Robert, P.C. (2002). Precision agriculture: a challenge for crop nutrition management. *Plant and soil*, 247, 143-149.

Robertson, P.G., Vitousek, P.M.. (2009). Nitrogen in agriculture: balancing the cost of an essential resource. *Annual review of environment and resources*, 34, 97-125.

Rouse, J.W., Haas, R.H., Schell, J.A., Deering, D.W. (1973). Monitoring vegetation systems in the great plains with ERTS. 3rd ERTS Symposium, NASA SP-351, Washington DC, 309-317.

Salgó, A., Gergely, S., Juhász, R. (2005). Characterising the maturation and germination processes in wheat by NIR methods. In Woodhead publishing series in food science, technology and nutrition. Woodhead Publishing, 212-219. doi:10.1533/9781845690632.6.212

Sarvia, F., De Petris, S., Orusa, T., Borgogno Mondino, E. (2021). MAIA S2 versus Sentinel 2: spectral issues and their effects in the precision farming context. Computational science and its applications (ICCSA), 21st international conference, Cagliari, Italy. doi:10.1007/978-3-030-87007-2_5

Sawyer, J.E. (1994). Concepts of variable rate technology with considerations for fertilizer application. *Journal of production agriculture*, 7, 195-201.

SCB (2022). Retrieved [05-09-2022] from: www.scb.se

Schils, R., Olesen, J.E., Kersebaum, K., Rijk, B., Oberforster, M., Kalyada, V. *et al.* (2018). Cereal yield gaps across Europe. *European journal of agronomy*, 101, 109-120. doi:10.1016/2018.09.003

Silva, J.V., Tenreiro, T.R., Spätjens, L., Anten, N.P.R., van Ittersum, M.K., Reidsma, P. (2020). Can big data explain yield variability and water productivity in intensive cropping systems? *Field crops research*, 255. doi:10.1016/j.fcr.2020.107828

Sinclair, T.R., Rufty, T.W. (2012). Nitrogen and water resources commonly limit crop yield increases, not necessarily plant genetics. *Global food security*, 1, 94-98. doi:10.1016/j.gfs.2012.07.001

Simmonds, N.W. (1996). Yields of cereal grain and protein. *Experimental agriculture*, 32, 351-356.

Spearman, C. (1904). The proof and measurement of association between two things. *American journal of psychology*, 15, 72-101.

Stafford, J.V. (2000). Implementing precision agriculture in the 21st century, *Journal of agricultural engineering research*, 76, 267-275.

Stern, K.R., Bidlack, J.E., Jansky, S. (2010). Stern's introductory plant biology, McGraw-Hill, USA

Swedish board of agriculture (Jordbruksverket) (2022). Retrieved from: www.jordbruksverket.se

Söderström, M., Piikki, K., Stenberg, M, Stadig, H., Martinsson, J. (2017). Predicting nitrogen uptake 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 and plant science*, 67, 637-650.

Tatham, A.S., Shewry, P.R. (1985). The conformation of wheat gluten proteins. The secondary structures and thermal stabilities of α-, β-, γ- and ω-gliadins. *Journal of cereal science*, 3, 103-113. doi:10.1016/S0733-5210(85)80021-7

Taylor, J.A. *et al.* (2021). Applications of optical sensing of crop health and vigour. In: Kerry, R., Escolà, A. (eds.) Sensing approaches for precision agriculture. progress in precision agriculture. Springer, 333-367. doi:10.1007/978-3-030-78431-7_12

Terman, G.L. (1979). Yields and protein content of wheat grain, as affected by cultivar, N, and environmental growth factors. *Agronomy journal*, 71, 437-440.

Thenkabail, P.S., Lyon, J.G., Huete, A. (2019). Fifty years of advances in remote sensing of agriculture and vegetation: summary, insights and highlights of volume IV'. In: Advanced applications in remote sensing of agricultural crops and natural vegetation. Fourth edition, volume four of series hyperspectral remote sensing of vegetation. CRC press, USA, 339-378.

Triboi, E., Abad, A., Michelena, A., Lloveras, J., Ollier, J. L., Daniel, C. (2000). Environmental effects on the quality of two wheat genotypes: 1. quantitative and qualitative variation of storage proteins. *European journal of agronomy*, 13, 47-64. doi:10.1016/S1161-0301(00)00059-9

Uthayakumaran, S., Wrigley, C. (2007). Chapter 5 wheat: grain-quality characteristics and management of quality requirements, woodhead publishing series in food science, technology and nutrition, cereal grains (2nd. ed.), 91-134. doi:10.1016/B978-0-08-100719-8.00005-X

Verrelst, J., Camps-Valls, G., Muñoz-Marí, J., Pablo Rivera, J., Veroustraete, F., Clevers, J.G.P.W., Moreno, J. (2015). Optical remote sensing and the retrieval of terrestrial vegetation bio-geophysical properties: a review, ISPRS, *Journal of photogrammetry and remote sensing*, 108, 273-290. doi:10.1016/j.isprsjprs.2015.05.005

Wang, Y., Zhang, Z., Feng, L., Du, Q., Runge, T. (2020). Combining multi-source data and machine learning approaches to predict winter wheat yield in the conterminous United States. *Remote sensing*, 12, 1232. doi:10.3390/rs12081232

Went, F.W. (1953). The effect of temperature on plant growth. *Annual review of plant physiology*, 4, 347-362.

Witzenberger, A., Hack, H., van den Boom, T. (1989). Erläuterungen zum BBCH-dezimal-code für die entwicklungsstadien des getreides. *Gesunde Pflanzen*, 41, 384-388. (in German)

Wolters, S., Söderström, M., Piikki, K., Stenberg, M. (2019). Near-real time winter wheat N-uptake from a combination of proximal and remote optical measurements: how to refne Sentinel-2 satellite images for use in a precision agriculture decision support system. In: Staford, J. (Ed.), Precision agriculture, *Precision agriculture* (2021) proceedings of the 12th European conference on precision agriculture, Wageningen Academic Publishers, Wageningen, the Netherlands, 1001-1007. doi:10.3920/978-90-8686-888-9

Wuest, S.B., Cassman, K.G. (1992). Fertilizer-nitrogen use efficiency of irrigated wheat II: Partitioning efficiency of preplant versus late-season application. *Agronomy journal*, 84, 689-694.

Zadoks, J.C., Chang T.T., Konzak. C.F. (1974). A decimal code for growth stages of cereals, *Weed research*, 14, 415-421.

Zhang, N., Wang, M., Wang, N. (2002). Precision agriculture: A worldwide overview, *Computers and electronics in agriculture*, 36, 113-132. doi:10.1016/S0168-1699(02)00096-0

Zhou, X., Kono, Y., Win, A., Matsui, T., Takashi S., Tanaka T. (2021). Predicting within-field variability in grain yield and protein content of winter wheat using UAV-based multispectral imagery and machine learning approaches, *Plant production science*, 24, 137-151. doi:10.1080/1343943X.2020.1819165

Popular science summary

Winter wheat nitrogen (N) fertilisation is often done in a fixed rate per field. Alternatively, the application of N fertilizer on a field can be done in a way that the rate of fertilizer application is adjusted to the requirements of crops in smaller fields sub-parts. Variable rate fertilization is a more precise, and therefore more efficient alternative.

Decision support systems collect and analyse different types of data (for example satellite data) for advising farmers in practise accessible via a website. In order to instruct fertiliser spreading machinery, maps (called taskmaps) are used that hold the information on the desired dosage of fertiliser for a field. Taskmaps can be generated from the support system to use on a fertiliser spreader.

In Paper I, a new model from satellite images helps to understand how much nitrogen has been taken up by the wheat crop. This can then be used on farms as support for N fertilisation decision making. This information can also be used to pinpoint fields where variable rate application is likely to bring the most benefit. Paper II looks into predicting the quality of winter wheat before harvest, which can help to make decisions on when and possibly where to harvest. The protein level present in harvested grain kernels is important for the quality of winter wheat products, such as bread. It was also tested whether knowledge from unmanned aerial vehicles (drones) could be used in satellite decision support systems. In the last Paper, the winter wheat crop development was predicted from weather information, this could help make existing fertilisation advice more accurate. Some of the outcomes of this work were implemented in a Swedish system called CropSAT.

Taken together, the different parts in this work have improved decision support in Sweden, however, there are also remaining challenges described.

Populärvetenskaplig sammanfattning

Vete är en av de viktigaste grödorna i världen. I Sverige sås höstvete på hösten och skördas i slutet av sommaren året därpå. För att det ska växa bra måste grödan ha tillgång till många olika näringsämnen, bland annat kväve. Gödsling med kväve till höstvete görs ofta med en jämn giva inom varje fält. Emellertid kan grödans behov av kväve variera inom ett fält, t ex beroende på varierande jordart. Därför är ett alternativ till en jämn giva, att variera tillförseln av kväve så att den anpassas efter behovet i olika delar av fältet. Varierad tilldelning, som är en central del inom det som kallas precisionsodling, leder till att kvävet kan utnyttjas mer effektivt.

Genom webbaserade beslutsstödsystem kan lantbrukare lättare utnyttja olika typer av data, till exempel data från satelliter, för att kunna anpassa åtgärder efter varierande behov. Gödningsspridare kan styras av digitala tilldelningskartor som kan innehålla information om hur mycket gödning som varje fältdel behöver. Beslutsstödsystem för precisionsodling kan generera sådana tilldelningskartor.

Den här avhandlingen består av tre vetenskapliga artiklar. I den första artikeln har en modell tagits fram som gör att man från satellitbilder kan uppskatta hur mycket kväve som tagits upp av en höstvetegröda mitt i växtsäsongen. Varje 20 x 20 m yta i alla höstvetefält i en satellitbild får ett beräknat värde, och informationen kan sedan användas som stöd till lantbrukare när man ska bestämma hur mycket ytterligare kväve som grödan behöver. Informationen kan också användas för att lokalisera fält där varierad tilldelning är mest effektiv. Den andra artikeln handlar om möjligheten att bedöma kvaliteten i höstvete innan skörd, något som skulle kunna vara användbart för en lantbrukare när man planerar skörden. Halten av protein i det skördade vetet är en viktig kvalitetsegenskap, och avgör till exempel om vetet är lämpligt att användas till bakning av bröd. I studien

undersöktes om data insamlade med hjälp av sensorer på drönare kunde användas för att beräkna proteinhalten i vete, och om denna information sedan kunde överföras till satellitbilder för användning i beslutsstödsystem. I den sista studien togs en modell fram för att man från väderdata kan beräkna höstvetets utvecklingsstadier. Olika åtgärder i fält, som till exempel kompletterande tilldelning av kväve, görs lämpligast när vetet nått vissa stadier. Därför kan en sådan modell vara användbar i beslutsstödsystem.

En del av resultaten i detta arbete har implementerats i ett fritt tillgängligt beslutsstödsystem som används i Sverige som heter CropSAT. Sammanfattningsvis så exemplifierar resultaten från de olika studierna hur olika typer av data kan användas för att förbättra funktionaliteten i beslutsstödsystem för precisionsodling, men de visar också att många utmaningar återstår, och att fortsatt forskning på området behövs.

Acknowledgements

Firstly, I like to thank my supervisors for advising, reviewing and all other work involved in this project:
My main supervisor Mats Söderström and co-supervisors Thomas Börjesson, Carl-Göran Pettersson and Heather Reese.

Kristin Piikki, while not an official supervisor, you played a large role in the work for this project.

I am also endlessly grateful for my colleagues in Skara:
Johanna Wetterlind, you gave me great advice during our many lunches.

Lena Engström, you are a fantastic chef and you showed me that I need to run more often to keep up with you.

Karin Andersson, it was always nice to talk to you. You gave me lovely tomato plants and made me feel very welcome on your "kräftskiva".

Omran Alshihabi, it was nice to "fika", have diner and drinks and exchange culture.

Evelin Axelsson, you helped me out with smart advice, many times. I hope we can play a boardgame soon, although the next event happens to coincide with my defense date (I'm sorry).

Annika Holm, you have been a great administrative support.

My colleagues: Karl Adler, Sofia Delin, Katarzyna Marzec-Schmidt and any others I have met in Skara during this period.

I am also very thankful for support from:
Karin Blombäck, you have shown outstanding management skill and have been a great support.

Björn Lindahl, Johan Stendahl and Lisbet Lewan, you did an outstanding job in advising and supporting the project, especially in the last phases.

Anna Mårtensson, Magnus Simonsson and Mats Larsbo, thank you for reviewing my thesis, this was of much help.

A special thank you for the Skara innebandy group, it was a pleasure to play this nice sport with you.

My colleagues at Dataväxt for the implementation of the work and many interesting meetings.

Thank you also to Mary McAfee for providing rapid and good language checks.

I am also very thankful for all personal support from a distance:
Former colleagues at Agro, thank you for sparking my interest in precision agriculture, software development and databases.

Dear Humans, Alex, your passing brought me your music the last months of my thesis work and with this many youth memories.

Stephen Donkor, your senior research insight was worth a lot.

In my private life I wish to show my enormous gratitude to:
Erik, you have been a greater support than I could possibly wish for.

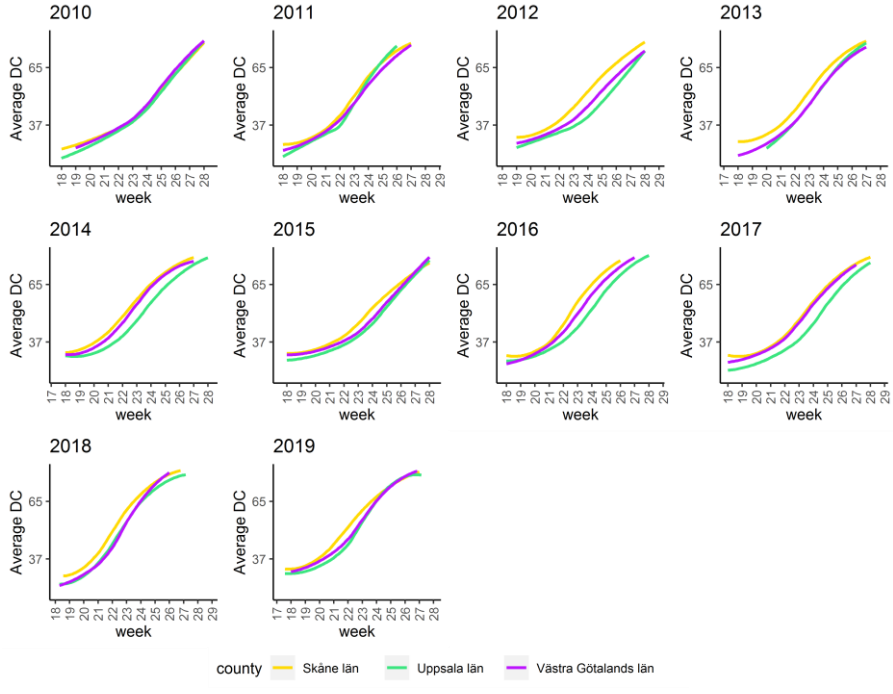
My dear friends: Aafke, Bart, Berry, Dick (not my father), Erwin, Floris, Freek, Henk, Jacco, Lianne, Maris, Michael, Rob, Robert, Sjors, Timo, Willem and the +ones.

My mother and father, Laura, Daniëlle and Vanessa.
Sweet hugs for Elijah, Hylke and Linde.

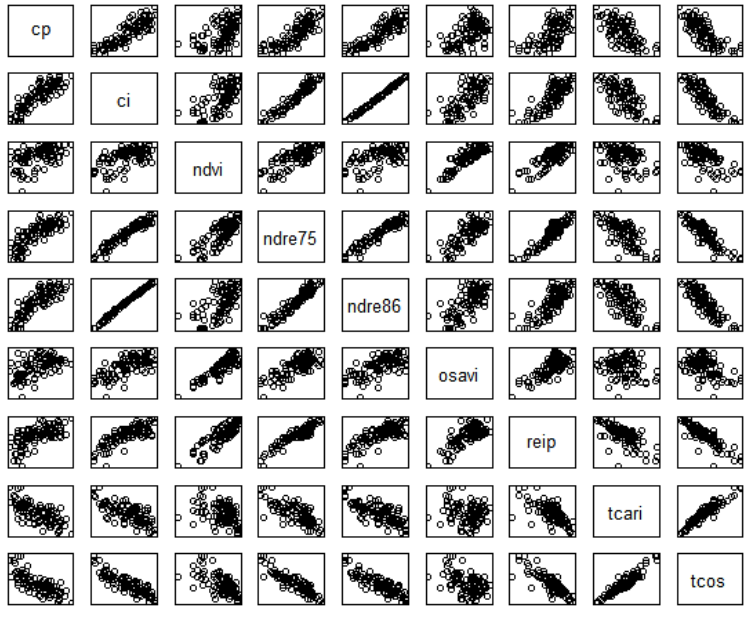
Appendix

Index	Formula
CI (red-edge)	(ρ7/ρ6)-1
MSAVI2	1/2 (((2 × ρ8 + 1) √((2 × ρ8 + 1)2 − 8 × (ρ8 − ρ4)))
NDRE75	(ρ7−ρ5)/ (ρ7+ρ5)
NDRE85	(ρ8−ρ5)/ (ρ8+ρ5)
NDRE86	(ρ8−ρ6)/ (ρ8+ρ6)
NDVI	(ρ8−ρ4)/ (ρ8+ρ4)
OSAVI	1+0.16((ρ7-ρ4)/((ρ7+ ρ4)+0.16))
REIP	700+40 (((ρ4-ρ7)/(2-ρ5))/(ρ6-ρ5))
TCOS	TCARI/OSAVI
TCARI	3((ρ5-ρ4)-0.2(ρ5-ρ3)ρ5/ρ4)

Appendix A1. Formula's for the vegetation indices used in the relevant Paper. Band (ρ) indicates the band number in the respective study where the index is used.



Appendix A2. Illustration of observed data in different regions in the period May 1 to August 15 for the 10 different years with applied local polynomial regression (LOESS) quadratic function.



Appendix A3. Intercorrelation for crude protein (CP %) and vegetation indices (Paper II) data; chlophyll index (CI), normalised-difference vegetation index (NDVI). Normalised-difference red-edge NDRE with alternative bands 8 and 6 (NDRE86), normalised difference red-edge (with bands 7 and 5 NDRE75), optimised soil adjusted vegetation index (OSAVI), red-edge inflection point (REIP), transformed chlorophyll absorption reflectance index (TCARI), and the TCARI/ OSAVI ratio (TCOS).

