



MONASH University

**Coupling Fused High Spatio-Temporal
Resolution Remote Sensing Data and
Crop Modelling to Predict Wheat Yield
at the Field Scale**

by

Yuval Sadeh

M.A.

A thesis submitted for the degree of Doctor of Philosophy at
Monash University in (2021)

*School of Earth, Atmosphere and Environment
Faculty of Science
Monash University
Australia*

Copyright notice

© The author (2021).

I certify that I have made all reasonable efforts to secure copyright permissions for third-party content included in this thesis and have not knowingly added copyright content to my work without the owner's permission.

Abstract

Among all cereal crops traded in international markets, wheat is the most important. Accurate production forecasts, as early as possible prior to the harvest, are critical for the market stability, as well as for farmers, grains companies and governments. Providing reliable, consistent and scalable crop yield data is one of the major challenges in monitoring food security. While early crop yield estimations may help to control food prices in the western world, they could also be used to prevent hunger crisis in third world countries, by giving local authorities and international organisations enough time to prepare. For years, scientists around the world have tried to assess yields from space using remote sensing and crop models, however with only limited success. One factor limiting success has been the trade-off between high spatial and temporal resolutions in remotely-sensed data, particularly in attempts to estimate crop yield at field and sub-field scales. Over the last decade, an increasing number of nano-satellites (known as CubeSats) were launched, which opens the door to a new era of crop monitoring from space.

This study aims to improve in-season yield estimations by coupling crop modelling and satellite images, with a focus on wheat in Australia. In this thesis, I propose a new method named VeRsatile Crop Yield Estimator (VeRCYe), which seeks to overcome the limitation of missing data to estimate wheat yield at the field and pixel scales, by combining the advantages of both high spatio-temporal resolution remote sensing and crop model simulations. In this process, I have developed a satellites-based sowing date detection method at the field scale and fused PlanetScope images (with a spatial resolution of ~3 m) and Sentinel-2 images (with a spatial resolution of 10 m) to create daily Leaf Area Index (LAI) datasets at 3 m resolution. Finally, I have used the LAI datasets and the detected sowing dates with The Agricultural Production Systems simulator (APSIM)-Wheat model to predict wheat yield at the field and subfield scales.

The sowing date detection method developed uses Planet's PlanetScope data to detect changes on field surface caused by sowing. The method detected 85% of the sown fields with a very high correlation ($R^2 = 0.99$) between actual and estimated dates. Time of sowing was detected with a median gap of 0 days while achieving RMSE of 0.9 and 1.9 days in a national set of data and in a representative commercial farm, respectively. I have also tested the ability of this method for detecting harvested area and its timing, resulting with 0-day gap for the median between the detected and the reported harvest dates (RMSE = 2.6 days).

The fusion method proposed in this study, fuses time series imagery sourced from Sentinel-2 and Planets' PlanetScope CubeSat constellation. This enabled daily Sentinel-2 consistent, cloud free, surface reflectance RGB-NIR images and crop Green-LAI to be generated at a 3 m resolution. Overall, the results from the study demonstrated that the new fused time-series data combined the spatial, temporal and spectral advantages of both sensors, allowing wheat Green-LAI to be monitored on a daily basis with an RMSE = 0.35-0.63 and R^2 of 0.92 in wheat.

VeRCYe was tested over multiple wheat fields located in the Australian wheat-belt, covering a large range of environmental conditions and farm management practices across three growing seasons (2017 - 2019). VeRCYe not only successfully estimated field-scale yield with $R^2 = 0.88$ (RMSE of 757 kg/ha), but was also found to be effective for generating yield maps at 3 m resolution ($R^2 = 0.32$, RMSE of 1,213 kg/ha), up to four months before crop harvest.

In contrast to most of the previous studies, my PhD project has developed a new approach to estimate yield without ground calibration data, which will make it applicable across different regions and environments. The advantages of VeRCYe is that it can be used to estimate yield without the need for ground calibration, theoretically can be applied for other crop types and with any remotely sensed LAI. Furthermore, VeRCYe can generate useful information which may help to identify yield gaps, understand yield variability, its causes and scale from the pixel-level to a regional-level.

Thesis including published works declaration

I hereby declare that this thesis contains no material which has been accepted for the award of any other degree or diploma at any university or equivalent institution and that, to the best of my knowledge and belief, this thesis contains no material previously published or written by another person, except where due reference is made in the text of the thesis.

This thesis includes two original papers published in peer reviewed journals. The core theme of the thesis is predicting wheat yield using fused high spatio-temporal resolution remote sensing. The ideas, development and writing up of all the papers in the thesis were the principal responsibility of myself, the student, working within the School of Earth, Atmosphere and Environment under the joint supervision of Dr Xuan Zhu, Prof. David Dunkerley, Prof. Jeffrey Walker and Dr Karine Chenu (The University of Queensland).

In the case of chapters 2, 3 and 4 my contribution to the work involved the following:

Thesis Chapter	Publication Title	Status	Nature and % of student contribution	Co-author name(s) Nature and % of Co-author's contribution*	Co-author(s), Monash student Y/N*
2	<i>Sowing Date Detection at The Field Scale Using CubeSats Remote Sensing</i>	<i>Published</i>	<i>85%. Concept, collecting data, algorithm development and writing first draft</i>	1) <i>Xuan Zhu, input into manuscript 5%.</i> 2) <i>David Dunkerley, input into manuscript 5%.</i> 3) <i>Karine Chenu, Data interpretation, input into manuscript 5%.</i>	No
3	<i>Fusion of Sentinel-2 and PlanetScope time-series data into daily 3 m surface reflectance and wheat LAI monitoring</i>	<i>Published</i>	<i>73%. Concept, collecting data, algorithm development and writing first draft</i>	1) <i>Xuan Zhu, input into manuscript 2.5%.</i> 2) <i>David Dunkerley, input into manuscript 5%.</i> 3) <i>Karine Chenu, Data interpretation, input into manuscript 7.5%.</i> 4) <i>Jeffrey Walker, input into manuscript 5%.</i> 5) <i>Yuxi Zhang, Data collection input into manuscript 2.5%.</i> 6) <i>Offer Rozenstein, Data collection input into manuscript 2.25%.</i> 7) <i>V.S. Manivasagam, Data collection input into manuscript 2.25%.</i>	Yes
4	<i>Crop Yield Prediction at The Field Scale Using Fused High Spatio-Temporal Resolution Remote Sensing and Crop Modelling</i>	<i>not submitted</i>	<i>75%. Concept, collecting data, algorithm development and writing first draft</i>	1) <i>Xuan Zhu, input into manuscript 2.5%.</i> 2) <i>David Dunkerley, Data interpretation, input into manuscript 5%.</i> 3) <i>Karine Chenu, Data interpretation, input into manuscript 12.5%.</i> 4) <i>Jeffrey Walker, Data interpretation, input into manuscript 5%.</i>	No

I have renumbered sections of submitted or published papers in order to generate a consistent presentation within the thesis.

Student name: Yuval Sadeh

Student signature: **Date:** 29 August 2021

I hereby certify that the above declaration correctly reflects the nature and extent of the student's and co-authors' contributions to this work. In instances where I am not the responsible author I have consulted with the responsible author to agree on the respective contributions of the authors.

Main Supervisor name: Dr. Xuan Zhu

Main Supervisor signature: **Date:** 29 August 2021

Acknowledgements

I would like to acknowledge and express my sincere gratitude to my supervisors Dr. Xuan Zhu and Prof. David Dunkerley, Prof. Jeffrey Walker and Dr. Karine Chenu. They have supported me both professionally and personally throughout the numerous years of my PhD. I would like to thank Xuan for teaching me so much about the scientific world, David, for your wise and wide view on academic life and research, which helped me to “separate the wheat from the chaff”, and Karine, for your guidance, patience, and for teaching me so much on agronomy, agriculture and crop modelling, which more than once opened my eyes to look at things from a different perspective. I also want to thank Jeff for welcoming me to your research group and later accepting my nomination request to join my supervision panel. Your uncompromising scientific excellence challenged me time after time, and helped me to become a better researcher and academic.

This research was supported by the Co-Funded Monash Graduate Scholarship and Monash University Graduate Research Completion Award. I would like to thank the staff of the School of Earth, Atmosphere and Environment for their support throughout the years.

I really need to thank my parents Eitan and Hemda for all of their support. Thanks to you I was fortunate enough to grow up on a Kibbutz. This experience taught me about hard work, open-mindedness, resilience, and the importance of being able to identify goodness in every person and experience. My parents and the kibbutz provided me with the foundations for what I needed to succeed in this PhD journey. For instance, waking up long before dawn to drive the old tractor to the fields regardless of whether it was raining or cold, not dropping out from English lessons in high-school despite that I couldn't yet read nor write, not to give up even when I cast away stones in an endless field, and to find time to socialise every evening, understanding that there can be more in this life than work and study. I am thankful for everything you taught me.

A special thank you goes to Anita and Gary Frankel, who welcomed us in Melbourne and adopted us as their children. I am so grateful for their kindness and for being our Aussie family, especially in these challenging times.

Finally, I would like to thank my amazing wife for being crazy enough to leave everything behind in Israel (family, a career as a lawyer, quiet life in the country, and our dog) and to join me on this adventure together with our three-year-old daughter Ruth and our two and a half-

month-old daughter Zohara (not-to-mention five suitcases). I know this wasn't always easy for you, but I couldn't hope for a better partner than you on this journey! Without you and your endless support, none of this would become a reality. To Ruth and Zohara, thank you for teaching me the most important lesson in life, the kind of lesson that no university in the world could teach.

Publications and Prizes During the PhD Candidature

Following a full list of all publications that have resulted from the studies undertaken for this degree:

Peer-reviewed Journal Paper

Sadeh, Y., Zhu, X., Dunkerley, D., Walker, J. P., Zhang, Y., Rozenstein, O., ... & Chenu, K. (2021) Fusion of Sentinel-2 and PlanetScope time-series data into daily 3 m surface reflectance and wheat LAI monitoring. *International Journal of Applied Earth Observation and Geoinformation*, 96, 102260. <https://doi.org/10.1016/j.jag.2020.102260>

Manivasagam, V. S., **Sadeh, Y.**, Kaplan, G., Bonfil, D. J., & Rozenstein, O. (2021). Studying the Feasibility of Assimilating Sentinel-2 and PlanetScope Imagery into the SAFY Crop Model to Predict Within-Field Wheat Yield. *Remote Sensing*, 13(12), 2395. <https://doi.org/10.3390/rs13122395>

Sadeh, Y., Zhu, X., Chenu, K., & Dunkerley, D. (2019). Sowing date detection at the field scale using CubeSats remote sensing. *Computers and electronics in agriculture*, 157, 568-580. <https://doi.org/10.1016/j.compag.2019.01.042>

Peer-reviewed Conference Paper

Sadeh, Y., Zhu, X., Walker, J. P., Dunkerley, D & Chenu, K. (2021). “*Wheat Yield Prediction Using Fused CubeSat and Sentinel-2 Imagery*”. The 13th European Conference on Precision Agriculture. Budapest, Hungary.

Sadeh, Y., Zhu, X., Walker, J. P., Dunkerley, D., Chen, Y & Chenu, K. (2021). “*Paddock-scale yield estimation using fused PlanetScope and Sentinel-2 imagery and crop modelling*”. The 20th Australian Agronomy Conference. Toowoomba, Australia.