I

Precision Agriculture (2021) 22:1263–1283
https://doi.org/10.1007/s11119-020-09783-7

Upscaling proximal sensor N-uptake predictions in winter wheat (*Triticum aestivum* L.) with Sentinel-2 satellite data for use in a decision support system

S. Wolters[1] · M. Söderström[1] · K. Piikki[1] · H. Reese[2] · M. Stenberg[3]

Accepted: 30 December 2020 / Published online: 21 January 2021

Abstract
Total nitrogen (N) content in aboveground biomass (N-uptake) in winter wheat (*Triticum aestivum* L.) as measured in a national monitoring programme was scaled up to full spatial coverage using Sentinel-2 satellite data and implemented in a decision support system (DSS) for precision agriculture. Weekly field measurements of N-uptake had been carried out using a proximal canopy reflectance sensor (handheld Yara N-Sensor) during 2017 and 2018. Sentinel-2 satellite data from two processing levels (top-of-atmosphere reflectance, L1C, and bottom-of-atmosphere reflectance, L2A) were extracted and related to the proximal sensor data (n = 251). The utility of five vegetation indices for estimation of N-uptake was compared. A linear model based on the red-edge chlorophyll index (CI) provided the best N-uptake prediction (L1C data: $r^2 = 0.74$, mean absolute error; MAE = 14 kg ha^{-1}) when models were applied on independent sites and dates. Use of L2A data, rather than L1C, did not improve the prediction models. The CI-based prediction model was applied on all fields in an area with intensive winter wheat production. Statistics on N-uptake at the end of the stem elongation growth stage were calculated for 4169 winter wheat fields > 5 ha. Within-field variation in predicted N-uptake was > 30 kg N ha^{-1} in 62% of these fields. Predicted N-uptake was compared against N-uptake maps derived from tractor-borne Yara N-Sensor measurements in 13 fields (1.7–30 ha in size). The model based on satellite data generated similar information as the tractor-borne sensing data ($r^2 = 0.81$; MAE = 7 kg ha^{-1}), and can therefore be valuable in a DSS for variable-rate N application.

Keywords Decision support system · L2A · Nitrogen fertilisation · Precision agriculture · Sentinel-2 · Variable rate application

✉ S. Wolters
sandra.wolters@slu.se

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

[2] Department of Earth Sciences, University of Gothenburg, Gothenburg, Sweden

[3] Swedish Board of Agriculture, Skara, Sweden

Introduction

Winter wheat (*Triticum aestivum* L.) is an important crop globally, and is often the main crop in northern European cropping schemes. Much of the arable land in Sweden is dedicated to winter wheat production (19% of total cropping area and 48% of total grain production) (Swedish Board of Agriculture 2019a). Fertilisation with nitrogen (N) is often performed in two or three steps between the developmental stages of tillering and booting, to match crop requirements. Making decisions on the frequency, timing and quantity of this supplementary fertilisation can be a challenge for farmers, who try to find an economically optimum level in fertilisation. The estimation of the N concentration in aboveground plant tissues multiplied by above-ground dry matter mass (here denominated as N-uptake) during the period of supplementary fertilisation is an important component in the formulation of a fertilisation strategy (Schils et al. 2018). There are both economic and environmental benefits in optimising N fertilisation, since it optimises the quantity and quality of the crop in relation to the production costs and at the same time helps prevent losses of excess N through leaching or denitrification (e.g. Delin and Stenberg 2014; Swedish Board of Agriculture 2019a). Optimisation can be done on different scales, by fertilising different fields in different ways or by fertilising individual fields using a variable rate (variable rate application, VRA). It is known that growing conditions are often non-uniform between and within cropping fields (e.g. Sawyer 1994; Stafford 2000). Thus VRA of N fertiliser based on estimated N-uptake at the time of fertilisation is likely to better meet crop demands than uniform application. Variable application of N fertiliser can also be carried out for the purpose of reaching target levels of grain protein content, which is an important quality aspect (Basnet et al. 2003; Börjesson et al. 2019; Söderström et al. 2010).

Estimation of N-uptake in winter wheat can be done in multiple ways. Optical remote sensing is a method that has gained interest in recent decades, due to relatively easy and affordable application in the field (Berger et al. 2020; Mulla 2013; Zhao 2005). The N-status of crops can be estimated with optical sensing instruments by estimation of leaf chlorophyll concentration and biomass volume (Curran 1989; Kokaly 2001). Optical sensors measure canopy reflectance in different regions of the electromagnetic spectrum. Proximal crop canopy optical sensors have been available to farmers and advisors for many years (Reusch 2003; Singh 2019). These can be mounted on a vehicle, such as a tractor, or can be used as a handheld instrument. Limitations of this approach are that users rely on costly equipment and that collection of sensor data in the field can be time-consuming, making this method less feasible or attractive for some farmers.

The Swedish Board of Agriculture (Jordbruksverket, Jönköping, Sweden) provides weekly assessments to farmers on N-uptake in winter wheat based on measurements using the handheld version of the Yara N-Sensor® (Yara Gmbh, Hanninghof, Germany; as described by Reusch 2005). The measurements are taken at about 40 point locations in each year across the major agricultural regions in Sweden. They have been carried out to create up-to-date advice for farmers and advisors based on N-uptake measurements in unfertilised plots (N-uptake in unfertilised plots, so called zero-plots, is an indication of soil N-supply) during the period of supplementary fertilisation. Measurements were also made in an area of the field judged to be uniform and with no experimental manipulation (i.e. managed by the farmer as usual). The results are reported to farmers and advisors through an internet service (www.greppa.nu). This campaign generated a continuous time series of N-uptake, as measured by the N-sensor, but with no spatial coverage.

A development could be to scale this point information using satellite remote sensing to full spatial coverage using optical satellite imagery. The Copernicus Sentinel-2 mission comprises a constellation of two polar-orbiting satellites (2A, 2B) placed in a sun-synchronous orbit, and phased 180° to each other (Fletcher 2012, Drusch 2012). The instrument has a swath width of 290 km and a revisit time of 5 days at the equator. At Swedish latitudes, the temporal frequency for obtaining a new satellite product is every 2–3 days. This temporal frequency may be sufficient for practical use in time-critical, within-season N-status monitoring in grain crops.

Sentinel-2 data are published in different levels of processing, as top-of-atmosphere (TOA) data (L1C) and bottom-of-atmosphere (BOA) data (L2A) (ESA 2020). L1C products have been disseminated by ESA since June 2015 and L2A products from May 2017. The L1C products are 100×100 km^2 tiles with radiometric and geometric corrections, including orthorectification and spatial registration (ESA 2020). L2A products are considered 'analysis-ready products' that should be ready for immediate analysis without the need for further processing, and thereby better estimate reflectance. The L2A correction algorithms used by ESA are based on a series of threshold tests that use as input TOA reflectance from various Sentinel-2 spectral bands and auxillary data, look-up tables derived from a radiative transfer model library and digital elevation model (DEM) data. The processing procedure encompasses six steps: (1) a scene classification procedure to identify clouds, their shadows and snow; (2) aerosol optical thickness (AOT) calculation using a dense dark vegetation algorithm; (3) usage of a differential absorption algorithm to retrieve water vapour (WV); (4) cirrus cloud correction; (5) surface reflectance retrieval; and (6) scene classification with a terrain correction using DEM data. Apart from the corrected result, outputs are an AOT map, a WV map and a scene classification map, together with quality indicators for cloud and snow probabilities (Gascon et al. 2017; Main-Knorn et al. 2017; Mueller-Wilm 2016).

To derive information on crop canopy health and vigour from multispectral satellite data and proximal crop sensors, vegetation indices are often used. Commonly used vegetation indices use bands in the near-infrared (NIR), red and red-edge regions of the electromagnetic spectrum. Red-edge indices in particular have been proposed for intensive crop management applications throughout the growing season (Gitelson et al. 2003; Reusch 2003, 2005; Söderström et al. 2017). Some red- and NIR- based indices lose sensitivity after reaching a threshold level of leaf coverage and/or chlorophyll content, whereas red-edge-based indices are still sensitive to variation in denser canopies (Barnes et al. 2000; Qi et al. 1994; Rouse et al. 1974). The short-wave infrared (SWIR) region has also been shown to be useful in assessment of N content in crops (e.g. Herrmann et al. 2010; Söderström et al. 2010), but many of the crop sensors currently in use lack bands in this region of the electromagnetic spectrum. With hyperspectral and multispectral sensor data, a limitation of the conventional vegetation indices is that only a limited portion of the available data is used. An alternative approach to indices is to apply more advanced multivariate or machine-learning methods to include all bands or, preferably, all relevant bands (Berger et al. 2020; Verrelst et al. 2015). Regardless of the approach used, it is important to properly validate the prediction models, and to make sure there is no overfitting. This is important for model deployment in a decision support system (DSS) for precision agriculture.

Crop canopy N-status information, in the form of vegetation index maps from satellite data, are already available to end-users (e.g. farmers and advisors) via an internet based DSS, such as CropSAT (www.cropsat.com) (Söderström et al. 2017). A DSS can function as a tool to translate reflectance based N-uptake data to N-rate maps (i.e. prescription maps) to be used for VRA of N fertilisers. Agricultural practitioners can

use a DSS to get instant access to inexpensive, yet timely and site-specific, decision support for precision N fertilisation.

Wolters et al. (2019) developed a prediction model based on Sentinel-2 L1C data for generation of N-uptake maps to be used in a DSS. That study demonstrated that a Sentinel-2 L1C-based model could be used to predict N-uptake with different vegetation indices for use in practical applications.

The aim in the present study was to test whether it is possible to scale up the weekly point measurements of the Yara N-Sensor to full spatial coverage of the arable land in southern Sweden, and make the result available in a DSS. Specific objectives were to:

- Develop N-uptake scaling models through scaling handheld proximal sensor data using Sentinel-2 data, and evaluate their performance at independent sites and dates;
- Evaluate differences in model performance between Sentinel-2 processing levels (L1C and L2A) and different vegetation indices;
- Quantify the spatial variation in N-uptake within and between fields, by applying the model in an area in Sweden with intensive winter wheat production; and
- Apply the best scaling model and evaluate the resulting satellite-based N-uptake maps by comparison with N-uptake maps from a commonly used tractor-borne sensor system.

The hypotheses are as follows:

- Handheld N-sensor measurements in winter wheat can be scaled up for use in a DSS for VRA of N.
- The scaled model performs better when based on Sentinel-2 L2A data than when based on L1C data.
- The scaled model produces N-uptake maps that are similar to maps from tractor-borne proximal reflectance sensors.

Materials and methods

Study area

The entire study area encompassed a large part of southern Sweden, from roughly 55° to 61° N and from 10° to 19° E (Fig. 1). The temperate climate makes this a suitable area for rain-fed grain production, e.g. the median winter wheat yield during 2015–2019 was 7400 kg ha^{-1} (Statistics Sweden, Örebro, Sweden; www.scb.se). Mean annual precipitation is around 700–1000 mm in the agricultural regions studied, with higher values in western parts, and mean annual temperature is 5–8 °C (Swedish Meteorological and Hydrological Institute, Norrköping, Sweden; www.smhi.se). The cropland (around 2×10^6 ha) in the region is mainly found on young lacustrine and marine sediments deposited after the Weichselian glaciation (Fredén 1994), with heavy clays in the northeast and mostly loam and sandy loam in the southwest (Piikki and Söderström 2019).

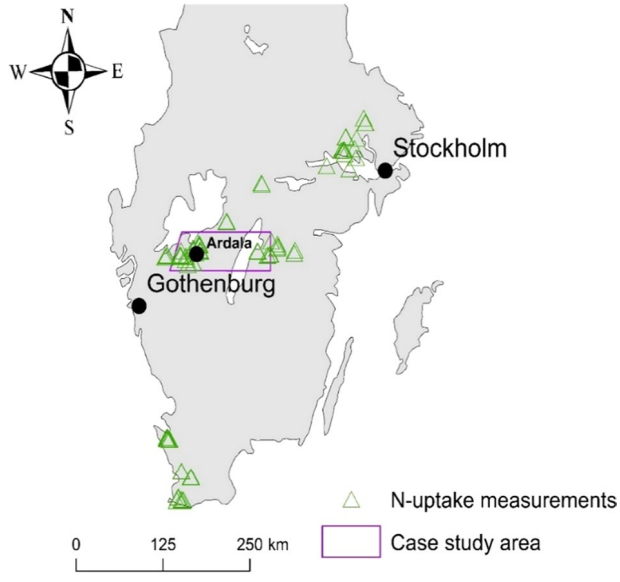


Fig. 1 The study area including the locations of proximal field data nitrogen (N)-uptake measurements in southern Sweden in winter wheat (*Triticum aestivum* L.). The area northeast of Gothenburg depicts the case study area in which the within-field variability of N-uptake was statistically assessed by deploying the satellite-based prediction model for all winter wheat fields. In the vicinity of the Ardala village, 13 winter wheat fields were scanned with a tractor-borne N-sensor for comparison with the Sentinel-2 based model

Handheld proximal sensing

The handheld version of the Yara N-Sensor was used by the Swedish Board of Agriculture to collect field data from 40 to 50 farms per year on N-uptake in winter wheat during growth stages DC22-53 (Zadok et al. 1974), in measurements conducted on a weekly basis from late April to early June 2017–2018. In this study, data was used from sensor measurements made in a uniform area of the winter wheat fields, in a part of the field that was managed by the farmer as usual (i.e. not the zero plots). Each proximal sensor measurement was the average of four recordings obtained in four directions, together representing an area of approximately 3 m × 3 m. The Yara N-Sensor is a sensor that records reflectance data in 45 bands 10 nm wide in the 400–900 nm region of the electromagnetic spectrum. Conversion to N-uptake is made through a built-in calibration function developed by Yara (Reusch 2003) and used by the Swedish Board of Agriculture in its monitoring programme. The selected farms cover the main winter wheat districts in Sweden (Fig. 1). Fourteen of the most commonly grown winter wheat cultivars in Sweden were included, seven from Lantmännen SW seed (Malmö, Sweden): 'Brons', 'Festival', 'Hereford', 'Julius', 'Linus', 'Norin', and 'Stava'; and seven from Scandinavian Seed (Lidköping, Sweden): 'Elvis', 'Frontal', 'Mariboss', 'Olivin', 'Praktik', 'RGT Reform' and 'Torp'.

Tractor-borne proximal sensing

Data on N-uptake were also gathered by tractor scanning in 13 fields (ranging in size from 2 to 30 ha) in the area around the village of Ardala (see Fig. 1) on 27–29 May 2017. A passive Yara N-Sensor® was used and the tractor was driven in a regular pattern over the fields on tramlines at 24 m spacing, recording N-uptake every second (with approximately 3 m between recordings). The N-uptake values from the sensor were

interpolated by ordinary block kriging (Burrough and McDonnell 1989) to the same 20 m × 20 m grid cell size as the Sentinel-2 data, to enable comparisons.

Satellite remote sensing

The Sentinel-2 satellites (2A and 2B) carry optical instrument payloads that sample 13 spectral bands with different spatial resolution: 10 m (2 [nominal blue], 3 [green], 4 [red] and 8 [broad-band NIR]); 20 m (5, 6, 7 [red-edge], 8A [narrow-band NIR], and 11, 12 [SWIR]); and 60 m (1 [coastal blue], 9 [NIR water vapour], and 10 [SWIR cirrus]) (ESA 2020). Details are given in Table 1.

Sentinel-2 data (both L1C and L2A) were downloaded from the Copernicus Open Access Hub (https://scihub.copernicus.eu/; ESA, EU, download period: 2018–2020). Sentinel-2 products were projected in WGS1984 UTM zone 33 N and organised by tiles following the military grid reference system (MGRS). After downloading, products were visually inspected for haze and clouds. Images that appeared cloud-free on field data points were paired (if within ± 3 days of acquisition) with the handheld proximal sensor data, and Sentinel-2 reflectance values were extracted (using nearest neighbour) from the pixel in which the field measurement was carried out. A total of 251 unique records had both handheld proximal data and remote sensing data. Some field measurements appeared to contain georeferencing errors (the points were outside the field) and were removed. The final dataset contained 242 records after exclusion of such records. Correlation of data from the L1C and L2A products was evaluated for four of the Sentinel-2 bands (3, 4, 6 and 8).

Table 1 Sentinel-2 satellite data bands with associated spatial resolution, central wavelength (λ) and bandwidth (Width) for sensor A (S2A) and sensor B (S2B) (ESA 2020)

Band	Spatial resolution	S2A		S2B	
	(m)	λ (nm)	Width (nm)	λ (nm)	Width (nm)
Band 1—Coastal aerosol	60	443.9	27	442.3	45
Band 2—Blue	10	496.6	98	492.1	98
Band 3—Green	10	560.0	45	559.0	46
Band 4—Red	10	664.5	38	665.0	39
Band 5—Vegetation red edge	20	703.9	19	703.8	20
Band 6—Vegetation red edge	20	740.2	18	739.1	18
Band 7—Vegetation red edge	20	782.5	28	779.7	28
Band 8—NIR	10	835.1	145	833.0	133
Band 8A—Narrow NIR	20	864.8	33	864.0	32
Band 9—Water vapour	60	945.0	26	943.2	27
Band 10—SWIR—Cirrus	60	1373.5	75	1376.9	76
Band 11—SWIR	20	1613.7	143	1610.4	141
Band 12—SWIR	10	2202.4	242	2185.7	238

Abbreviations in band descriptions are Near infrared (NIR) and shortwave infrared (SWIR)

Vegetation indices

To explore how different types of vegetation indices functioned in modelling, five different vegetation indices all based on bands within the spectral region of the proximal sensor were tested for this study. The vegetation indices were calculated from the Sentinel-2 data and were: normalised difference vegetation index (NDVI; Rouse et al. 1974); normalised difference red-edge vegetation index for two different band combinations, bands 8 and 5 (NDRE85), and bands 8 and 6 (NDRE86) (Barnes et al. 2000); modified soil-adjusted vegetation index (MSAVI2; Qi et al. 1994); and leaf chlorophyll index (CI; Gitelson et al. 2003). Two of these indices (NDVI and MSAVI2) are NIR-/red-based and both are already used in CropSAT (Söderström et al. 2017). The other three indices (NDRE85, NDRE86 and CI) use a red/red-edge band combination. Bands in the red-edge region have been shown to be useful in studies of N-uptake (Reusch 2003, 2005). Equations 1–5 show how the indices were calculated, where ρ is reflectance and the subscript indicates the Sentinel-2 band number:

$$NDVI = \frac{\rho_8 - \rho_4}{\rho_8 + \rho_4} \tag{1}$$

$$MSAVI2 = \frac{1}{2}\left[\left(2 \times \rho_8 + 1\right)\sqrt{\left[\left(2 \times \rho_8 + 1\right)^2 - 8 \times \left(\rho_8 - \rho_4\right)\right]}\right] \tag{2}$$

$$NDRE85 = \frac{\rho_8 - \rho_5}{\rho_8 + \rho_5} \tag{3}$$

$$NDRE86 = \frac{\rho_8 - \rho_6}{\rho_8 + \rho_6} \tag{4}$$

$$C = CI_{red-edge} = \left(\frac{\rho_7}{\rho_6}\right) - 1 \tag{5}$$

Modelling and validation

An overview of the data processing and analyses performed in the study is given in Fig. 2. Regression models for prediction of N-uptake were parameterised between the handheld proximal sensing N-uptake data and the vegetation indices derived from the Sentinel-2 data. To assess the prediction accuracy when applying the model to new sites and dates, a spatiotemporal cross-validation was designed and employed. This method of leave-one-out cross-validation is a model validation technique in which the records are repeatedly split into 'test data' (the record for which a prediction is made) and 'training data' (the records used to parameterise the model). With each iteration, one record was assigned to the test set and the remaining records were assigned to the training set. In order to validate the model in a manner which resembled a practical application in a DSS, any records from the same site as the test record and data collected later in the

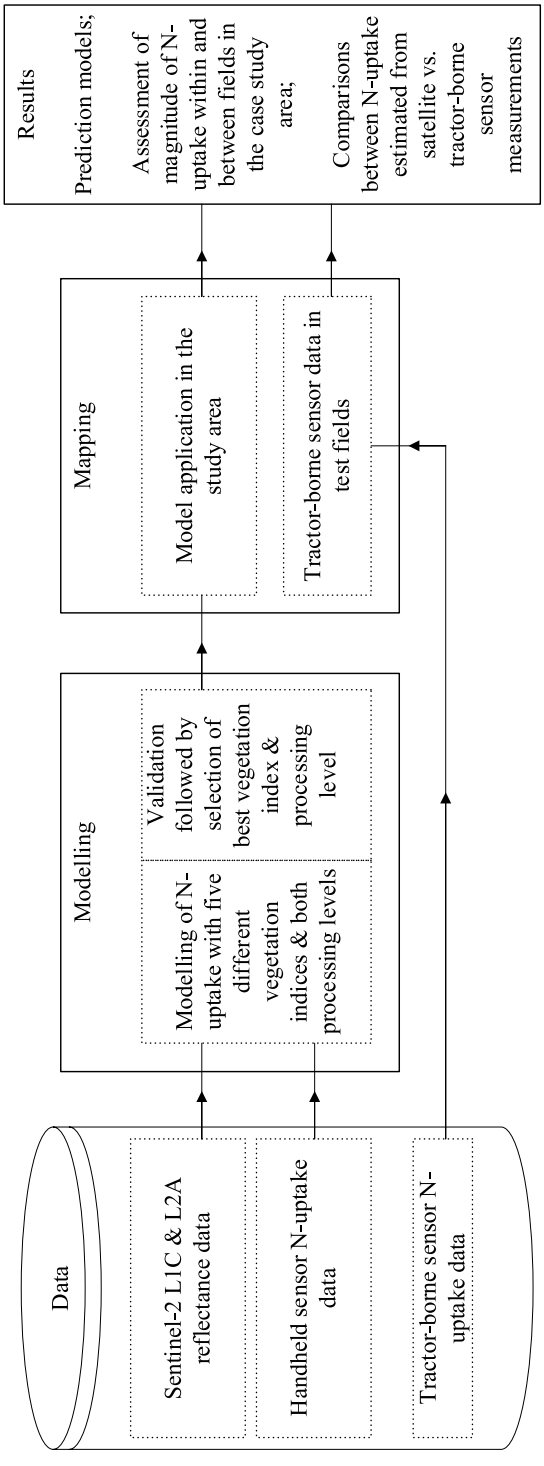


Fig. 2 Schematic overview of the study. Data are in top of atmosphere (L1C) and bottom of atmosphere (L2A) processing level

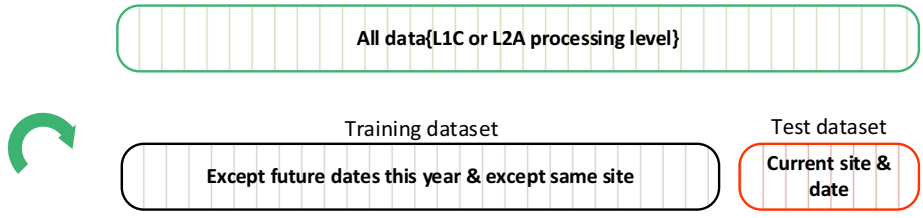


Fig. 3 Graphical display of the spatiotemporal leave-one-out cross-validation procedure used for regression predictions. A set of all data of either top of atmosphere (L1C) or bottom of atmosphere (L2A) processing level enters the procedure. For each record in the dataset, a model was trained leaving out that record and also leaving out other records from the same site and all records from later dates in the same year. The training set was used to calibrate prediction models and nitrogen (N)-uptake was predicted for the test record

same year (than the date of the test record) were also removed from the training dataset. This was repeated until all 242 records had been tested (Fig. 3).

To determine prediction accuracy, validation measures were calculated from the measured and predicted N-uptake values. The Nash–Sutcliffe modelling efficiency (E) can theoretically take values between $-\infty$ and 1, where an E value of 0 means that the model is just as accurate as a mean of the measured data and an E value of 1 is a perfect fit, which means that the predicted values are equal to the measured values (Nash and Sutcliffe 1970). The mean error (ME) is a measure of the overall prediction bias. The mean absolute error (MAE) is the average of the absolute prediction error. The coefficient of determination (r^2) explains the goodness-of-fit of the prediction.

To exemplify possible use of modelling results and to quantify the magnitude of within- and between-field variation in N-uptake, the best fitting prediction model was applied to predict N-uptake within a case study area (Fig. 1), covering 7045 km^2 with a high density of winter wheat fields (as determined by winter wheat fields reported in the EU agricultural subsidies system; Swedish Board of Agriculture 2019b). For this, cloud- and haze-free satellite data from 27 May 2017 were used (corresponding to approximately winter wheat growth stage DC37 in this area according to an online service by the Swedish Board of Agriculture; https://etjanst.sjv.se/). Modelled data were extracted for each field (which was reduced in size using a buffer of 15 m along field boundaries to avoid mixed pixel effects) and the magnitude of within-field variation in N-uptake was calculated for the inter-percentile range 2.5–97.5% (to exclude other potential outlier effects).

The satellite-based model predictions of N-uptake were compared with maps resulting from interpolation of the tractor-borne N-sensor measurements. This comparison consisted of two parts: (1) a visual comparison of spatial variation patterns and (2) field-wise correlation analyses between satellite-based and tractor-based values (r^2 and MAE).

Software

To determine which MGRS files matched the field measurement coordinates, the Python programming language was used (Python software foundation, Wilmington, Delaware, USA). Data were stored in a SQLserver database (Microsoft, Redmond, Washington, USA) and analyses were performed using the R programming language (R core team 2018). ArcGIS 10.7 (Esri Inc., Redlands, California, USA) was used for data analysis and display.

Results

Data exploration

The data from the L1C and L2A processing levels were found to be linearly correlated to each other in bands, 3, 4, 6 and 8 (Fig. 4). In bands 3 and 4, the L1C values were larger than the L2A values. In bands 6 and 8, the opposite pattern could be seen, where L2A values were smaller than L1C values. In all cases, Pearson correlation coefficient was higher than 98%, with the lowest correlation for band 3. Reflectance values in the individual bands for the two different processing levels were thus very similar in this dataset.

The vegetation indices studied correlated differently with N-uptake measured by the handheld proximal canopy reflectance sensor for both the L1C and L2A processing level datasets (Fig. 5). NDVI and the two different NDRE indices showed weak correlations, with $r^2 < 0.44$ for both L1C and L2A data. As can be seen from the diagram, the CI values were better correlated to the proximal sensor measurements (L1C: $r^2 = 0.78$; L2A: $r^2 = 0.76$). The MSAVI2 values were non-linearly correlated with the field data (L1C: $r^2 = 0.55$; L2A: $r^2 = 0.50$).

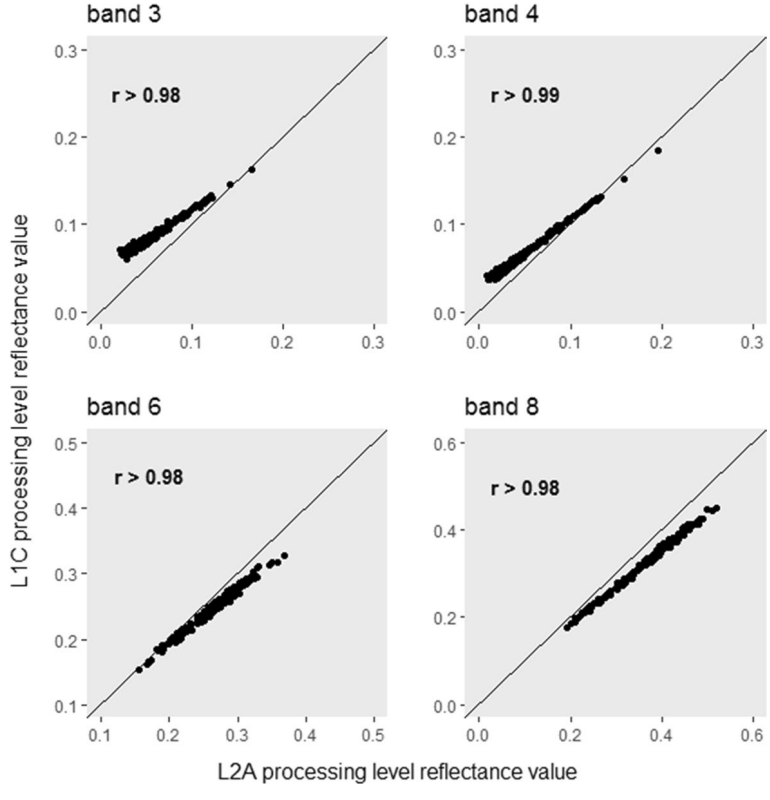


Fig. 4 Comparison of reflectance values in top of atmosphere (L1C) versus bottom of atmosphere (L2A) processing level products for bands 3, 4, 6 and 8

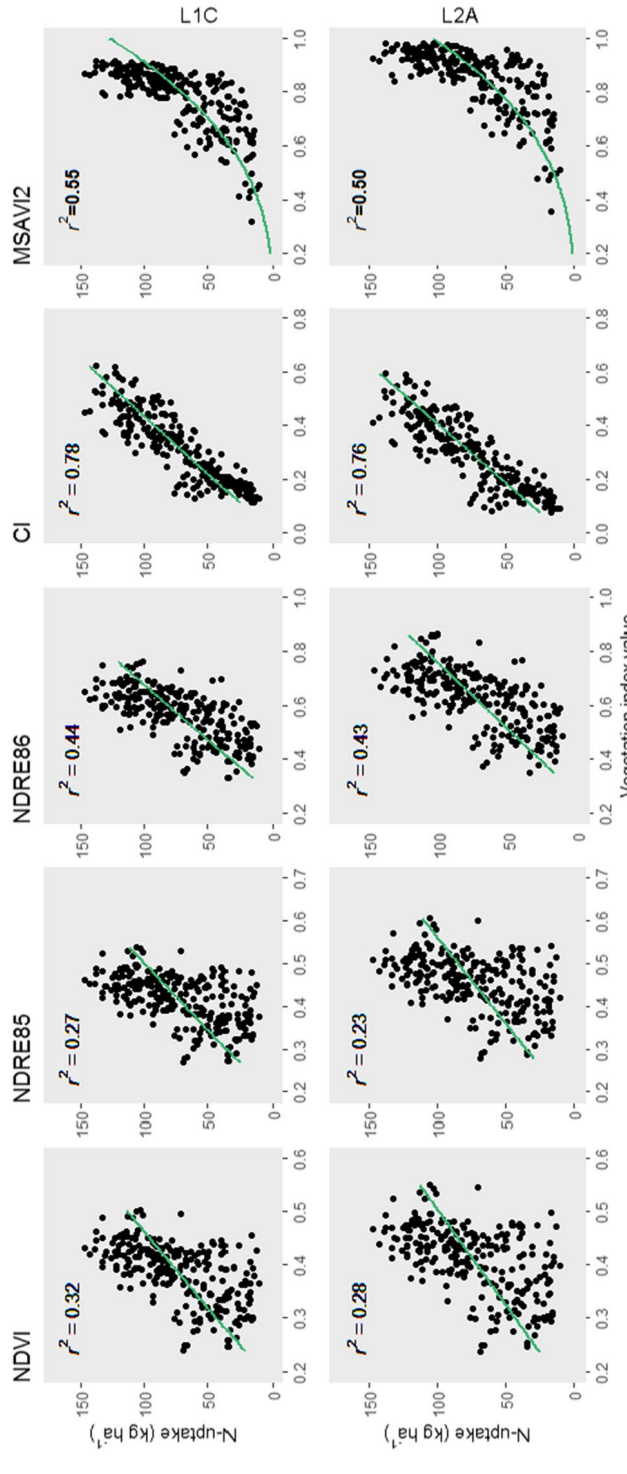


Fig. 5 Measured nitrogen (N)-uptake plotted against the five different vegetation indices tested for all 14 winter wheat cultivars in growth stages DC22-53 (Zadok et al. 1974). The horizontal axis are vegetation index values, on the vertical axis shows N-uptake (kg ha^{-1}) from the ground measurements made with a proximal canopy reflectance sensor, for the two different processing levels top of atmosphere (L1C) and bottom of atmosphere (L2A). A linear model fitted the indices: normalised difference vegetation index (NDVI), normalised difference red-edge index based on band 8 and band 5 (NDRE85), normalised difference red-edge index based on band 8 and band 6 (NDRE86) and chlorophyll index (CI). A non-linear power function fitted the modified soil adjusted vegetation index (MSAVI2)

N-uptake prediction

The results of the spatiotemporal cross-validation procedure for the different models are summarised in Table 2. Use of L2A data instead of L1C data resulted in only minimal differences in the correlation statistics and modelling was therefore done only with the L1C satellite data. There was a relatively strong linear relationship ($r^2 = 0.74$) between the predicted and observed values. The Nash–Sutcliffe efficiency index confirmed relatively good model performance (E = 0.72). The MAE was 14 kg N ha^{-1} and ME was 4 kg N ha^{-1}, reflecting the spread around the 1:1 line in Fig. 6, and the prediction bias was small.

Model application

The prediction model based on the CI values was applied at around development stage DC37 to all winter wheat fields in the case study area (see Fig. 1). As the indicator maps in Fig. 7 show, fields with high or low within-field variation were spread across the area, but some spatial trends were distinguishable. For example, in the intense cultivation area in the west (the circles to the west in Fig. 7a, c), there was a tendency for fields to have relatively low within-field variability, whereas in the eastern part of the case study area fields with higher within-field variability in N-uptake were common. Of all fields larger than 5 ha, 62% showed variation in N-uptake greater than 30 kg N ha^{-1}. The within-field variation in N-uptake (inter-percentile range 2.5–97.5%) in the 4169 different fields ranged between 0 and 105 kg ha^{-1} (summarised in Fig. 8). Less than 50% of the fields were between 5 and 10 ha in size and these had the smallest within-field variation, on average 30 kg N ha^{-1}. The magnitude of within-field variation increased with field size, up to 41 kg N ha^{-1} for fields > 30 ha. The histograms in Fig. 8 show this variation, which shows a normal distribution. Field mean values of N-uptake, as measured with the Yara N-sensor, ranged between 28 and 149 kg N ha^{-1}, with an average of 90 kg N ha^{-1} (Fig. 9).

Comparisons between N-uptake predicted by the satellite model and the tractor-borne N-sensor are shown in Figs. 10 and 11; Table 3. The r^2 values ranged from 0.29 to 0.85 for the 13 fields. When all fields were considered together, r^2 was 0.81 and MAE was

Table 2 Validation statistics (modelling efficiency (E), mean error (ME), mean absolute error (MAE), coefficient of determination (r^2)) for prediction models for different vegetation-based indices at two different processing levels, top of atmosphere (L1C) and bottom of atmosphere (L2A)

	NDVI	NDRE85	NDRE86	CI	MSAVI2
L1C					
E	0.30	0.16	0.41	0.72	0.46
MAE (kg ha^{-1})	23	25	21	14	20
ME (kg ha^{-1})	2	− 1	2	4	8
r^2	0.31	0.20	0.42	0.74	0.55
L2A					
E	0.15	0.08	0.41	0.70	0.36
MAE (kg ha^{-1})	25	26	21	15	22
ME (kg ha^{-1})	− 4	− 4	1	5	9
r^2	0.21	0.15	0.41	0.73	0.48

Normalised difference vegetation index (NDVI); normalised difference red-edge index (NDRE85) based on band 8 and band 5; normalised difference red-edge index (NDRE86) based on band 8 and band 6; chlorophyll index (CI) and modified soil adjusted vegetation index (MSAVI2)

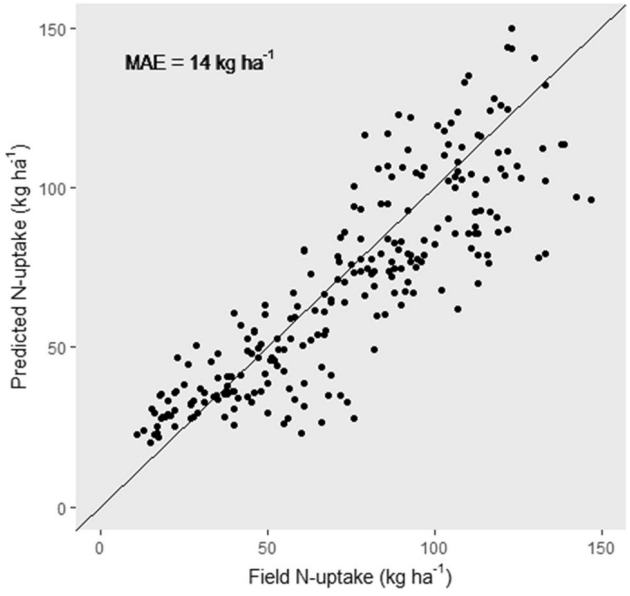


Fig. 6 Spatiotemporal cross-validation prediction based on a chlorophyll index (CI) model for top of atmosphere (L1C) satellite data versus field nitrogen (N)-uptake obtained from the proximal canopy reflectance sensor. The mean absolute error (MAE) is also shown

7 kg N ha^{-1}. High r^2 values were more common when the within-field variation was larger. The MAE varied slightly (4–12 kg N ha^{-1}) between the fields, and this variation did not appear to be clearly related to field size. Visual comparison revealed that the spatial pattern of within-field variability was fairly similar between maps produced by data from the tractor-borne N-sensor and maps generated from the Sentinel-2 L1C CI-based prediction model. A close-up view of a few fields is shown in Fig. 11.