Sadeh, Y., Zhu, Z., Dunkerley, D., Walker, J., Zhang, Y., Rozenstein, O., V.S. Manivasagam and Chenu, K. (2020). “*Estimating daily Leaf Area Index with a 3 m resolution using CubeSat satellites*”. In Proceedings of the 15th International Conference on Precision Agriculture. Monticello, IL: International Society of Precision Agriculture, Minneapolis, USA.

Sadeh, Y., Zhu, Z., Dunkerley, D., Walker, J., Zhang, Y., Rozenstein, O., V.S. Manivasagam and Chenu, K. (2020). “*Sentinel-2 And PlanetScope Data Fusion into Daily 3 M Images for Leaf Area Index Monitoring*”, IGARSS 2020 - 2020 IEEE International Geoscience and Remote Sensing Symposium, Waikoloa, Hawaii, USA.

Sadeh, Y., Zhu, X., Chenu, K., & Dunkerley, D. (2019). “*Coupling Fused High Spatio-Temporal Resolution Remote Sensing Data and Crop Modelling to Predict Wheat Yield at the Field Scale*”. The 19th Australian Agronomy Conference. Wagga Wagga, New South Wales, Australia.

Sadeh, Y., Zhu, X., Chenu, K., & Dunkerley, D. (2019). “*Satellite-based detection of sowing dates in field crops*”. The 12th European Conference on Precision Agriculture. Montpellier, France.

Conference Oral Presentations

2021 Sadeh, Y., Zhu, X., Walker, J. P., Dunkerley, D & Chenu, K. Talk. “*Wheat Yield Prediction Using Fused CubeSat and Sentinel-2 Imagery*”. The 13th European Conference on Precision Agriculture. Budapest, Hungary.

2021 Sadeh, Y., Zhu, X., Walker, J. P., Dunkerley, D., Chen, Y & Chenu, K. Talk. “*Paddock-scale yield estimation using fused PlanetScope and Sentinel-2 imagery and crop modelling*”. The 20th Australian Agronomy Conference. Toowoomba, Australia.

2020 Sadeh, Y., Zhu, X., Dunkerley, D., Walker, J., & Chenu, K. Talk: “*Innovative Application of CubeSat Imagery to Predict Wheat Yield Without Ground-Based Data*”. AGU Fall Meeting 2020. Virtual Conference.

2020 Sadeh, Y., Zhu, Z., Dunkerley, D., Walker, J., Zhang, Y., Rozenstein, O., V.S. Manivasagam and Chenu, K. “*Estimating daily Leaf Area Index with a 3 m resolution using CubeSat satellites*”. In Proceedings of the 15th International Conference on Precision Agriculture (unpaginated, online). Monticello, IL: International Society of Precision Agriculture.

2020 Sadeh, Y., Zhu, Z., Dunkerley, D., Walker, J., Zhang, Y., Rozenstein, O., V.S. Manivasagam and Chenu, K. “*Sentinel-2 And PlanetScope Data Fusion into Daily 3 M Images for Leaf Area Index Monitoring*”, IGARSS 2020 - 2020 IEEE International Geoscience and Remote Sensing Symposium, Waikoloa, Hawaii, USA (Virtual Symposium) 2020.

2019 Sadeh, Y., Zhu, X., Chenu, K., & Dunkerley, D. Talk: “*Estimating Leaf Area Index from space at 3 m*”. The Australian Society of Plant Scientists. Melbourne, Australia.

2019 Sadeh, Y., Zhu, X., Chenu, K., & Dunkerley, D. Talk: “*Coupling Fused High Spatio-Temporal Resolution Remote Sensing Data and Crop Modelling to Predict Wheat Yield at the Field Scale*”. the 19th Australian Agronomy Conference. Wagga Wagga, New South Wales, Australia.

2019 Sadeh, Y., Zhu, X., Chenu, K., & Dunkerley, D. Talk: “*Satellite-based detection of sowing dates in field crops*”. The 12th European Conference on Precision Agriculture. Montpellier, France.

2018 Sadeh, Y. Talk: “*Implementing Satellite-Based Espionage Techniques in Agriculture: Detection of Sowing at The Field Scale Case Study*”. The 31st Victorian Universities Earth & Environmental Sciences Conference (VUEESC), Melbourne, Australia.

Awards

2020 - American Geophysical Union (AGU) Outstanding Student Presentation Award

2020 - Virtual Student Travel Grant - the 2020 American Geophysical Union (AGU) Fall Meeting

2020 - Monash University Graduate Research Completion Award

2020 - Best Oral Presentation - VUEESC 2020

2020 - The Upper Galilee Regional Council, Israel scholarship (2020) for PhD research in advanced agriculture technologies

2019 - The Australian Society of Agronomy - 2019 Conference Bursary

2019 - The European Conference on Precision Agriculture (France, 2019) - Student travel grant

2017 - Co-Funded Monash Graduate Scholarship – Monash University (fully funded PhD)

Abbreviations

API: Application programming interface

APSIM: Agricultural Production Systems sIMulator

ASTER: Advanced Spaceborne Thermal Emission and Reflection Radiometer

AVHRR: Advanced Very High Resolution Radiometer

BOA: Bottom of atmosphere

CF: Conversion factor

CWC: Canopy Water Content

ESA: European Space Agency

EVI2: Enhanced Vegetation Index 2

FAO: Food and Agriculture Organization

FAPAR: Fraction of Absorbed Photosynthetically Active Radiation

FCOVER: Fraction of vegetation cover

GCVI: Green Chlorophyll Vegetation Index

GEOGLAM: GEO Global Agricultural Monitoring

GNDVI: Green NDVI

Green-WDRVI: Green Wide Dynamic Range Vegetation Index

GSR: Green Simple Ratio

IDW: Inverse distance weighted

LAI: Leaf Area Index

MODIS: Moderate Resolution Imaging Spectroradiometer

MSAVI: Modified Soil-Adjusted Vegetation Index

MTVI2: Modified Triangular Vegetation Index 2

NASA: National Aeronautics and Space Administration

NDVI: Normalized Difference Vegetation Index

NIR: Near infra red

NVT: National variety trials

OSAVI: Optimized Soil-Adjusted Vegetation Index

PCA: Principal component analysis

PS: PlanetScope

RDVI: Renormalized Difference Vegetation Index

RGB: Red, Green, Blue

RMSE: Root mean square error

ROI: Region of interest

RS: Remotely Sensed

S2: Sentinel-2

SAFY: Simple Algorithm For Yield

SAR: Synthetic Aperture Radar

SCYM: Scalable satellite-based crop yield mapper

SIF: Solar-induced fluorescence

SNAP: Sentinel Application Platform

SR: Simple Ratio

STAIR: SaTellite dAta IntegRation

SWIR: Short-wavelength infrared

TIR: Thermal infrared

TOA: Top of atmosphere

TVI: Transformed Vegetative Index

VeRCYe: VeRsatile Crop Yield Estimator

VI: Vegetation index

VIIRS: Visible Infrared Imaging Radiometer Suite

WDRVI: Wide Dynamic Range Vegetation Index

WOFOST: World Food Studies Simulation Model

Table of Contents

Coupling Fused High Spatio-Temporal Resolution Remote Sensing Data and Crop Modelling to Predict Wheat Yield at the Field Scale	I
Abstract	III
Thesis including published works declaration	V
Acknowledgements	VII
Publications and Prizes During the PhD Candidature	IX
Abbreviations	XI
Table of Contents.....	XIII
1. Chapter 1 – Introduction	1
1.1. Background.....	1
1.2. Objectives and scope.....	3
1.3. Research Significance	4
Sowing date detection (Chapter 2)	4
Fusion of PlanetScope and Sentinel-2 into daily 3 m LAI (Chapter 3)	4
Yield prediction at the field and sub-field scales (Chapter 4).....	5
1.4. Thesis Outline.....	5
2. Chapter 2 – Sowing date detection	6
2.1. Abstract	6
2.2. Introduction	7
2.3. Methodology	11
2.3.1. Study fields and regions.....	11
2.3.2. Imagery	13
2.3.3. Semi-automated sowing detection.....	15
2.3.4. Change detection.....	15
2.4. Results.....	20
2.5. Discussion	22
2.5.1. Detecting no-tillage sowing with CubeSat satellites	22
2.5.2. Unprecedented levels of sowing detection	24
2.5.3. Limitations and perspectives	26
2.6. Conclusions	28
3. Chapter 3 – Fusion of PlanetScope and Sentinel-2 into daily 3 m LAI.....	29
3.1. Abstract	29

3.2.	Introduction	30
3.3.	Methodology	33
3.3.1.	Field trials and in-situ LAI measurements.....	33
3.3.2.	Imagery	34
3.3.3.	Data fusion of reflectance.....	35
3.3.4.	Fused LAI in 3 m.....	36
3.3.5.	Adjustment of remotely-sensed Green LAI.....	38
3.4.	Results.....	39
3.4.1.	Fused surface reflectance accuracy	39
3.4.2.	Fused vegetation indices.....	41
3.4.3.	LAI estimations.....	43
3.4.4.	Adjustment of S2 LAI to estimate Green LAI in wheat.....	46
3.5.	Discussion	48
3.5.1.	The advantages of the proposed fusion approach	48
3.5.2.	Estimating wheat Green LAI.....	50
3.5.3.	Limitations and prospects.....	52
3.6.	Conclusions	53
4.	Chapter 4 – The Versatile Crop Yield Estimator.....	55
4.1.	Abstract	55
4.2.	Introduction	55
4.3.	Methodology	59
4.3.1.	Sowing and harvest date detection	60
4.3.2.	APSIM model simulations	60
4.3.3.	Remotely sensed LAI	63
4.3.4.	Coupling APSIM model simulation with remotely sensed LAI for field scale yield estimations....	63
4.3.5.	Generating yield maps at 3 m spatial resolution	68
4.4.	Results.....	70
4.4.1.	Sowing and harvest date detection accuracy	70
4.4.2.	Field-scale yield estimations accuracy	71
4.4.3.	Yield map accuracy.....	73
4.5.	Discussion	75
4.5.1.	Wheat yield estimation at the field-scale	75
4.5.1.	The potential of generating 3 m yield maps.....	79
4.5.2.	Sowing dates as model inputs	80
4.5.3.	Limitations and prospects.....	82
4.6.	Conclusions	84
5.	Chapter 5 – Conclusions and future directions	86
5.1.	Sowing date detection.....	86
5.2.	Data fusion into high-resolution LAI.....	87
5.3.	VeRCYe.....	88
6.	References	90

Chapter 1 – Introduction

1.1. Background

Accurate yield estimations, as early as possible prior to harvest, are critical for market stability, farm management, grain companies and governments. Wheat is the largest broadacre crop in Australia with 18 million tonnes harvested in 2018-19 (Australian Bureau of Statistics 2020). However, risks and uncertainties within the global food system are growing with the projected increase in extreme weather events due to climate change (Ray et al. 2015). These may affect the variability of food prices in both short and long-term future.