Discussion

Using handheld proximal crop sensing data to build the model, it was possible to develop a well-performing, simple linear prediction model for N-uptake from Sentinel-2 data ($r^2 = 0.74$ for the CI-based model with L1C data). Prediction of N-uptake by sensors is done in reality through its correlation with total canopy chlorophyll content, which in turn is closely correlated with total canopy N content (e.g. Schlemmer et al. 2013). Using N-uptake models based on satellite data is a low-cost method to derive decision support for building N fertilisation strategies, quickly and inexpensively, for large cropping areas. In this case, the N-uptake prediction model was general, and 14 different winter wheat varieties were included. In addition, a relatively long period of crop development (DC22-53) was covered. Earlier work by Wolters et al. (2019) and Söderström et al. (2017) showed that satellite-based N-uptake models could be slightly improved if they were cultivar-specific, but this was not done in the present study.

Estimation of N-uptake by means of reflectance data from optical satellite data brings common challenges that arise when using remote sensing data. Irregular product quality

Fig. 7 Spatial representation of within-field variation in nitrogen (N)-uptake in the case study area (for location, see Fig. 1). One point represents one winter wheat field > 5 ha. The red circles to the left (west) and right (east) highlight areas with visible differences in the magnitude of within-field variation in N-uptake (kg ha^{-1}). Background: Sentinel-2 true colour composite with sensing date 27 May 2017

due to interference by clouds and cloud shadows can present difficulties in model calibration or model implementation in a DSS. Satellite data suppliers, in this case ESA, devote much effort to providing high-quality satellite data and are continually striving to improve their available products. However, regardless of ongoing development, some issues like persistent periods of cloud cover are difficult to overcome. Before implementation in a DSS, visual assessment of satellite data quality may be required by DSS providers and users.

With regard to development of models in the present study, there were enough cloud-free images available to get good estimates of N-uptake throughout the season

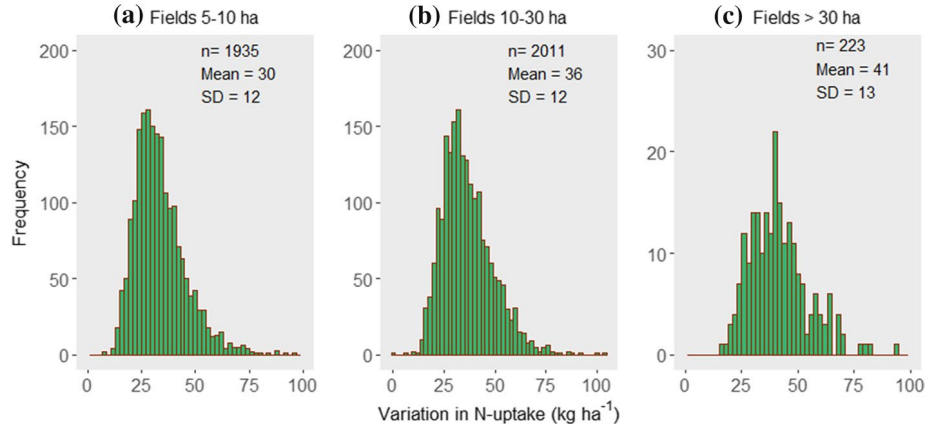


Fig. 8 Range of variation (inter-percentile range 2.5–97.5%) of nitrogen (N)-uptake within fields (kg ha^{-1}) for field sizes **a** between 5 and 10 ha, **b** between 10 and 30 ha and **c** more than 30 ha in the case study area

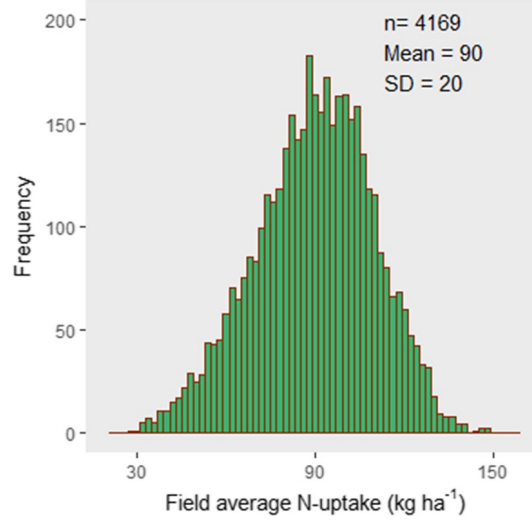


Fig. 9 Average nitrogen (N)-uptake (kg ha^{-1}) in fields > 5 ha in the case study area

of supplementary N fertilisation at the field data locations. The very small differences between models based on the L1C and L2A data (Fig. 5) suggest that the visual inspection of quality of the L1C images performed during the selection process was sufficient to yield directly usable data. L1C products gave slightly lower index values for the MSAVI2 vegetation index than the L2A products (Fig. 5). This was due to somewhat higher reflectance values in band 4 and lower reflectance values in band 8 (in L1C compared with L2A).

It was shown that, using a Sentinel-2 prediction model of N-uptake, it was possible to map the range of within-field variation over large areas. This can be valuable information for farmers, advisory workers and precision agriculture retailers, and in environmental protection-based programmes such as the Swedish 'Focus on Nutrients' Initiative (OECD 2018). In highly variable fields, the potential benefit of precision agriculture practices is likely to be greater from both an economic and an environmental point of view. Many fields in this study (> 62%) showed within-field variation exceeding 30 kg N ha^{-1}. The

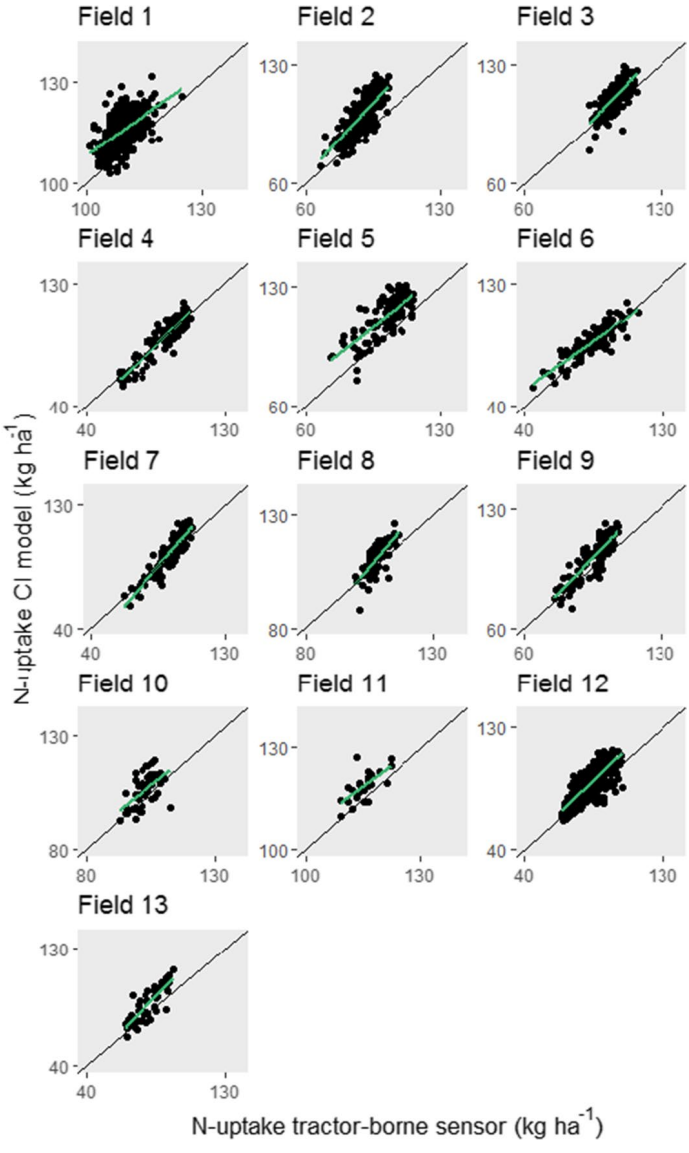


Fig. 10 Nitrogen (N)-uptake from the tractor-borne sensor compared with satellite-based prediction model N-uptake for the Ardala fields (see also Table 3; Fig. 11)

data presented therefore support previous conclusions by e.g. the European Union (EEA 2019; EPRS 2016) that VRA of N fertilisers is the way forward in using resources more efficiently and thereby also potentially mitigating emissions of greenhouse gases.

A future application of N-uptake prediction models is as part of N fertilisation algorithms, with the aim of deriving an economically optimum N-rate. For other crops and small fields that require high spatial detail, prediction models based on instruments mounted on other vehicles or stationary systems may be applicable. Having more

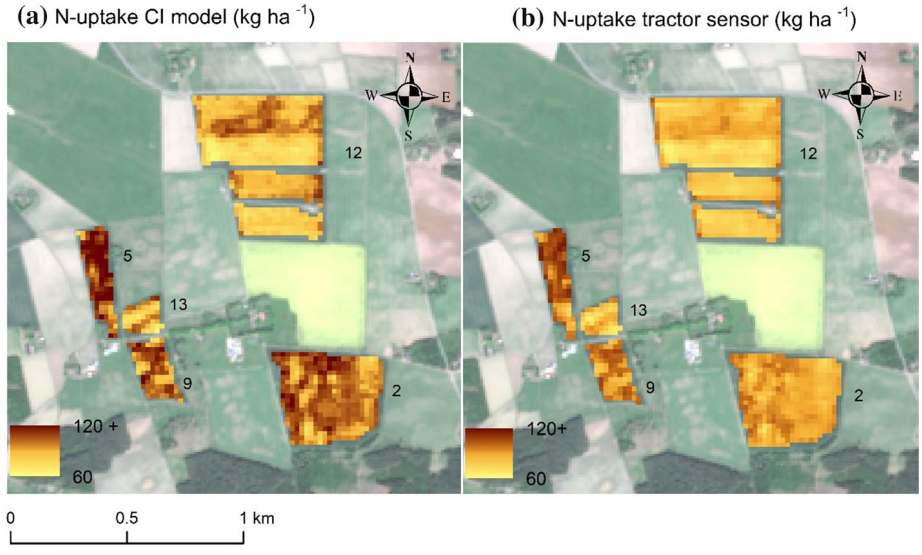


Fig. 11 Close-up view of some of the Ardala fields (location in Fig. 1). Left: the nitrogen (N)-uptake calculated from the satellite prediction model. Right: Interpolated N-uptake data from the tractor-borne N-sensor within two days of the aquisition date of the satellite image. The different fields are labelled with their field identification (ID) number (cf. Table 3; Fig. 10). Background: Sentinel-2 true colour composite with sensing date 27 May, 2017

Table 3 Comparison of nitrogen (N)-uptake (kg ha^{-1}) values produced by tractor-borne Yara N-sensor measurements (N-sensor) with N-uptake maps predicted by a model based on satellite data (Satellite) for 13 fields around Ardala village, arranged by field size (large to small)

ID no	Field size (ha)	N-sensor (kg ha^{-1})		Satellite (kg ha^{-1})		r^2	MAE (kg ha^{-1})
		Min	Max	Min	Max		
12	30.4	64	103	61	113	0.64	7
1	20.7	101	125	103	132	0.29	7
2	17.2	68	103	71	124	0.65	12
3	11.7	93	118	80	130	0.52	6
4	9.2	62	108	54	116	0.85	4
7	6.5	62	108	57	119	0.85	5
5	5.9	73	116	75	131	0.57	11
6	5.6	45	113	54	116	0.77	6
9	5.0	75	108	71	121	0.71	8
8	4.8	99	116	88	126	0.56	5
10	4.2	93	113	93	120	0.32	5
13	3.0	65	97	62	114	0.70	8
11	1.7	109	122	110	127	0.47	4
All	125.9	45	125	54	132	0.81	7

Coefficient of determination (r^2) and mean absolute error (MAE) are also shown

reflectance data available from different platforms, for example drones, could help increase model performance via downscaling of algorithms.

Models for N-uptake should ideally be continuously updated to include new cultivars and varying growing conditions. Weekly scanning of crops with proximal sensors over several seasons (as done by the Swedish Board of Agriculture for their advisory programme) provides an invaluable dataset of observations that are useful for producing models that are applicable in seasons with different growing conditions.

Winter wheat is an important staple crop and is widely cultivated in northern Europe (often in large fields), so there is great potential for using modelling techniques based on Sentinel-2 data for this crop. Considering the good resemblance between N-uptake maps from the model based on satellite sensing and those from the commonly used tractor N-sensor (Table 3), it can be inferred that satellite data may also be useful in practical VRA applications.

Conclusions

It was possible to scale proximal canopy reflectance sensor data to full spatial coverage in wheat production areas in Sweden using Sentinel-2 satellite data. A linear model based on the CI showed the best N-uptake prediction performance for new sites and dates (L1C data: $r^2 = 0.74$ and MAE = 14 kg ha^{-1}). Model predictions were not improved when using BOA reflectance values from Sentinel-2 L2A data compared with L1C data.

When Sentinel-2 satellite-based maps of N-uptake were compared with maps based on data from a tractor-borne sensor in 13 fields, a correlation of $r^2 = 0.81$ and MAE = 7 kg N ha^{-1} was found. Visual comparison of the maps showed a similar pattern of spatial variation. It was concluded that the satellite-based model, using CI values from L1C Sentinel-2 data for N-uptake prediction, gave results comparable to a tractor-borne in-field reflectance instrument.

Satellite-based scaling of proximal N-uptake measurements is useful for general assessments of within and between-field variation, making it possible to pinpoint fields where VRA of N would be most useful. The within-field variation in N-uptake exceeded 30 kg N ha^{-1} in 62% of all fields larger than 5 ha, indicating potential for major economic and environmental benefits from within-field VRA.

Acknowledgements Thanks to the project "Focus on Nutrients" at the Swedish Board of Agriculture for the georeferenced data on nitrogen uptake obtained by handheld sensor measurements. Thanks also to farmer Henrik Stadig for giving us access to the tractor-scanning data.

Author contributions Conceptualisation and method: WS, SM and PK; data collection and preparation: WS and SM; data analysis: WS; first draft preparation: WS; review and editing: WS, SM, PK, RH and SM; All authors have read and agreed to the published version of the manuscript.

Funding Open Access funding provided by Swedish University of Agricultural Sciences. This project was funded by Västra Götalandsregionen and the Swedish University of Agricultural Sciences (contract: RUN 2018-00141) together with Dataväxt AB, Sweden.

Data availability Data are not openly available, but raw data are accessible: Sentinel-2 data from the European Space Agency (https://scihub.copernicus.eu/dhus/), and compilations of the nitrogen uptake (kg ha^{-1}) measurements in newsletters (in Swedish) from the Swedish Board of Agriculture (https://greppa.nu/vara-tjanster/sasongsnytt.html).

Code availability Programming code is not openly available.

Compliance with ethical standards

Conflict of interest The authors declare no conflicts of interest. The funders had no role in the research design; in the collection, analyses, or interpretation of data; in the writing of the manuscript or in the decision to publish the results.

References

Barnes, E. M., Clarke, T. R., Richards, S. E., Colaizzi, P. D., Haberland, J., Kostrzewski, M., et al. (2000). Coincident detection of crop water stress, nitrogen status and canopy density using ground based multispectral data. In P. C. Robert, R. H. Rust, & W. E. Larson (Eds.), *Proceedings of the 5th international conference on precision agriculture* (pp. 16–19). Madison: American Society of Agronomy.

Basnet, B., Apan, A., Kelly, K., Strong, W., & Butler, B. (2003). Relating satellite imagery with grain protein content. In *Proceedings of the spatial sciences conference* (pp. 22–27). Los Angeles: Spatial Sciences Institute.

Berger, K., Verrelst, J., Féret, J., Wang, Z., Wocher, M., Strathmann, M., et al. (2020). Crop nitrogen monitoring: Recent progress and principal developments in the context of imaging spectroscopy missions. *Remote Sensing of Environment, 242*, 111758.

Börjesson, T., Wolters, S., & Söderström, M. (2019). Satellite-based modelling of protein content in winter wheat and malting barley. In J. Stafford (Ed.), *Precision agriculture, Proceedings of the 12th European conference on precision agriculture* (pp. 581–587). Wageningen: Wageningen Academic Publishers.

Burrough, P. A., & McDonnell, R. A. (1989). *Principles of geographical information systems*. New York: Oxford University Press.

Curran, P. J. (1989). Remote sensing of foliar chemistry. *Remote Sensing of Environment, 30*, 271–278.

Delin, S., & Stenberg, M. (2014). Effect of nitrogen fertilization on nitrate leaching in relation to grain yield response on loamy sand in Sweden. *European Journal of Agronomy, 52*, 291–296.

Drusch, M., et al. (2012). Sentinel-2: ESA's optical high-resolution mission for GMES operational services. *Remote Sensing of Environment, 120*, 25–36.

European Environment Agency (EEA). (2019). *Climate change adaptation in the agriculture sector in Europe*, Report: 04/2019, ISSN 1977-8449.

European Parliamentary Research Service (EPRS). (2016). *Precision agriculture and the future of farming in Europe, Brussels, European parliament*, ISBN 978-92-846-0475-3.

European Space Agency (ESA). (2020). *Copernicus open access hub*. Retrieved February 22, 2020 from https://www.sentinel-hub.com/.

Fletcher, K. (2012). Sentinel-2: ESA's *Optical High-Resolution Mission for GMES Operational Services*, ESA SP-1322/2.