Satellite-based remote sensing is considered a reliable, affordable, and timely source to improve crop yield prediction (Becker-Reshef et al. 2020), therefore many yield prediction methods using satellite data have been developed in the last few decades (e.g. Idso et al. 1977; Ferencz et al. 2004; Prasad et al. 2006; Franch et al. 2015). Traditionally, these methods are based on the correlation between Vegetation Indices (VIs) and crop yield (e.g. Raun et al. 2001; Labus et al. 2002; Becker-Reshef et al. 2010; Bognár et al. 2017). Many studies have shown a linear relationship between the photosynthetic capacity estimated from spectral responses and the crop phenology, which can be used to predict wheat yields using satellite remote sensing (e.g. Becker-Reshef et al. 2010; Franch et al. 2015; Bognár et al. 2017; Zhao et al. 2020). In recent years, methods which combine satellite images and machine learning techniques have become very popular (e.g. Cai et al. 2019; Jeffries et al. 2019; Feng et al. 2020; Kamir et al. 2020). However, reliance upon a unique and local relationship is not ideal, especially when crops experience highly variable environmental conditions, as in Australia, where crops are frequently stressed by heat waves, frosts and droughts (Chenu et al. 2013).

Crop production in Australia consists of two distinct cropping seasons, with wheat being the main winter crop across the whole grain belt. Wheat is traditionally planted from March to June and harvested from October to December of the same year (Potgieter et al. 2016). In Australia wheat is grown mainly in water-limited environments, which can cause substantial variations in crop yield, especially when water stress occurs around flowering (Chenu et al. 2011; Chenu et al. 2013). Furthermore, considerable yield losses can result from frosts around flowering, which prevent farmers from sowing their fields early to minimize risk of heat and drought stresses later in the growing season (Zheng et al. 2015). These conditions make it very challenging to make within-season yield predictions in Australia.

Over the last decade, the number of companies developing nano-satellites (also known as CubeSats) has increased. These new satellites, such as Planet Labs' PlanetScope CubeSats (Planet Team 2020), are relatively inexpensive to build including mass production, enabling both high spatial (<5 m) and temporal resolution (<1 week) imagery at low cost (Jain et al. 2016). However, in contrast to large expensive satellites such as Sentinel-2 or Landsat, CubeSat constellations frequently suffer from inconsistency in data collected by the different satellites in the constellation (Houborg and McCabe 2016). These inconsistencies limit the accuracy of surface reflectance-based applications such as estimation of VIs and hinder the use of these satellites to estimate leaf area index (LAI); defined as the total one-sided green leaf area per unit of soil area. Because photosynthesis takes place in the green parts of the plant, LAI is considered as an important plant characteristic. LAI has been found to be a good indicator of crop status and leaf abundance, as well as phenological stage, and can be used as an indicator of different farm management methods, or the impact of stresses and pests (Huang et al. 2019). Therefore, LAI also plays an important role in crop monitoring and can be used in crop growth models to better forecast yield (Clevers 1991; Bøgh et al. 2004; Lobell et al. 2015).

Recent studies have shown that improvement in crop yield model forecasting can be expected by using more frequent high-spatial and high-temporal satellite images per growing season (Jain et al. 2016; Jin et al. 2017a; Waldner et al. 2019). Therefore, the ability to generate high spatio-temporal resolution images and LAI datasets by fusing images acquired by well-studied satellites such as MODIS, Landsat or Sentinel-2 with CubeSats is needed.

Despite the growing availability of Earth observing data to monitor crop development and yield estimation, use of spaceborne sensors are limited by the type of data they can retrieve. Conversely, crop growth models can be used to simulate key physiological processes (Holzworth et al. 2014; Huang et al. 2019). Therefore, integrating the capabilities of remote sensing with crop model simulations has a great potential for improving capabilities in monitoring crop development and yield estimation through space and time. Two ways of merging the abilities of crop models and remotely sensed data are to either i) use data assimilation techniques, as assimilation of LAI into the models (Ines et al. 2013; Huang et al. 2015; Huang et al. 2019; Pan et al. 2019) or ii) by using satellite data to extract key model parameters, such as sowing dates, and use them as model inputs.

Sowing dates are a major input for crops models, which are commonly used to explore the expected yield effects of different management practices (Zheng et al. 2012; Holzworth et al. 2014; Chenu et al. 2017; Flohr et al. 2017). However, sowing dates are a source of

considerable uncertainty for regional studies (Mathison et al. 2017). Therefore, accurate information about sowing dates at a farm scale can be used to reduce the uncertainty of crop model simulations (Mathison et al. 2017).

While most studies have estimated crop yield at regional, county or state scales (e.g. Ines et al. 2013; Huang et al. 2015; Azzari et al. 2017; Jin et al. 2017b; Cai et al. 2019; Jin et al. 2019), few studies have attempted to estimate yields without calibration to ground data (e.g. Becker-Reshef et al. 2010; Franch et al. 2015; Lobell et al. 2015; Azzari et al. 2017; Jin et al. 2019). Moreover, despite the extensive scientific effort to utilize remote sensing for crop yield forecasting, a relatively small number of studies tried to estimate yield at the field and to map its within-field variability (e.g. Donohue et al. 2018; Lai et al. 2018; Chen et al. 2020; Manivasagam et al. 2021; Sagan et al. 2021), and only very few attempted doing so without any ground-based data for calibration (e.g. Jain et al. 2016; Burke and Lobell 2017; Dado et al. 2020; Deines et al. 2021), achieving limited success. To overcome these limitations it is essential to develop new methods which combine earth observation data with data generated from crop growth models, to eliminate the need for *in-situ* yield measurement and to preform global yield monitoring (Waldner et al. 2019).

1.2. Objectives and scope

The principal objective of this research was to develop a method to predict wheat yield at the field and pixel (i.e. sub-field) scales using remote sensing without using ground-based data. While most of the common satellite-based yield estimation methods extensively rely on in-situ data for training their models, they typically provide only a local solution for the area they have been calibrated in. The idea behind the method proposed in this thesis was therefore to develop a scalable and flexible yield estimation method which combines the power of both, i) high resolution spaceborne remote sensing and ii) crop modelling. In this process, the remotely sensed data provide continuous information of the crop's health, development and their spatial variability within the field, while the crop model simulates the crop growth processes and the farm management practices which can't be observed from space. To minimise the uncertainty of the model simulations, this research developed a method to detect the fields' sowing dates and to monitor the crop development using daily Leaf Area Index (LAI) at 3 m during the season, derived from fused remotely sensed data. This was later used as the link to APSIM crop model simulations.

The expected result from this study was to create a robust method that can ultimately estimate crop yields from the pixel level to regional scales for different types of broadacre crops, which theoretically can be used anywhere around the globe.

To achieve this, the main objective was divided into three sub-objectives:

- I. Detect sowing dates at a field scale using daily Planet Labs' PlanetScope images.
- II. Fuse PlanetScope and Sentinel-2 images to create a high spatio-temporal resolution LAI (Leaf Area Index) dataset.
- III. Couple remote sensing data and APSIM crop model simulations and to investigate the potential of using detected sowing dates and high spatio-temporal resolution LAI maps to predict wheat yields within Australia at the field and pixel scales. It is important to emphasise that the method employs no ground-based data.

1.3. Research Significance

This study illustrates that the method proposed in this thesis is capable of identifying when a field was sown, monitoring the crops performance and health daily at 3 m in form of VIs and LAI, estimating its field-scale yield, producing a yield map of the field at the pixel level months before the harvest and finally, detecting when the field is harvested. This section highlights the main findings and research significance of each chapter.

Sowing date detection (Chapter 2)

- A. High-resolution CubeSats images were used to detect sowing dates at the field scale for the first time.
- B. The sowing detection method achieved an unprecedented accuracy with RMSE of 0.9 and 1.9 days ($R^2 = 0.99$).
- C. The study overcame signal inconsistencies existing among the constellation's sensors.

Fusion of PlanetScope and Sentinel-2 into daily 3 m LAI (Chapter 3)

- A. A new method to fuse time series of images from two different satellite constellations was developed.
- B. The method combines the spatial, temporal and spectral advantages of both sensors.
- C. Daily Sentinel-2 consistent, surface reflectance RGB-NIR images generated at a 3 m.
- D. Daily monitoring of crop Leaf Area Index (LAI) at a 3 m resolution (R^2 of 0.94, 86% relative accuracy and RMSE of 1.37).

Yield estimation at the field and pixel scales (Chapter 4)

- A. A new method to estimate crop yield at the field-scale, without ground calibration data, was developed.
- B. The new method is able to estimate wheat yield at the field-scale in high correlation with the reported yield on average two months before the harvest.
- C. The method is able to produce estimated yield maps at the pixel level two to four months before the harvest.

1.4. Thesis Outline

This thesis is organized into five chapters with Chapter 1 being an introductory chapter on the background, objectives and scope of the research. Chapter 2 focuses on the sowing date detection method developed to address sub-objective number 1. This chapter is a reproduction of the published paper:

Sadeh, Y., Zhu, X., Chenu, K., & Dunkerley, D. (2019). Sowing date detection at the field scale using CubeSats remote sensing. *Computers and electronics in agriculture*, 157, 568-580. <https://doi.org/10.1016/j.compag.2019.01.042>

Chapter 3 describes the data fusion method developed to address sub-objective number 2. This chapter is a reproduction of the published paper:

Sadeh, Y., Zhu, X., Dunkerley, D., Walker, J. P., Zhang, Y., Rozenstein, O., ... & Chenu, K. (2021) Fusion of Sentinel-2 and PlanetScope time-series data into daily 3 m surface reflectance and wheat LAI monitoring. *International Journal of Applied Earth Observation and Geoinformation*, 96, 102260. <https://doi.org/10.1016/j.jag.2020.102260>

Chapter 4 presents the yield estimation method developed in this study along with a comprehensive accuracy evaluation of its ability to estimate wheat yield at the field scale and to produce yield maps at the pixel level. Chapter 5 summarizes the conclusions and the lessons learnt from this study, but also includes suggestions for future works that need to be addressed in order to take full advantage of the potential of the proposed yield estimation method.

As each of the chapters 2 – 4 are written as a paper, which include their own introduction and conclusions, the background section in Chapter 1 has been kept relatively brief and is intended therefore to provide a general background on the main topic of the thesis. Similarly, Chapter 5 summarizes only the main conclusions and the lessons learnt from this study.

Chapter 2 – Sowing date detection

This Chapter is a reproduction of the paper:

Sadeh, Y., Zhu, X., Chenu, K., & Dunkerley, D. (2019). Sowing date detection at the field scale using CubeSats remote sensing. *Computers and electronics in agriculture*, 157, 568-580. <https://doi.org/10.1016/j.compag.2019.01.042>,

with section and figure numbers changed to fit the thesis structure.

2.1. Abstract

Sowing dates have a great influence on crop yields as they affect what environmental conditions the plants will experience. Therefore, sowing dates are important to many individuals and organizations on the food production chain, including food manufacturers and traders. Despite their importance, large-scale comprehensive data on sowing dates are currently rare and often available only as broad estimates at the state, county or district level. Continuously obtaining sowing dates at the field scale is expensive, time-consuming and prone to human errors. Remote sensing on the other hand has the potential of conducting rapid, cost-effective and continuous surveys of farm management practices over large scales. Over the last decade, a new era in Earth observation satellites began with the production and deployment of smaller, lighter and cheaper nano-satellites known as CubeSats. Images from these satellites can potentially be used to capture sowing dates over time. In this study, we developed an innovative semi-automated sowing date detection methodology, which uses high spatio-temporal resolution CubeSat images to detect sowing dates at the field scale. We identified the sowing dates by using Planet's PlanetScope data to detect changes on the fields' surface caused by no-tillage sowing. Our approach overcame the signal inconsistencies existing among the numerous sensors in the constellation, and detected 85% of the sown fields with $R^2 = 0.99$. We succeeded to identify the actual sowing dates of individual fields with a median gap of 0 days within an unparalleled RMSE of 0.9 and 1.9 days in a set of national trials and in fields of a commercial farm, respectively. The methodology presented in this study can be used to produce and update field level crop statistics on a near daily-basis at a low cost. It is robust and simple and can be applied over a wide range of soil types, atmospheric conditions, crop types and sensors, to detect sowing dates in regions where no-tillage sowing is practiced.