Fredén, C. (1994). *Geology, National Atlas of Sweden*. Stockholm, Sweden: SNA Publishing.

Gascon, F., Bouzinac, C., Thépaut, O., Jung, M., Francesconi, B., Louis, J., et al. (2017). Copernicus sentinel-2A calibration and products validation status. *Remote Sensing, 9*(6), 584.

Gitelson, A. A., Gritz, Y., & Merzlyak, M. N. (2003). Relationships between leaf chlorophyll content and spectral reflectance and algorithms for non-destructive chlorophyll assessment in higher plant leaves. *Journal of Plant Physiology, 160*, 271–282.

Herrmann, I., Karnieli, A., Bonfil, D. J., Cohen, Y., & Alchanatis, V. (2010). SWIR-based spectral indices for assessing nitrogen content in potato fields. *International Journal of Remote Sensing*. https://doi.org/10.1080/01431160903283892.

Kokaly, R. F. (2001). Investigating a physical basis for spectroscopic estimates of leaf nitrogen concentration. *Remote Sensing of Environment, 75*, 153–161.

Main-Knorn, M., Pflug, B., Louis, J., Debaecker, V., Müller-Wilm, U., & Gascon, F. (2017). Sen2Cor for Sentinel-2. In *Conference: Image and signal processing for remote sensing* (3rd ed.). Bellingham: International Society for Optics and Photonics. https://doi.org/10.1117/12.2278218.

Mueller-Wilm, U. (2016). *S2 MPC: Sen2Cor configuration and user manual*, reference: S2-PDGS-MPC-L2A-SUM-V2.3.

Mulla, D. J. (2013). Twenty five years of remote sensing in precision agriculture: Key advances and remaining knowledge gaps. *Biosystems Engineering, 114*(4), 358–371.

Nash, J. E., & Sutcliffe, J. V. (1970). River flow forecasting through conceptual models part I: A discussion of principles. *Journal of Hydrology, 10*(3), 282–290.

Organisation for Economic Co-operation (OECD). (2018). Retrieved February 28, 2020 from https://read.oecd-ilibrary.org/.

Piikki, K., & Söderström, M. (2019). Digital soil mapping of arable land in Sweden, validation of performance at multiple scales. *Geoderma, 352*, 342–350.

Qi, J., Chehbouni, A., Huete, A. R., & Kerr, Y. H. (1994). Modified Soil Adjusted Vegetation Index (MSAVI). *Remote Sensing of Environment, 48*, 119–126.

R Core Team. (2018). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved February 22, 2020 from http://www.R-project.org/.

Reusch, S. (2003). Optimisation of oblique-view remote measurement of crop N-uptake under changing irradiance conditions. In J. Stafford & A. Werner (Eds.), *Precision agriculture. Proceedings of the 4th European conference on precision agriculture* (pp. 573–578). Wageningen: Wageningen Academic Publishers.

Reusch, S. (2005). Optimum waveband selection for determining the nitrogen uptake in winter wheat by active remote sensing. In J. Stafford & A. Werner (Eds.), *Precision agriculture 05* (pp. 261–266). Wageningen: Wageningen Academic Publishers.

Rouse, J. W., Haas, R. H., Scheel, J. A., & Deering, D. W. (1974). Monitoring Vegetation Systems in the Great Plains with ERTS. In *Proceedings, 3rd Earth Resource Technology Satellite (ERTS) Symposium* (1st ed., pp. 48–62).

Sawyer, J. E. (1994). Concepts of variable rate technology with considerations for fertilizer application. *Journal of Production Agriculture, 7*, 195–201. https://doi.org/10.2134/1994.0195.

Schils, R., Olesen, J. E., Kersebaum, K., Rijk, B., Oberforster, M., Kalyada, V., et al. (2018). Cereal yield gaps across Europe. *European Journal of Agronomy, 101*, 109–120. https://doi.org/10.1016/2018.09.003.

Schlemmer, M., Gitelson, A. A., Schepers, J., Ferguson, R., Peng, Y., Shanahan, J., & Rundquist, D. C. (2013). Remote estimation of nitrogen and chlorophyll contents in maize at leaf and canopy levels. *International Journal of Applied Earth Observation and Geoinformation, 25*, 47–54.

Singh, S. P. (2019). Site specific nutrient management through nutrient decision support tools for sustainable crop production and soil health. In *Soil fertility management for sustainable development* (pp. 13–23). Singapore: Springer.

Söderström, M., Börjesson, T., Pettersson, C. G., Nissen, K., & Hagner, O. (2010). Prediction of protein content in malting barley using proximal and remote sensing. *Precision Agriculture, 11*, 587–599.

Söderström, M., Piikki, K., Stenberg, M., Stadig, H., & Martinsson, J. (2017). Predicting nitrogen uptake 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 and Plant Science, 67*, 637–650.

Stafford, J. V. (2000). Implementing precision agriculture in the 21st century. *Journal of Agricultural Engineering Research, 76*(3), 267–275. https://doi.org/10.1006/2000.0577.

Swedish Board of Agriculture. (2019a). *Use of agricultural land 2019, Final statistics*. Retrieved March 18, 2020 from www.Jordbruksverket.se.

Swedish Board of Agriculture. (2019b). *The block map of agricultural land*. Retrieved February 22, 2020 from www.Jordbruksverket.se.

Verrelst, J., Camps-Valls, G., Muñoz-Marí, J., Rivera, J. P., Veroustraete, F., Clevers, J. G. P. W., & Moreno, J. (2015). Optical remote sensing and the retrieval of terrestrial vegetation bio-geophysical properties - A review. *ISPRS Journal of Photogrammetry and Remote Sensing*, *108*, 273–290. https://doi.org/10.1016/j.isprsjprs.2015.05.005.

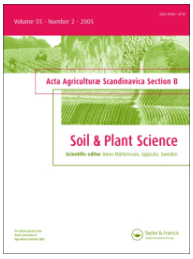
Wolters, S., Söderström, M., Piikki, K., & Stenberg, M. (2019). Near-real time winter wheat N-uptake from a combination of proximal and remote optical measurements: How to refine Sentinel-2 satellite images for use in a precision agriculture decision support system. In J. Stafford (Ed.), *Precision agriculture,*

proceedings of the 12th European conference on precision agriculture (pp. 1001–1007). Wageningen: Wageningen Academic Publishers. https://doi.org/10.3920/978-90-8686-888-9.

Zadok, J. C., Chang, T. T., & Konzak, C. F. (1974). A decimal code for the growth stages of cereals. *Weed Research, 14*, 415–421.

Zhao, C., Liu, L., Wang, J., Huang, W., Song, X., & Li, C. (2005). Predicting grain protein content of winter wheat using remote sensing data based on nitrogen status and water stress. *International Journal of Applied Earth Observation and Geoinformation, 7*, 1–9.

II



Taylor & Francis
Taylor & Francis Group

Acta Agriculturae Scandinavica, Section B — Soil & Plant Science

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/sagb20

Predicting grain protein concentration in winter wheat (*Triticum aestivum* L.) based on unpiloted aerial vehicle multispectral optical remote sensing

Sandra Wolters, Mats Söderström, Kristin Piikki, Thomas Börjesson & Carl-Göran Pettersson

To cite this article: Sandra Wolters, Mats Söderström, Kristin Piikki, Thomas Börjesson & Carl-Göran Pettersson (2022) Predicting grain protein concentration in winter wheat (*Triticum aestivum* L.) based on unpiloted aerial vehicle multispectral optical remote sensing, Acta Agriculturae Scandinavica, Section B — Soil & Plant Science, 72:1, 788-802, DOI: 10.1080/09064710.2022.2085165

To link to this article: https://doi.org/10.1080/09064710.2022.2085165

Published online: 25 Jun 2022.

ACTA AGRICULTURAE SCANDINAVICA, SECTION B — SOIL & PLANT SCIENCE
2022, VOL. 72, NO. 1, 788–802
https://doi.org/10.1080/09064710.2022.2085165

Taylor & Francis
Taylor & Francis Group

RESEARCH ARTICLE

OPEN ACCESS Check for updates

Predicting grain protein concentration in winter wheat (*Triticum aestivum* L.) based on unpiloted aerial vehicle multispectral optical remote sensing

Sandra Wolters [a], Mats Söderström[a], Kristin Piikki[a], Thomas Börjesson[b] and Carl-Göran Pettersson[c]

[a]Department of Soil and Environment, Swedish University of Agricultural Sciences (SLU), Skara, Sweden; [b]Agroväst Livsmedel AB, Skara, Sweden; [c]Lantmännen, Stockholm Sweden

ABSTRACT

Prediction models for crude protein concentration (CP) in winter wheat (*Triticum aestivum L.*) based on multispectral reflectance data from field trials in 2019 and 2020 in southern Sweden were developed and evaluated for independent trial sites. Reflectance data were collected using an unpiloted aerial vehicle (UAV)-borne camera with nine spectral bands having similar specification to nine bands of Sentinel-2 satellite data. Models were tested for application on near-real time Sentinel-2 imagery, on the prospect that CP prediction models can be made available in satellite-based decision support systems (DSS) for precision agriculture. Two different prediction methods were tested: linear regression and multivariate adaptive regression splines (MARS). Linear regression based on the best-performing vegetation index (the chlorophyll index) was found to be approximately as accurate as the best performing MARS model with multiple predictor variables in leave-one-trial-out cross-validation (R^2 = 0.71, R^2 = 0.70 and mean absolute error 0.64%, 0.60% CP respectively). Models applied on satellite data explained to a small degree between-field variations in CP (R^2 = 0.36), however did not reproduce within-field variation accurately. The results of the different methods presented here show the differences between methods used and their potential for application in a DSS.

ARTICLE HISTORY
Received 10 November 2021
Accepted 30 May 2022

KEYWORDS
Decision support system; multispectral; protein; Sentinel-2; unpiloted aerial vehicle (UAV); wheat

Introduction

Grain crude protein concentration (CP) is an important baking quality indicator in bread wheat (*Triticum aestivum* L.). The quantity and also quality, of CP (primarily the proteins glutenin and gliadin) affect gluten formation and the physical properties of bread dough (Gooding and Davies 1997). In many countries, wheat grain intended for milling and baking is sold for a higher price if a certain CP threshold is exceeded. In Sweden, that threshold is often 11.5% protein on a dry matter basis. CP concentration is therefore determined on all grain deliveries, using near infrared transmittance (NIT) sensing. It is also routinely determined in grain samples from winter wheat field trials, using the same technology.

Yield maps from monitors on combine harvesters are already used as a tool for farmers to evaluate crop management and for guidance in precision management in coming seasons, e.g. by splitting fields into management zones (Mulla 1993; Martínez-Casasnovas et al. 2018; Miao et al. 2018). Combine harvesters equipped with NIT sensors for CP mapping during harvest are also available (as described in Taylor et al. (2005) and Thylén and Algerbo (2001)), however, these are not yet as widely used as the yield mapping counterpart. Field zoning based on expected CP is an alternative method. With CP estimation before harvest, there is the option to split fields into harvesting zones with different expected CP in the current year. This would provide the option to exploit the spatial heterogeneity by selling some grain loads as bread wheat for a higher price and other loads as fodder wheat. Such models could also be useful for the grain industry, providing information on available quality at harvest. For example, Freeman et al. (2003) have shown how pre-harvest prediction of winter wheat grain yield and/or protein using the normalised difference vegetation index (NDVI; Rouse et al. 1974) could assist farmers in generating yield maps and reliable product marketing. However, the spatial pattern of CP within fields, and the relationship with for example yield, is complex and can vary substantially between years (e.g. Delin 2004).

Grain proteins are synthesised from nitrogen (N) that is translocated from other plant organs (50–70%) and from ongoing N-uptake during grain filling (30–50%) (Gooding 2009). It has been demonstrated in field trials that CP can be increased when additional N is applied in Zadok's growth stage (Zadok et al.

CONTACT Sandra Wolters sandra.wolters@slu.se Gråbrödragatan 19, 532 31 Skara, Sweden

1974) DC37 or later (Finney et al. 1957; Gooding and Davies 1992; Hamnér et al. 2017; Rodrigues et al. 2018; Hu et al. 2021. Sieling and Kage 2021), even though N use efficiency may be lowered by very late applications (Gooding 2009). If CP could be predicted from spectral reflectance data sensed by cameras mounted on UAVs, satellites, ground vehicles or hand-held instruments, CP predictions could serve as decision support for late-season N fertilisation aiming at meeting certain CP targets. Bastos et al. (2021) made a review of CP predictions, including remote and proximal sensing as well as on-combine sensors. They concluded that using on-combine protein measurements generated more accurate predictions than what could be achieved using proximal or remote sensing during the growing season. Prediction modelling of CP based on proximal and remote sensing data is often less successful than e.g. prediction models of yield (e.g. Freeman et al. 2003; Øvergaard et al. 2013; Barmeier et al. 2017). Barmeier et al. (2017) used a hyperspectral field sensor in field experiments during anthesis (DC 65) in malting barley, and used partial least squares regression (PLSR) for developing prediction models of protein content. Validation was done in independent field trials, but the models performed poorly ($R^2 = 0.28$), indicating the challenges in estimating protein content in harvested grain. Similar to Hansen et al. (2002) they did not find a clear effect of nitrogen fertiliser level on protein content. Prey and Schmidhalter (2019) tested a large number of vegetation indices and their correlation to grain N concentration, and found that only few parts of the electromagnetic spectrum within the visible-near infrared (NIR) region proved to be useful. Bands corresponding to around 780 nm in the lower part of the NIR region, combined with a band in the upper part of the red edge region performed best. Zhou et al. (2021) compared linear regression and some machine learning (ML; Liakos et al. 2018) methods for CP prediction in four fields in Japan, using data from a multispectral UAV based sensor. The results showed that it was difficult to predict protein accurately, although some ML methods seemed to perform better than the linear models. Bastos et al. (2021) reported that most studies on CP predictions in grain based on remote and proximal sensors probably overestimated the model accuracy and precision since the models were often not tested across different spatio-temporal scales.

Börjesson and Söderström (2003) showed that the best time for making protein predictions for winter wheat and malting barley (*Hordeum vulgare* L.) when using spectral data from handheld sensors is at the end of anthesis (Zadok's DC69). Basnet et al. (2003) and Prey and Schmidhalter (2019) also suggested that canopy reflectance derived just after anthesis are best correlated with grain protein content, and Bastos et al. (2021) found that this was most commonly reported in the studies included in their review. The reason may be that it provides information on sources of N available for protein formation through late season N remobilisation to the grain (e.g. Hansen et al. 2002). Söderström et al. (2010) demonstrated that it is possible to map CP in malting barley based on remotely sensed crop canopy reflectance together with ancillary variables. Börjesson et al. (2019) demonstrated that CP prediction in winter wheat could be performed based on a combination of early (DC37) and late (DC73) satellite reflectance data. Their prediction models had mean absolute error (MAE) of <1% CP when tested on independent fields.

The seasonal dynamics of N supply, from fertilisers and from mineralisation of organic matter in the soil in relation to the dynamics of (other) yield-influencing environmental conditions, such as water availability, disease, temperature and radiation, affect the CP content in the harvested grain (see e.g. Gooding 2009). Small on-farm experiments (OFE; Lacoste et al. 2022) called zero-plots and max-plots can be used for determining optimal N fertilisation rates (see e.g. Lory and Scharf 2003; Raun et al. 2011). Zero-plots are strategically placed plots in commercial fields (often 10–100 m^2) left without any N fertiliser application, which are used to monitor N-limited crop growth as a proxy for soil N supply. Max-plots are plots with an N-fertilisation level high enough for N not to limit yield, and crop growth in max-plots can be used as a proxy for potential N-uptake or potential yield (Johnson and Raun 2003; Piikki and Stenberg 2017). Since it is now relatively common to use zero-plots and sometimes also max-plots in cereal fields, it is interesting to investigate how these plots can be useful in CP prediction modelling. It has been shown, e.g. by Pettersson and Eckersten (2007), that there are differences between grain crop cultivars in terms of both canopy properties (e.g. structure or albedo) that affect vegetation indices and CP levels. Thus, it is possible that cultivar-specific models would perform better than general models. On the other hand, when general models are parameterised based on all cultivars used in a field trial, the resulting larger calibration dataset may give more robust models.

Satellite-based decision support systems (DSS) can reach many farmers and cover very large areas in comparison with handheld or tractor-based sensor systems. In agricultural research, an abundance of data on crop

qualities are available from field trials. Sensors carried by unpiloted aerial vehicles (UAVs) are suitable for collecting data in single fields for different crop management purposes (Du et al. 2017). It could be beneficial to use UAVs to collect spectral data in field trials, generate prediction models for crop properties and then apply such models on satellite images used in DSS to upscale the research results in practical implementation.

The aim of the present study was to develop and evaluate prediction models for CP at harvest in winter wheat, based on optical remote sensing data from UAV-borne multispectral sensors acquired in development stages DC69-73, in field trials. In this study we used a selection of vegetation indices based on spectral bands in the red to NIR region which in previous studies have been useful for describing the protein content (e.g. Freeman et al. 2003; Pettersson et al. 2006; Prey and Schmidhalter 2019). For modelling we used both linear regression and multivariate adaptive regression splines (MARS; Friedman 1991). MARS is a non-parametric ML method that in previous studies has been useful when dealing with remote sensing data in relatively small datasets (e.g. Filippi et al. 2014; Söderström et al. 2015). Special attention was paid to validation in order to validate on independent data. Moreover, the models developed were applied on satellite images currently used in DSSs and performance was evaluated against field observations. The following tests were carried out:

- *Screening for best CP modelling approach*, by leave-one-trial-out cross-validation of eight different modelling approaches comprising linear regression models or MARS models with different predictors (combinations of vegetation indices in trial plots with realistic N rates and chlorophyll index in zero-plots and max-plots)
- *Test of cultivar-specific models*, by leave-one-trial-out cross-validation
- *Test of models on Sentinel-2 data*, by comparing predictions with CP values from grain samples collected across five fields.