2.2. Introduction

Sowing dates have a large impact on crop development and growth and ultimately crop yields as they affect what environmental conditions the plants will experience (Coventry et al. 1993; Flohr et al. 2017). Sowing dates are thus important to organizations and individuals who are interested in production, processing, marketing and trade of food and its products (Guo 2013). Sowing dates are one of the factors under the farmers' control to influence the crop environment and thus the yield (Ortiz-Monasterio and Lobell 2007), as earlier or later sowing frequently decreases crop survival and grain yields (Ozturk et al. 2006; Flohr et al. 2017). For instance, sowing too early increases the risk of frost in crops like wheat (*Triticum aestivum*) grown in post-heading frost prone regions, such as Australia (Zheng et al. 2012; Flohr et al. 2017). On the other hand, late sowing can result in a reduced growing season due to greater temperature (Duchemin et al. 2015) and greater risk of drought and heat stresses (e.g. Zheng et al. 2012; Chenu et al. 2013). Studies showed that the difference between 'best' the and 'worst' sowing dates could affect wheat yields by more than 1 t/ha (Ortiz-Monasterio and Lobell 2007) and even cause a crop failure, with a reduction of up to 100% of the yield (Zheng et al. 2015). Knowing the actual time when farmers sowed their crop could be used to better understand and improve current sowing practices and potentially increase yields (Ortiz-Monasterio and Lobell 2007). For instance, an analysis of sowing date records indicated that wheat farmers from the major cropping parts of Australia have shifted their sowing dates by around 1.5 days per annum over the last decade, and could still sow earlier to increase their yield (Flohr et al. 2018). Given the significance of sowing dates for crop production, detecting sowing dates would be valuable for different crops from a local to global scale (Manfron et al. 2017; Urban et al. 2018).

Sowing dates are major inputs for crops models, which are commonly used to explore the expected yield effects of different management practices (including sowing dates across a range of locations (Zheng et al. 2012; Holzworth et al. 2014; Chenu et al. 2017; Flohr et al. 2017)). However, sowing dates are a source of considerable uncertainty for regional studies (Mathison et al. 2017). Accurate regional information about sowing dates at a farm scale can be used to reduce the uncertainty of crop simulations (Mathison et al. 2017), however, obtaining accurate sowing dates from ground reports is very difficult and time consuming (Sacks et al. 2010; Marinho et al. 2014).

Traditionally, information associated with crop phenology (including sowing dates) is collected in farmers' surveys and is used to produce cropping systems or agro-ecological zoning maps that can be found as a form of national census. Even though sowing dates at a field scale are very important agronomically, only a few census datasets contain this information or are made

publicly available (in both developed and developing countries) (Urban et al. 2018). Furthermore, most of these data when they exist, are often infrequent, not quantitative, and available at the national scale or at the sub-national scale such as the state, county or district level (Sacks et al. 2010; Guo 2013). Producing statistics at the field level may improve our understanding of the reasons of yield gaps (i.e. gaps between achieved and achievable yields) (Hochman et al. 2012). Moreover, such a database could assist to identify fields with low-performing management where greater productivity could easily be achieved by adapting the sowing date (Jain et al. 2016). To overcome this issue, many studies have used satellite remote sensing to map sowing dates over large areas using different spatial and temporal resolutions (e.g. Lobell et al. 2003; Sakamoto et al. 2005; Ortiz-Monasterio and Lobell 2007; Lobell et al. 2013; Marinho et al. 2014; Jain et al. 2016; Manfron et al. 2017; Urban et al. 2018).

The advantages of remote sensing over the traditional approaches to collecting data on crop management practices have long been recognised (Ortiz-Monasterio and Lobell 2007). However, a methodology for detecting sowing dates directly from remote sensing imagery has not yet been developed. That is mainly because there is a lag-time between the sowing date and the plant emergence. Jin et al. (2016) argued that it is impossible to detect the sowing dates directly. Most existing studies attempted first to detect the crop “green-up”, i.e. the earliest reliable evidence of vegetation that can be sensed on satellite images (Lobell et al. 2013; Manfron et al. 2017), and then to backcast emergence and sowing dates by assuming the lag time between the sowing and the green-up detection to be constant (e.g. Lobell et al. 2013; Marinho et al. 2014; Duchemin et al. 2015; Jin et al. 2016; Gao et al. 2017; Manfron et al. 2017). In India, Lobell et al. (2013) used time series of MODIS and SPOT satellite Vegetation Index (VI) products, both with a spatial resolution of 1 km, to estimate the green-up date each year at a district level. This allowed them to infer sowing dates within two-day difference from reported dates at the district level. In China, Jin et al. (2016) estimated wheat sowing dates using Normalized Difference Vegetation Index (NDVI) time series derived from the Chinese satellite HJ-1A/B (with a spatial resolution of 30 m and 2 days revisit time). They estimated sowing dates based on the relationship between the green-up and the sowing dates following the method from Lobell et al. (2013), and found sowing-dates estimates differing by an average of 6 days from the ground-based estimates. In the U.S.A, Urban et al. (2018) compared three satellite-based sowing estimation methods applied for maize and soybean at a county-level by using three different sensors that cover a large range of the electromagnetic spectrum. They used MODIS (optical) Enhanced Vegetation Index (EVI), solar-induced fluorescence (SIF) from GOME-2 (fluorescence) and Ku-band backscattering (dB) from QuikSCAT (radar) to detect the initial green-up stage. In their case, both EVI and SIF based estimations of sowing dates reached R^2 of 0.75 and RMSE < 10 days from their county-level