Materials and methods

Sites

The study area was Skåne (Scania) county in southern Sweden (approx. 55–56°N, 12–14°E). Data were used from five field trials, conducted at different locations. Three of the trials were performed in 2019 (at Stora Markie, Tommarp and Alnarp) and two in 2020 (at Lund and Brantevik) (Figure 1). All field trials were part of an ongoing national trial series (L7-0150, Nitrogen demand of different winter wheat cultivars[1]). The winter wheat cultivars grown were different between years, and 10 different cultivars were tested in each year. Each trial had six split N treatments (the latest applied in Zadok's development stage DC37). Total N rates were: 0, 80, 140, 200, 260 and 320 kg N ha^{-1}. The trial was replicated four times. In total, one trial consisted of 240 plots, of approximately 2 × 10 m in size (Figure 2). The focus in this study was on winter wheat cultivars intended for bread wheat, for which CP is critically important. Six bread wheat cultivars were grown in the field trials in 2019 and 2020 (cultivars: Etana, Hallfreda, Julius, Linus, Praktik and RGT Reform). The data for the satellite application were collected from five sites in the same region (A-E in Figure 1).

UAV measurements

Data were collected with a UAV octocopter (Explorian-8, Pitchup AB, Gothenburg, Sweden), equipped with a MAIA-S2 multispectral camera (Eoptis Srl, Trento, Italy), a global navigation satellite system (GNSS) unit and an incoming light sensor (ILS). The MAIA-S2 sensor has nine bands with centre wavelengths in the range 443–865 nm (Table 1). These bands correspond to nine of the bands in the Sentinel-2 satellite system (ESA, EU) (Nocerino et al. 2017).

The UAV was flown autonomously (at 80 m height and flight speed 5 m s^{-1}), using a pre-planned flight mission, in DC65-75, covering anthesis to medium milk development stages. Details of the flights are summarised in Table 2. The flights were mostly carried out between 12:00 and 16:30 h local time, when the solar incidence angle varied between about 40 and 55°. Each flight took about 10 minutes, during a period with uniform cloud conditions (clear sky or complete overcast), no precipitation the day before the flight, and not windier than a gentle breeze. At each trial site, 10 reflectance calibration panels (MosaicMill Oy, Vantaa, Finland) measuring 50 cm × 50 cm were placed on the ground (five along each end of the trial). These panels had known near-lambertian reflectance characteristics (2%, 9%, 23%, 44% and 75%) within the 400–900 nm range of the electromagnetic spectrum.

The UAV images were collected with at least 80% overlap both along and between flight lines. During a flight, the UAV's position was logged using GNSS and the incoming light from the ILS was logged at each photo point. Post-processing of images was performed using Multicam Stitcher Pro 1.1 provided by the manufacturer of the MAIA-S2 sensor. This software corrects for geometrical distortion and radial distortion of the

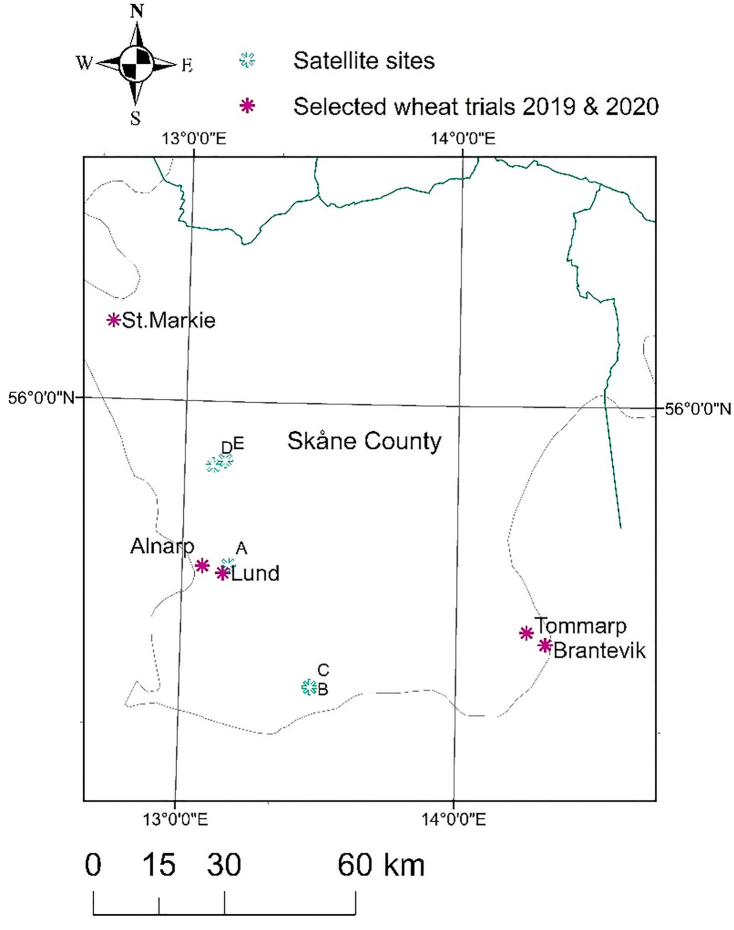


Figure 1. Field trial locations in Skåne (Scania) County, Sweden, and locations where satellite data were collected. Numbers (A-E) indicate locations of winter wheat field trials where trial-data were collected for testing crude protein (CP) prediction models using satellite data.

raw images and stitches the images of each of the nine bands into one multispectral image. The software also incorporates data from the ILS to radiometrically calibrate the images. The output images had pixel size of about 3×3 cm. Orthomosaics were created from the output images using the web application Solvi (https://solvi.ag; Solvi AB, Gothenburg, Sweden). These were downloaded and further processed in ArcGIS (ESRI Inc., Redlands, CA, USA). Using the reflectance panels, a linear function was derived to empirically recalculate the digital numbers of the orthomosaics to reflectance. The median reflectance of each band and for each trial plot was calculated (excluding a 0.2–0.3 m buffer zone along the plot edges). Data on harvested yield (kg ha^{-1}) and protein content (CP in % of dry matter) were extracted for each plot in the Nordic Field Trials System (NFTS; https://nfts.dlbr.dk; Danish Technological Institute and SEGES, Aarhus, Denmark). CP in this dataset was determined by FOSS Infratec1241 NIT equipment (FOSS, Hillerød, Denmark). These data were combined with the UAV data for the statistical modelling and analyses.

Statistical analyses and modelling

An initial screening process revealed potential problems with the data from the flight at Alnarp. In this particular field trial there was an unusually large amount of soil

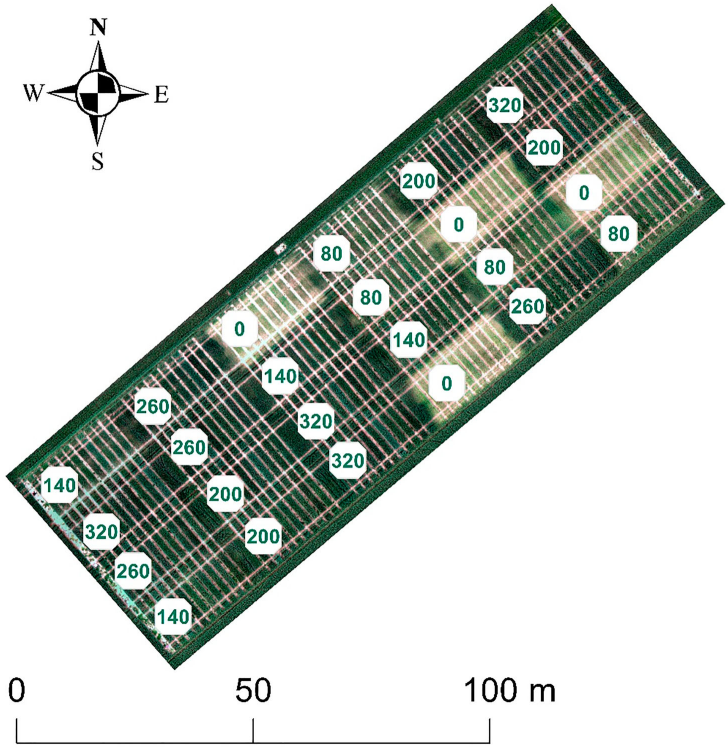


Figure 2. A field trial orthomosaic example from the unmanned aerial vehicle (UAV) sensor in Brantevik, on sensing date June 18, 2020. The different nitrogen (N) treatments are displayed in the trial blocks.

nitrogen available (the zero-plot yield was about 10 tonnes ha^{-1}) which made treatments in this trial similar and not representative for Swedish conditions, therefore this trial was excluded from analysis. In addition, two plots in the Brantevik trial and 14 plots in the Lund trial were excluded, due to missing data in the NFTS.

Plots where the applied N rate was >0 kg N ha^{-1} and <320 kg ha^{-1} were selected for model calibration and validation. The plots with 0 kg N ha^{-1} and 320 kg N ha^{-1} were denoted as zero-plots and max-plots, respectively and were not used in general modelling. A remote sensing-based vegetation index from these plots was introduced as a predictor. Reflectance values of different replicates (blocks) were averaged. The initial dataset contained 1184 records, and cleaning, treatment subsetting and block aggregation resulted in 96 records for building predictive models.

The correlations among and between predictors and CP content were explored using the Spearman

Table 1. Band specifications of the MAIA-S2 camera used in this study. NIR = near infrared.

MAIA-S2 sensor	Centre wavelength (nm)	Width of band (nm)	Band name	Corresponding band in Sentinel-2 constellation
S1	443	20	Violet	Band 1
S2	490	65	Blue	Band 2
S3	560	35	Green	Band 3
S4	665	30	Red	Band 4
S5	705	15	Red Edge 1	Band 5
S6	740	15	Red Edge 2	Band 6
S7	783	20	NIR 1	Band 7
S8	842	115	NIR 2	Band 8
S9	865	20	NIR 3	Band 8A

Table 2. Crop development stage, date, time and weather on the day of the unpiloted aerial vehicle (UAV) flights.

Trial	UAV flight date (yyyy-mm-dd)	Crop development stage	Local flight time (hh:mm)	Weather
Alnarp	2019-06-22	DC75	16:00-16:10	Clear sky
Stora Markie	2019-06-22	DC75	13:50-14.00	Clear sky
Tommarp	2019-06-23	DC71	11:55-12:05	Slight haze
Brantevik	2020-06-18	DC69	14:15-14:25	Overcast
Lund	2020-06-18	DC69	16:30-16:40	Overcast

correlation coefficient (r). The statistical procedures linear regression modelling (as explained in Hastie et al. 2009) and multivariate adaptive regression splines (MARS, Friedman 1991) were used in this study for building predictive models. Linear regression is a simple statistical procedure of fitting a linear model with one, or more (if multiple linear regression), predictor variables. MARS is a more flexible, non-parametric regression method which can predict using non-linear relationships with multiple predictors.

Vegetation indices

Seven vegetation indices (VIs) were calculated from the reflectance values in the different bands in the electromagnetic spectrum (Table 3). These were: (i) Optimised soil-adjusted vegetation index (OSAVI) (Huete 1988; Rondeaux et al. 1996), which was developed in an effort to minimise soil brightness influence by use of red and near-infrared (NIR) wavelengths. (ii) Red-edge inflection point (REIP), the computed wavelength where the crop canopy reflectance spectrum has its inflection point in the red-edge wavelength region, which is strongly related to chlorophyll content (Reusch 1997). A method to approximate this point presented by Guyot et al. (1988) was used in this study. (iii) Transformed chlorophyll absorption in reflectance index (TCARI) (Kim et al. 1994; Haboudane et al. 2002), which is one of several 'CARI' indices and indicates the relative abundance of chlorophyll. (iv) TC/OS, a ratio introduced by Haboudane et al. (2002) that is very sensitive to chlorophyll content variations and to variations in leaf area index (LAI: defined as half the total area of green elements of the canopy per unit horizontal ground area). This index is not sensitive to altitude. TC/OS has been found to show good results for protein predictions in malting barley (Pettersson et al. 2006). (v) Chlorophyll index (CI) (Gitelson et al. 2003), is another relevant simple ratio index. (vi). Normalised difference vegetation index (NDVI) (Rouse et al. 1974), is a very common NIR-visible-based ratio calculation introduced by Rouse et al. (1974). (vii) Normalised difference red-edge index (NDRE) (Barnes et al. 2000) is calculated in a similar manner to NDVI and includes a red-edge band instead of a red band, making it less sensitive to saturation if biomass is high compared with NDVI.

The equations in Table 3 show how these VIs were calculated, with reflectance (ρ) followed by the MAIA-S2 band number.

Table 3. The seven vegetation indices used in this study, with equations and references. Reflectance is expressed in spectral bands (ρ).

Index	Full name	Equation	References
OSAVI	Optimised soil-adjusted vegetation index	$1 + 0.16\frac{\rho 7 - \rho 4}{(\rho 7 + \rho 4) + 0.16}$	Huete 1988; Rondeaux et al. 1996
REIP	Red-edge inflexion point	$700 + 40\frac{\frac{\rho 4 - \rho 7}{2 - \rho 5}}{\rho 6 - \rho 5}$	Guyot et al. 1988
TCARI	Transformed chlorophyll absorption in reflectance index	$3(\rho 5 - \rho 4) - 0.2(\rho 5 - \rho 3)\frac{\rho 5}{\rho 4}$	Kim et al. 1994; Haboudane et al. 2002
TC/OS	Ratio calculation	$\frac{tcari}{osavi}$	Haboudane et al. 2002
CI	Chlorophyll index	$\frac{\rho 7}{\rho 6} - 1$	Gitelson et al. 2003
NDVI	Normalised difference vegetation index	$\frac{\rho 8 - \rho 4}{\rho 8 + \rho 4}$	Rouse et al. 1974
NDRE75	Normalised difference red-edge index	$\frac{\rho 7 - \rho 5}{\rho 7 + \rho 5}$	Barnes et al. 2000

Modelling strategies

The eight modelling strategies tested (Table 4) were combinations of two different model types, linear regression (a) and MARS modelling (b), and four different predictor sets (1–4). Linear regression models were based on a VI highly correlated with CP content, and MARS models were based on all seven VIs in Table 3. In some strategies, in addition to the VIs, the CI values in the zero-plots (denoted CI-zero) and/or the max-plots (denoted CI-max) were included as predictors. The index CI was pre-selected since this index was found to have the highest correlation with protein in this dataset. Irrespective of model type, the predicted values were constrained within reasonable limits. Predicted CP values below 8% were set to 8% and predicted CP values higher than 13.5% were set to 13.5%.

Table 4. The eight different modelling strategies with different model types and predictor sets tested in this study.

Strategy	Predictors	Model type	Validation
1a	Best single VI	LR	LTO + field test
2a	Best single VI + CI-zero	MLR	LTO
3a	Best single VI + CI-max	MLR	LTO
4a	Best single VI + CI-zero + CI-max	MLR	LTO
1b	All seven VIs	MARS	LTO + field test
2b	All seven VIs + CI-zero	MARS	LTO
3b	All seven VIs + CI-max	MARS	LTO
4b	All seven VIs + CI-zero + CI-max	MARS	LTO

All models were evaluated by leave-one-trial-out cross-validation (LTO). Two strategies were also evaluated by comparison with independent crude protein (CP) observations in five production fields. VI, vegetation index; LR, linear regression; MLR, multiple linear regression; MARS, multivariate adaptive regression splines. CI-zero, chlorophyll index in zero-plot; CI-max, chlorophyll index in max-plot.

Model cross-validation

To assess the prediction accuracy when applying a modelling strategy to new sites, leave-one-trial-out (LTO) cross-validation was performed. Records were repeatedly split into 'test data' (the record(s) for which a prediction was made) and 'training data' (the records used to parameterise the model). With each iteration, all data from one trial were assigned to the test set and the remaining records were assigned to the training set.

To determine prediction accuracy, two validation measures (mean absolute error, coefficient of determination) were calculated from the measured and predicted N-uptake values. Mean absolute error (MAE) is the average of the absolute prediction errors. Coefficient of determination (R^2) quantifies the prediction goodness-of-fit.

Model parameterisation

After LTO cross-validation, the final models were parameterised for each of the eight strategies using all data (no trials left out). In linear regression models (1a-4a in Table 4), all predictors obtained were included in the final model, however in the parameterisation of MARS model predictors with a little predictive power were discarded. Therefore, the final model may not include all predictors. Modelling was also carried out separately for each specific cultivar in the dataset. The cultivar-specific models were validated for the best-performing strategies in the general models.

Model field application

Model application was tested for general models from strategies 1a and 1b (Table 4) using satellite data for five winter wheat fields in Skåne county (locations A-E in Figure 1). These models were implemented on satellite data, Sentinel-2, L2A processing type, from June 25, 2020 (the cloud-free image expected to most closely correspond to DC 69-75). Field data were collected just before harvest in the year 2020. On each sampling location, one sample consisted of seven subsamples (for sampling stability), covering an area of approximately 3×3 m. There were eight samples obtained on each field. A total of 34 records remained after removal of sample points with incorrect location references, or those collected very close to field boundaries.

Software

The data were stored in a SQLserver database (Microsoft, Redmond, Washington, USA) accessed via SQL Server Management Studio (Microsoft, Redmond, Washington, USA) and analyses were carried out using R (R Core Team, 2021), including package 'Earth' (Milborrow 2021). ArcGIS 10.7 (Esri Inc., Redlands, California, USA) was used for spatial data analysis and display.

Results

In most compiled predictor sets, there were strong correlations between variables (Figure 3). For example, NDRE75 proved to be highly correlated with three other VIs, CI, TCOS and REIP. The VIs TCARI and TCOS were also very highly correlated, while TCARI and OSAVI showed very low correlation. All except two predictor pairs r was above 0.5. For the indices, the Spearman correlation ranged between −0.90 and 0.96. Wolters et al. (2021) have shown a good relationship with reflectance calculated to CI and N-uptake. In this study, the CI index also had the best correlation with grain CP ($r = 0.87$) content and was selected as the index to use in linear models.

Performance of general models

A summary of the results from LTO cross-validation for the six different bread wheat varieties is given in Table 5. Using the linear regression method and the best-performing index (CI), $R^2 = 0.71$ the MAE was 0.64% CP. Inclusion of CI values from zero-plots reduced the accuracy of the model slightly, to $R^2 = 0.60$ with MAE 0.71% CP. The max-plot CI value (CI-max) did not influence the prediction outcome substantially, $R^2 = 0.71$, MAE = 0.64% CP. A model with both the zero-plots and the max plot gave $R^2 = 0.60$, MAE = 0.71% CP for the linear method.

Leave-one-trial-out cross-validation with the MARS method gave lower accuracy in the results. The number of indices selected by the MARS method varied from three to five. Using all seven indices at the start of the procedure resulted in a model with $R^2 = 0.50$, MAE = 0.90% CP. The VIs: CI, OSAVI, TCOS and REIP were selected in this model. In the MARS method, the model was improved with the introduction of zero-plot CI values $R^2 = 0.63$, MAE = 0.70% CP. The opposite was the case when max-plots were used for modelling, $R^2 = 0.36$, MAE = 1.19% CP. The model performance improved again when VI values for both zero-plots and max-plots were introduced to the model $R^2 = 0.70$, MAE = 0.60% CP. In all models, for both methods, MAE was <1.20% CP.

Prediction results for all eight modelling strategies summarised in Table 5 are also presented in Figure 4. In two cases (1a, 3a), linear models were generally a better fit than the predictions with MARS modelling

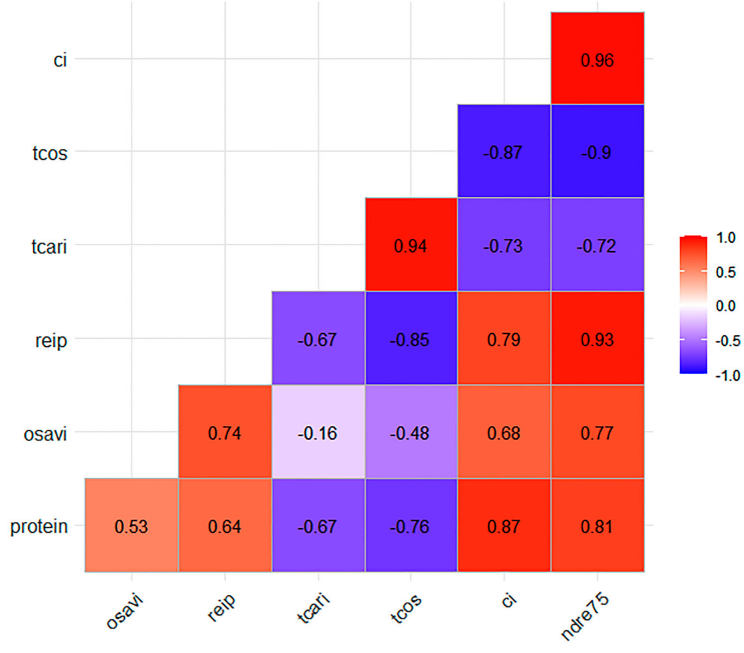


Figure 3. Spearman correlation coefficient (r) for the different predictors and crude protein (CP) content.

and in two cases (2b, 4b) the MARS models performed slightly better. In the prediction graphs for linear predictions, there was less variation between sites than in the MARS method graphs. For the site 'Tommarp', predicted values were lower than observed in the first three models (Figure 4).

Performance of cultivar-specific models

Results from LTO cross-validation of cultivar-specific models for strategies 1a and 4b (the best-performing strategies for general models for the two model types, Figure 4) are summarised in Table 6. Using the linear modelling method with CI (strategy 1a), it was possible to parameterise well-performing cultivar-specific models. Cultivar-specific linear models performed better than general models. For MARS models, more variation in the performance of different cultivar-specific models was found (Table 6). In both methods, the varieties 'RGT Reform' and 'Etana' stood out as good-performing varieties.

Table 5. Validation results (leave-one-trial-out) of crude protein (CP) content in wheat for the eight modelling strategies, using two model types (linear (a), MARS (b)), and four different combinations of predictor variables (1-4).

Strategy	R^2	MAE (% CP)	Predictors in final model
1a	0.714	0.64	CI
2a	0.602	0.71	CI, CI-zero
3a	0.710	0.64	CI, CI-max
4a	0.601	0.71	CI, CI-zero, CI-max
1b	0.504	0.90	CI, OSAVI, TCOS, REIP
2b	0.630	0.70	CI, CI-zero, OSAVI, TCOS
3b	0.363	1.19	CI, CI-max, OSAVI, TCOS, REIP
4b	0.703	0.60	CI, CI-zero, CI-max, NDRE75, REIP, TCARI, OSAVI

Indices used are shown in Table 3. R^2, goodness of fit; MAE, mean absolute error.

Satellite application

Model performance evaluation for a satellite dataset on five fields (strategy 1a-b) showed that the linear regression model based on one VI performed better (higher R^2) than the multivariate model based on four VIs (Figure 5). However, it is clear from Figure 5 that the model mainly explained between-field variation in those cases. The CP variation within the fields was small and could not be predicted well by the model. In application of the linear model there was a general prediction bias, with most predictions being higher than the observed values.

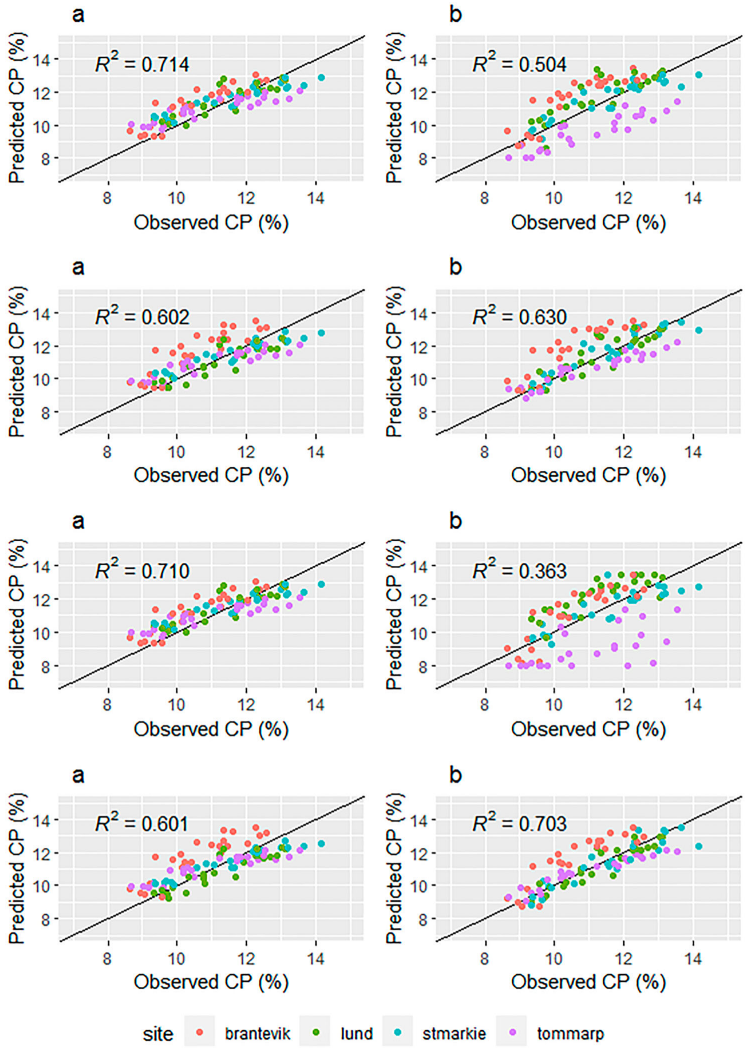


Figure 4. Leave-one-trial-out cross-validation for the eight modelling strategies (see Table 5). Predicted crude protein (CP) percentage vs. observed CP percentage.

Discussion

Field trials and decision support systems

Field trials are important for agricultural research. From a practical application perspective, it is advantageous if results can quickly reach end-users, in the form of advice and recommendations (Söderström et al. 2021). One part of the ongoing digitalisation of agriculture is increasing use of digital DSSs in precision agriculture. Some of these systems are using remote sensing multispectral data derived with satellites, or less commonly, UAVs. It is important to develop different types of relevant models that are suitable for implementation in DSSs. Thus, the aim in the present study was to develop a model for protein prediction at harvest in winter wheat using data from a UAV-borne sensor flown over a number of field trials during two seasons and transfer that model to a satellite data processing DSS and test it on a few fields. This approach, moving from

Table 6. Validation statistics from leave-one-trial-out cross-validation of linear or multiple linear regression models based on the best single index for strategy 1a (see Table 4) and Multivariate Adapted Regression Splines (MARS) prediction results for strategy 4b.

Strategy	Cultivar	R^2	MAE (% CP)
1a	Etana	0.86	0.41
1a	Hallfreda	0.70	0.69
1a	Julius	0.81	0.64
1a	Linus	0.79	0.60
1a	Praktik	0.80	0.50
1a	RGT Reform	0.87	0.40
4b	Etana	0.75	0.48
4b	Hallfreda	0.33	1.00
4b	Julius	0.70	0.69
4b	Linus	0.71	0.69
4b	Praktik	0.65	0.56
4b	RGT Reform	0.84	0.46

R^2, goodness of fit; MAE, mean absolute error; CP, crude protein content.

trials to DSS application, revealed opportunities and challenges.

UAV data collection

Ideally, UAV data is collected in all field trials during the same crop development stage, at the same time of day, and under similar weather conditions (e.g. Souza de et al. 2021). In reality this is not always feasible, especially in a project with trial sites located far apart (Figure 1). We aimed for a relatively narrow crop development window (DC69-73; as suggested in earlier studies, e.g. Bastos et al. 2021; Börjesson and Söderström 2003; Prey and Schmidhalter 2019;) however, in one of the trials used in the modelling measurements were made at DC75. Note also that the DC specified is the manually assessed average DC of each trial, although there were some differences within each trial, both random and between varieties and N rates. Since weather conditions were different, the aim was at least for uniform weather conditions where possible during individual flights, although three different types of weather were encountered during the flights: overcast, sunny and hazy (Table 2).

Trial design

The purpose of the trial design (Figure 2 & Table 2), was to test the response of different wheat varieties to different N rates. This means that the trials were, when possible, located in places where the crop variation was only driven by the treatment (i.e. different N rates and cultivars). In farmers' fields, this of course rarely is the case (Colaço and Bramley 2018). Conventional plot experiments are likely not the ideal method to evaluate variable rate technologies implemented to accommodate the effects of spatial variability in CP. Other factors may limit crop growth to different extents in different parts of a field, such as availability of other soil nutrients or water. A difficulty with the use of reflectance data is that the causes of variation in reflectance from the crop are usually not known. Studies separating e.g. water status from N status may provide valuable context in this case (Reese et al. 2010; Kusnierek and Korsaeth 2015).

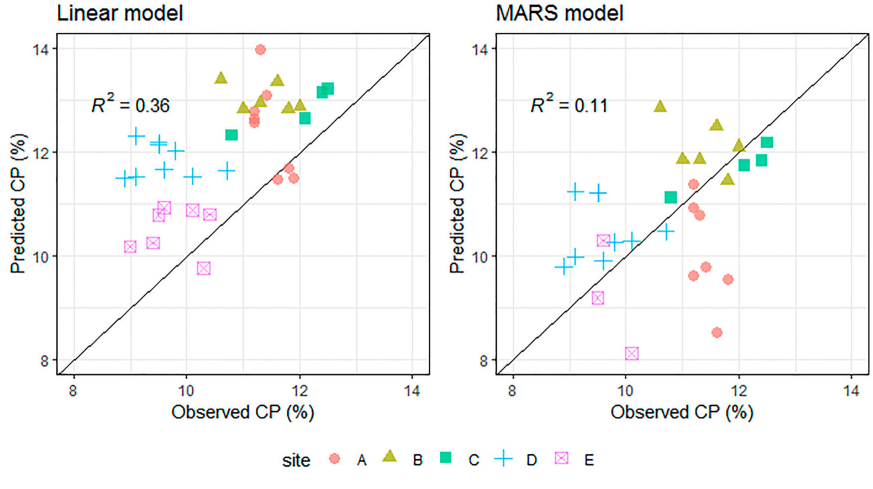


Figure 5. Predicted crude protein (CP) values versus observed CP values when the best-performing (left) linear model (strategy 1a) and (right) multivariate adaptive regression splines (MARS) model (strategy 1b) (see Table 4) as applied on Sentinel-2 data from the five test fields (locations A-E in Figure 1).

Protein prediction models

Even with the well-known relationship between CP and N rate (Terman et al. 1969; Hamnér et al. 2017; Sieling and Kage 2021), it is challenging to model protein content in field crops (Colaço and Bramley 2018), even when distinct fertilisation steps are present in the dataset (Pettersson et al. 2006). In this study we did not find benefit in using a more complex CP prediction model (MARS model based on multiple indices including zero-plots and max-plots) ($R^2 = 0.70$; MAE = 0.60% CP) over using the best linear model (based on one vegetation index, CI) ($R^2 = 0.71$; MAE = 0.64% CP). It appeared that the ML model did not optimise with this small dataset, and the independent validation. Given that it can be difficult to compare validation statistics between different studies, these models seem to perform relatively well compared to a statistical summary of reported research in this field (Bastos et al. 2021). In terms of best-performing index (CI), our study is in line with earlier work (e.g. Reusch 2005; Prey and Schmidhalter 2019) which indicates that an index based on two bands in the NIR-red-edge region is better correlated to N concentration and N-uptake than commonly used indices, such as NDVI and MSAVI2. There were in general small or no improvements in model performance when zero-plots and/or max-plots were included as predictors in the models (Table 5). In practical agricultural production in Sweden, the use of zero-plots and max-plots as a means of improving supplementary N fertilisation rate at DC37-45 in winter wheat appears to be increasing (Hushållningssällskapet 2021). However, data collection was done here at a considerably later stage, when the benefit of zero-plots and max-plots for protein prediction models may be limited. . There was much variation in the model accuracy between the different models including these predictors, and there was no proven benefit in preparing zero-plots and max-plots for this type of CP prediction.

The linear cultivar-specific protein prediction models worked better than the general model (Tables 5 and 6). The LTO validation statistics showed good performance, with the best variety being RGT Reform ($R^2 = 0.87$; MAE = 0.40% CP). When modelling was performed for each variety individually, the MARS model performed less well, possibly due to the relatively small number of observations used for calibration. When the MARS method was used, the variety 'RGT Reform' again performed best ($R^2 = 0.84$; MAE = 0.46% CP). With a more extensive dataset, the results would likely have been more clear. The response to N varies between cultivars (Fowler 2003). Some cultivars tend to be both high-yielding and at the same time have high CP content, whereas others can be high-yielding but the CP content will be lower. The yield-CP interaction differs among the cultivars tested in this study. For example, the cultivar Etana is a high-yielding high-protein cultivar, and in comparison the cultivar Hallfreda tends to have lower CP at the same yield, whereas the cultivar Praktik has a lower yield at the same CP (Hammarstedt 2021). This interaction is likely important when using canopy reflectance for modelling CP, and modelling yield variation is often easier than modelling CP (e.g. Barmeier et al. 2017).

Transfer of UAV data to satellite data

A question raised in the section on model field application, is whether it is possible to transfer models between platforms. Application of the UAV-based models on Sentinel-2 satellite data revealed that transfer of the model resulted in a low prediction accuracy and bias in the predictions. We used a UAV-borne sensor (MAIA-S2), which sensed bands with similar spectral characteristics as the intended satellite to be used in a DSS (Sentinel-2), however there are likely still discrepancies between the acquired reflectance values from the two sensing systems (shown in e.g. Bukowiecki et al. 2021; Peng et al., 2021; Rasmussen et al. 2021; Sarvia et al. 2021). Panels with known reflectance properties were used in the field trials and the digital numbers were calculated using an empirical relationship derived in each trial. Here we applied the protein prediction models on Sentinel-2 L2A data, which presumes that data from the two platforms are comparable. The prediction accuracy was not good when the models were applied, which indicates that the comparability in the field of data from MAIA-S2 and Sentinel-2 requires further investigation (Figure 5). An issue with applying models in satellite based DSS is that images from suitable dates (the growth stages for which the model was parameterised) must be available. The protein prediction model in this study was developed based on data collected at crop development stages DC69-75, a relatively large DC range, however a satellite image was not always available for that period. If the model is to be applied over large areas and the crop growth stages in individual fields are not known, the model performance may be limited. Earlier research suggested that spectral models during anthesis could be affected by large phenological shifts during that period, and that the most stable relationships between N concentration and canopy reflectance could be found during milk-

ripening stage (DC73; Prey and Schmidhalter 2019). In this case, the DC was not recorded in the fields. The Swedish Board of Agriculture has a website where weekly recordings of DC are reported (https://etjanst.sjv.se/povpub-gui/#/karta?produktionsinriktning=jordbruk, accessed March 10, 2022), and by following information from this page we noticed it is possible that in part of south-western Sweden the DC in winter wheat on June 25, 2020 was slightly higher than DC75.

Validation of the performance of a satellite-based model requires carefulness (e.g. Bastos et al. 2021). Collecting ground truth protein data in a growing wheat crop is difficult if the data are to be comparable to values sensed by a satellite such as Sentinel-2. The Sentinel-2 pixels are 400 m^2 (20 × 20 m) for the bands used in the best-performing models. Ground truth data in the field were subsamples collected within about 3 × 3 m areas. In this case, the sampling locations were not positioned according the outline of the pixels of the Sentinel-2 data. Even if this had been done, minor shifts in the georeferencing of the satellite image, practicalities during field work, and accuracy of positioning of the grain sampling location, still make it difficult to match a sample with a satellite image pixel. The sampling procedure is usually challenging and will inevitably have induced uncertainty in the results. In future research, the procedure of collecting ground truth data with modification will better match to the satellite data. The results reported in Figure 5 should be interpreted with the issues described above in mind. There appears to be some bias in the predictions (see linear model application in Figure 5), with the model over-predicting the protein content compared with the ground observations. If all observations are regarded collectively, the model made a poor prediction on the satellite dataset ($R^2 = 0.36$) and the within-field variation was not captured by the model. There was no field check done on the crop development stage for the different fields, and therefore the satellite data was selected from regional general information on reported crop growth stages in winter wheat, which will differ between fields. The simpler linear model performed better than the more flexible MARS model.

The modelling results of the different methods presented here add to the general knowledge base on CP estimation and protein models will maybe be applicable in a DSS in the future, when a different approach has been tested. Predicting and mapping the variation of protein content in grain crops remains a challenge, as also reported by Bastos et al. (2021). Further work on finding functional methods using proximal and remote sensing is recommended.

Note

1. The trial series forms part of Sverigeförsöken. It was conducted by Hushållningssällskapet and funded by Stiftelsen lantbruksforskning. Agronomic data from the trials are publicly available in the Nordic Field Trials System (NFTS; https://nfts.dlbr.dk; The Danish Technological Institute and SEGES, Aarhus, Denmark). More method details on the trials are also available in this database.

Acknowledgements

The authors wish to thank the funders, and the farmers who provided access to their fields. Also special thanks to Heather Reese, Associate Professor at the University of Gothenburg, for advice during the work on this study.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This research was funded by Västra Götalandsregionen and the Swedish University of Agricultural Sciences (contract: RUN 2018-00141) together with Dataväxt AB, Sweden. The remote sensing data were collected in a project funded by Stiftelsen Lantbruksforskning (contract: O-18-20-162). The test field data collection and laboratory analyses were conducted in a project funded by the Lantmännen Research Foundation (project number: 2016H040). The work by K. Piikki was funded by Formas National Research Programme for Food (contract: 2019-02280).

Data availability

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

Notes on Contributors

Sandra Wolters, PhD candidate, Swedish University of Agricultural Sciences (SLU). Previously affiliated with Wageningen Plant Research and Wageningen University, the Netherlands.

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

Kristin Piikki, plant physiologist with a PhD in Environmental Sciences (Gothenburg University, Gothenburg Sweden, 2006). Now researcher in precision agriculture and pedometrics. Holds a position as a Senior Lecturer at SLU and is an associate member of staff at The Alliance of Bioversity International and the International Center for Tropical Agriculture (CIAT), Nairobi, Kenya. Senior associate editor at Geoderma (Elsevier) and associate editor of Precision Agriculture (Springer).

Thomas Börjesson, agronomist with a PhD in Microbiology (Swedish University of Agricultural Sciences, Uppsala

Sweden, 1993). Now project leader at Agroväst Livsmedel AB, Skara Sweden. Agroväst is a NGO with emphasis on applied science and to promote agricultural development, including Precision Agriculture in West Sweden.

Carl-Göran Pettersson, agronomist with a PhD in Crop Science (Swedish University of Agricultural Sciences, Uppsala Sweden, 2007). Now Senior Project Manager Grain at Lantmännen Group R&D, Stockholm Sweden.

ORCID

Sandra Wolters http://orcid.org/0000-0001-7928-5026

References

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.

Barnes EM, Clarke TR, Richards SE, Colaizzi PD, Haberland J, Kostrzewski M, Lascano RJ. 2000. Coincident detection of crop water stress, nitrogen status and canopy density using ground based multispectral data. Proceedings of the fifth International Conference on Precision agriculture; Bloomington, MN, USA, 1619.

Basnet BB, Apan A, Kelly R, Jensen T, Strong W, Butler D. 2003. Relating satellite imagery with grain protein content. In *Proceedings of the 2003 Spatial Sciences Institute biennial conference: spatial knowledge without boundaries (SSC2003)*; Canberra, Australia. 22–27 September 2003, Spatial Sciences Institute, pp. 1–11.

Bastos M, de Borja Reis F, Sharda A, Wright A, Ciampitti Y, A I. 2021. Current status and future opportunities for grain protein prediction using On-and Off-combine sensors: A synthesis-analysis of the literature. Remote Sens (Basel). 13:5027. doi:10.3390/rs13245027.

Börjesson T, Söderström M. 2003. Prediction of protein content in cereals using canopy reflectance. In: Stafford JV, Werner A, editor. *Precision agriculture*. Wageningen: Wageningen Academic Publishers; p. 89–94.

Börjesson T, Wolters S, Söderström M. 2019. Satellite-based modelling of protein content in winter wheat and malting barley. In: Stafford J, editor. *Precision Agriculture, Proceedings of the 12th European conference on precision agriculture*. Wageningen: Wageningen Academic Publishers; p. 581–587.

Bukowiecki J, Rose T, Kage H. 2021. Sentinel-2 data for Precision agriculture? A UAV-based assessment. Sensors. 21:2861. doi:10.3390/s21082861.

Colaço AF, Bramley RGV. 2018. Do crop sensors promote improved nitrogen management in grain crops? Field Crops Res. 218:126–140. doi:10.1016/j.fcr.2018.01.007.

Delin S. 2004. Within-field variations in grain protein content, relationships to yield and soil nitrogen and consistency in maps between years. *Precis Agric*. 5(6):565–577.

Du MM, Noguchi N, Itoh A, Shibuya Y. 2017. Multi-temporal monitoring of wheat growth by using images from satellite and unmanned aerial vehicle. Int J Agric Biol Eng. 10(5):1–13. doi:10.25165/j.ijabe.20171005.3180.

Filippi AM, Güneralp I, Randall J. 2014. Hyperspectral remote sensing of aboveground biomass on a river meander bend using multivariate adaptive regression splines and stochastic gradient boosting. Remote Sens Lett. 5(5):432–441. doi:10.1080/2150704X.2014.915070.

Finney KF, Meyer JW, Smith FW, Fryer HC. 1957. Effect of foliar spraying on pawnee wheat with urea solutions on yield, protein content, and protein quality. Agron J. 49:341–347.

Fowler DB. 2003. Crop nitrogen demand and grain protein concentration of spring and winter wheat. Agron J. 95:260–265.

Freeman KW, Raun WR, Johnson GV, Mullen RW, Stone ML, Solie JB. 2003. Late-season prediction of wheat grain yield and grain protein. Commun Soil Sci Plant Anal. 34(13-14):1837–1852. doi:10.1081/CSS-120023219.

Friedman JH. 1991. Multivariate adaptive regression splines. Ann Stat. 19:1–67. doi:10.1214/aos/1176347963.

Gitelson AA, Gritz Y, Merzlyak MN. 2003. Relationships between leaf chlorophyll content and spectral reflectance and algorithms for non-destructive chlorophyll assessment in higher plant leaves. Journal of Plant Physics. 160:271–282.

Gooding, M.J. (2009). The wheat crop. In: Khan K, Shewry PR, editors. Wheat chemistry and technology (4th Edition). Cambridge: Woodhead Publishing and AACC International Press; p. 19–49.

Gooding MJ, Davies WP. 1992. Foliar urea fertilization of cereals: A review. Fertil Res. 32:209–222.

Gooding MJ, Davies WP. 1997. Wheat production and utilization. *systems*, *quality and the environment*. Cambridge: Cab International.

Guyot G, Baret F, Major DJ. 1988. High spectral resolution: determination of spectral shifts between the red and the near infrared. Int Arch Photogramm Remote Sens Spat Inf Sci. 11:750–760.

Haboudane D, Miller JR, Tremblay N, Zarco-Tejada PJ, Dextraze L. 2002. Integrated narrow-band vegetation indices for prediction of crop chlorophyll content for application to precision agriculture. Remote Sens Environ. 81:416–426.

Hammarstedt M. 2021. Mer än avkastning styr valet av höstvete (More than yield governs the choice of winter wheat). *Arvensis*, 4-5, 20-22. (In Swedish). Retrevied from: https://issuu.com/kb_arvensis/docs/arvensis_4-5_2021_inf_r_val_av_h_stvetesort.

Hamnér K, Weih M, Eriksson J, Kirchmann H. 2017. Influence of nitrogen supply on macro- and micronutrient accumulation during growth of winter wheat. Field Crops Res. 213:118–129. doi:10.1016/j.fcr.2017.08.002.

Hansen PM, Jørgensen JR, Thomsen A. 2002. Predicting grain yield and protein content in winter wheat and spring barley using repeated canopy reflectance measurements and partial least squares regression. J Agric Sci. 139:307–318.

Hastie T, Tibshirani R, Friedman JH. 2009. *The elements of statistical learning: data mining, inference, and prediction*, 2nd ed. New York: Springer.

Hu C, Sadras VO, Lu G, Zhang P, Han Y, Liu L, Xie J, Yang X, Zhang S. 2021. A global meta-analysis of split nitrogen application for improved wheat yield and grain protein content. Soil Tillage Res. 213: 105111. doi:10.1016/j.still.2021.105111.

Huete AR. 1988. A soil-adjusted vegetation index (SAVI). Remote Sens Environ. 25(3):295–309. doi:10.1016/0034-4257(88)90106-X. ISSN 0034-4257.

Hushållningssällskapet. 2021. Fieldreport. Retrieved from: https://hushallningssallskapet.se/wp-content/uploads/

2021/06/final-hushallningssallskapet-faltrapporten-nr-2-2021-low.pdf.
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.
Kim MS, Daughtry CST, Chappelle EW, McMurtrey JE, Walthall CL. 1994. The use of high spectral resolution bands for estimating absorbed photosynthetically active radiation (Apar). Proceedings of the 6th symposium on physical measurements and signatures in remote sensing; Val D'Isere, France, January 17-21, 299-306.
Kusnierek K, Korsaeth A. 2015. Simultaneous identification of spring wheat nitrogen and water status using visible and near infrared spectra and powered partial least squares regression. Comput Electron Agric. 117:200–213. doi:10.1016/j.compag.2015.08.001.
Lacoste M, Cook S, McNee M, Gale D, Ingram J, Bellon-Maurel V, MacMillan T, Sylvester-Bradley R, Kindred D, Bramley R, et al. 2022. On-Farm experimentation to transform global agriculture. Nat Food. 3:11–18. doi:10.1038/s43016-021-00424-4.
Liakos KG, Busato P, Moshou D, Pearson S, Bochtis D. 2018. Machine learning in agriculture: A review. Sensors. 18:2674. doi:10.3390/s18082674.
Lory JA, Scharf PC. 2003. Yield goal versus delta yield for predicting fertilizer nitrogen need in corn. Agron J. 95(4):994–999.
Martínez-Casasnovas J, Escolà A, Arnó J. 2018. Use of farmer knowledge in the delineation of potential management zones. In precision agriculture: A case study in maize (*Zea mays L.*). Agriculture. 8(6):84. doi:10.3390/agriculture8060084.
Miao Y, Mulla D, Pierre R. 2018. An integrated approach to site-specific management zone delineation. Front Agric Sci Eng. 5:432–441. doi:10.15302/J-FASE-2018230.
Milborrow S. 2021. Notes on the earth package. Retrieved from: http://www.milbo.org/doc/earth-notes.pdf.
Mulla DJ. 1993. Mapping and managing spatial patterns in soil fertility and crop yield. In: Robert PC, Rust RH, Larson WE, editor. *Soil specific crop management*. Madison: ASA/CSSA/SSSA; p. 15–26.
Nocerino E, Dubbini M, Menna F, Remondino F, Gattelli M, Covi D. 2017. Geometric calibration and radiometric correction of the maia multispectral camera. Int Arch Photogramm Remote Sens Spat Inf Sci. XLII-3/W3:149–156. doi:10.5194/isprs-archives-XLII-3-W3-149-2017.
Øvergaard SI, Isaksson T, Korsaeth A. 2013. Prediction of wheat yield and protein using remote sensors on plots—part I: assessing near infrared model robustness for year and site variations. J Near Infrared Spectrosc. 21:117–131.
Peng J, Manevski K, Kørup K, Larsen R, Neumann Andersen M. 2021. Random forest regression results in accurate assessment of potato nitrogen status based on multispectral data from different platforms and the critical concentration approach. Field Crops Res. 268:108158. doi:10.1016/j.fcr.2021.108158.
Pettersson CG, Eckersten H. 2007. Prediction of grain protein in spring malting barley grown in Northern Europe. Eur J Agron. 27: 205–214.
Pettersson CG, Söderström M, Eckersten H. 2006. Canopy reflectance, thermal stress, and apparent soil electrical conductivity as predictors of within-field variability in grain yield and grain protein of malting barley. Precis Agric. 7:343–359. s11119-006-9019-4.
Piikki K, Stenberg B. 2017. A modified delta yield approach for estimation of economic optimal nitrogen rate (EONR) for wheat (Triticum aestivum L.) and barley (Hordeum vulgare L.). Agric Food Sci. 26(4):233–241.
Prey L, Schmidhalter U. 2019. Temporal and spectral optimization of vegetation indices for estimating grain nitrogen uptake and late-seasonal nitrogen traits in wheat. Sensors. 19:4640. doi:10.3390/s19214640.
R Core Team. 2021. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. URL https://www.R-project.org/.
Rasmussen J, Azim S, Boldsen SK, Nitschke T, Jensen SM, Nielsen J, Christensen S. 2021. The challenge of reproducing remote sensing data from satellites and unmanned aerial vehicles (UAVs) in the context of management zones and precision agriculture. Precis Agric. 22:834–851. doi:10.1007/s11119-020-09759-7.
Raun W, Solie J, Stone M. 2011. Independence of yield potential and crop nitrogen response. Precis Agric. 12:508–518. doi:10.1007/s11119-010-9196-z.
Reese CL, Long DS, Clay DE, Clay SA, Beck DL. 2010. Nitrogen and water stress impacts hard Red spring wheat (Triticum aestivum) canopy reflectance. J Terr Obs. 2(1):1946–1143. Purdue University Press.
Reusch S. 1997. Entwicklung eines reflexionsoptischen Sensor zur Erfassung der Stickstoffversorgung landwirtschaftlicher Kulturpflanzen. Kiel Inst. für Landwirtschaftliche Verfahrenstechnik, Max-Eyth-Gesellschaft für Agrartechnik.
Reusch S. 2005. Optimum waveband selection for determining the nitrogen uptake in winter wheat by active remote sensing. In: JV Stafford, A Werner, editor. *Precision agriculture '05*. Wageningen: Wageningen Academic Publishers; p. 261–266.
Rodrigues FA, Blasch G, Defourny P, Ortiz-Monasterio J, Schulthess U, Zarco-Tejada PJ, Taylor JA, Gérard B. 2018. Multi-temporal and spectral analysis of high-resolution hyperspectral airborne imagery for precision agriculture: assessment of wheat grain yield and grain protein content. Remote Sens (Basel). 10(6):930. doi:10.3390/rs10060930.
Rondeaux G, Steven M, Baret F. 1996. Optimization of soil-adjusted vegetation indices. Remote Sens Environ. 55 (2):95–107. doi:10.1016/0034-4257(95)00186-7. ISSN 0034-4257.
Rouse JW, Haas RH, Schell JA, Deering DW. 1974. Monitoring vegetation systems in the Great Plains with ERTS. In: Freden SC, Mercanti EP, Becker M, editor. Third earth resources Technology satellite–1 syposium. volume I: technical presentations. Washington, DC: NASA SP-351, NASA; p. 309–317.
Sarvia F, De Petris S, Orusa T, Borgogno-Mondino E. 2021. MAIA S2 versus sentinel 2: spectral issues and their effects in the precision farming context. In: Gervasi O, Murgante B, Misra S, Garau C, Blečić I, Taniar D, Apduhan BO, Rocha AMAC, Tarantino E, Torre CM, editors. Computational science and its applications - ICCSA 2021. Cham: Springer; p. 63-77. doi:10.1007/978-3-030-87007-2_5.
Sieling K, Kage H. 2021. Apparent fertilizer N recovery and the relationship between grain yield and grain protein concentration of different winter wheat varieties in a long-term field trial. Eur J Agron. 124: 126246. doi:10.1016/j.eja.2021.126246.

Souza de R, Buchhart C, Heil K, Plass J, Padilla FM, Schmidhalter U. 2021. Effect of time of Day and Sky conditions on different vegetation indices calculated from active and passive sensors and images taken from UAV. Remote Sens (Basel). 13:1691. doi:10.3390/rs13091691.

Söderström M, Börjesson T, Pettersson C, Nissen K, Hagner O. 2010. Prediction of protein content in malting barley using proximal and remote sensing. Precis Agric. 11:587–599. doi:10.1007/s11119-010-9181-6.

Söderström M, Börjesson T, Roland B, Stadig H. 2015. Modelling within-field variations in deoxynivalenol (DON) content in oats using proximal and remote sensing. Precis Agric. 16:1–14. doi:10.1007/s11119-014-9373-6.

Söderström M, Pikki K, Stadig H. 2021. Yield maps for everyone. scaling drone models for satellite-based decision support. In: Stafford J, editor. *Precision Agriculture '21, Proceedings of the 13th European conference on precision agriculture*. Wageningen: Wageningen Academic Publishers; p. 911–918.

Taylor J, Whelan B, Thylén L, Gilbertsson M, Hassall J. 2005. Monitoring wheat protein content on-harvester: Australian experiences. In: Stafford JV, editor. *Precision Agriculture '05, Proceedings of the 5th European Conference on Precision Agriculture*. Wageningen: Wageningen Academic Publishers.

Terman GL, Ramig RE, Dreier AF, Olson RA. 1969. Yield-protein relationships in wheat grain, as affected by nitrogen and water. Agron J. 61:755–759.

Thylén L, Algerbo PA. 2001. Development of a protein sensor for combine harvesters. In: Grenier G, Blackmore S, editor. In ECPA 2001, *Proceedings of the 3rd European Conference on Precision agriculture*. Montpellier; p. 869–873.

Wolters S, Söderström M, Piikki K, Reese H, Stenberg M. 2021. Upscaling proximal sensor N-uptake predictions in winter wheat (Triticum aestivum L.) with Sentinel-2 satellite data for use in a decision support system. Precis Agric. 22:1263–1283. doi:10.1007/s11119-020-09783-7.

Zadok JC, Chang TT, Konzak CF. 1974. A decimal code for the growth stages of cereals. Weed Res. 14:415–421.

Zhou X, Kono Y, Win A, Matsui T, Tanaka TST. 2021. Predicting within-field variability in grain yield and protein content of winter wheat using UAV-based multispectral imagery and machine learning approaches. Plant Prod Sci. 24:137–151. doi:10.1080/1343943X.2020.1819165.

Acta Universitatis agriculturae Sueciae
Doctoral Thesis No. 2022:59

Nitrogen (N) fertilisation in crops can be made more efficient by moving from uniform application to meeting variable crop requirements within fields. Within field variable rate N fertilisation of winter wheat (*Triticum aestivum* L.) is practically feasible using information from web-based decision support systems (DSS). Data from different source platforms, such as satellite, unmanned aerial vehicle (UAV) or weather stations can be brought together in a DSS for fertilisation planning. In Sweden, satellite-based variable rate N fertilisation was available for winter wheat via a DSS, however, the existing module could be improved on different aspects.

Sandra Wolters carried out her PhD studies at the department of Soil and Environment at SLU in Skara, Sweden. She has a master of science from Wageningen University, the Netherlands.

Acta Universitatis agriculturae Sueciae presents doctoral theses from the Swedish University of Agricultural Sciences (SLU).

SLU generates knowledge for the sustainable use of biological natural resources. Research, education, extension, as well as environmental monitoring and assessment are used to achieve this goal.

ISSN 1652-6880
ISBN (print version) 978-91-7760-993-3
ISBN (electronic version) 978-91-7760-994-0