validation data, whereas the radar-based estimates were negatively correlated with observed dates for most counties. Although the green-up detection approach gained popularity in the last two decades, using a fixed lag-time limits this method to the regions where this constant is known or can be determined. Furthermore, this approach neglects the strong impact of the weather variability, such as the timing of rainfall, on crop emergence dates (Marais Sicre et al. 2016). Another method estimated the fraction of absorbed photosynthetically active radiation (fAPAR) to estimate the sowing date. But using this method, Lobell et al. (2003) only managed to predict planting months at a high confidence (up to 50% of the mean). Ortiz-Monasterio and Lobell (2007) used estimated fAPAR together with measured daily temperature and radiation as inputs in a crop growth model to estimate sowing dates. They achieved a high level of agreement ($R^2 = 0.85$, RMSE = 6.5 days) between their estimated and reported sowing dates. In Japan, Sakamoto et al. (2005) detected paddy fields which were ploughed and flooded before the rice planting, by identification of the decreased VI values during this period, using time series of MODIS-based EVI. They achieved an RMSE of 12.1 days between their sowing-date estimations and official regional statistics.

Remote sensing has also been applied to monitor other agricultural operations such as tillage (ploughing and harrowing), irrigation and harvesting (Mc Nairn et al. 1998; Sakamoto et al. 2005; Hadria et al. 2009; Pacheco et al. 2010; Guo 2013). In terms of sowing, sowing crops into untilled soil (also called 'no-tillage', 'zero tillage' farming or 'conservation agriculture' (Derpsch et al. 2010)) has increasingly been adopted around the world over the last decades (Hobbs et al. 2008). This technique presents a lot of advantages such as fuel, time and labour conservation, soil improvement (e.g. increase in soil nitrogen, organic matter, and water infiltration; decrease in soil erosion), lower costs and potential for yield increase (Baker and Saxton 2006; Pacheco et al. 2010). Due to these benefits (Zheng et al. 2014), the cropland area under conservation tillage increased worldwide by 47% in four years only, i.e. from 106.5 million ha in 2009 to 157 million ha in 2013 (Kassam et al. 2015).

With the no-tillage farming practice, seeds are directly sown into the previous crop residue and stubble. Therefore, we can assume that sowing represents the first change on the surface at the field-scale since the last harvest. From a remote sensing perspective, this means a distinct contrast between the sown ('disturbed') and un-sown (untilled) soil. Theoretically, this change can be used to detect sowing dates at the field scale using time-series of satellite images. Remotely-sensed tillage-mapping approaches already exist. They are based on the assessment of two features that are strongly modified by tillage systems: residue cover and surface roughness. Previous studies used optical, multi or hyperspectral sensors for qualitative and quantitative estimations of residue cover (Bégué et al. 2018). Daughtry (2001) showed that crop residues and soils often have similar spectral signature. However, crop

residue has a unique absorption of electromagnetic radiation near 2100 nm, which is associated with cellulose and lignin. The understanding of this absorption feature lays the foundation for a number of tillage indices derived from optical remote sensing with shortwave infrared (SWIR) bands (1200-2500 nm), such as ASTER, Landsat and MODIS (Serbin et al. 2009; Zheng et al. 2014).

To accurately detect sowing dates at a field scale from space using a no-tillage detection approach, a new high spatial and temporal resolution remote sensing source is needed. While satellites such as Sentinel-2 and Landsat may have adequate spatial resolution for different applications in agriculture (e.g. Battude et al. 2016; Skakun et al. 2017; Bégué et al. 2018), their temporal resolution (5 and 16 days revisit time, respectively) is not ideal for sowing detection, as there may be weeks between the acquisition of two clear-sky images (McCabe et al. 2017). To overcome these spatio-temporal limitations constellations of micro or nano-satellites known as CubeSats can be used (McCabe et al. 2017).

Over the last decade, the number of companies developing CubeSat satellites has increased. These new satellites, such as Planet Labs (“Planet”) PlanetScope and Skybox imaging SkySat satellites, are relatively inexpensive to build and allow mass production which enables the creation of a collection of both high spatial (<5 m) and temporal resolution (<1 week) imagery at lower cost (Dash and Ogutu 2016; Jain et al. 2016). These satellites have the potential to monitor and detect rapidly changing environments on the Earth surface (McCabe et al. 2017) and to obtain multiple measures of the same field, including sowing and harvesting dates throughout a single growing season, as demonstrated by Jain et al. (2016) in India. These authors established and calibrated linear relationships between sowing date (and yield) and the Green Chlorophyll Vegetation Index (GCVI) derived from SkySat images, which allowed them to estimate sowing dates (and wheat yields) with an R^2 of 0.41-0.62 and an RMSE of 6.68-12.41 days. However, a downside of CubeSat constellations is the frequent inconsistency in data collected by different satellites in the constellation (Houborg and McCabe 2016). Such inconsistencies may limit the accuracy of surface reflectance-based applications such as estimation of vegetation indices, and could hinder the use of CubeSat satellites to monitor changes on the Earth surface.

In this chapter, a new and robust method is proposed to use CubeSat satellites, which operate in the optical range of the electromagnetic spectrum, to estimate sowing dates at the field scale via detection of no-tillage operation. The methodology used Planet’s PlanetScope satellite images to detect sowing dates of (i) small experimental wheat fields sown across Australia in a wide range of soils and weather conditions, and (ii) in all fields of a typical commercial farm in Victoria.

2.3. Methodology

Sowing dates were detected at a field scale through unsupervised change detection using Planet's CubeSat data. Our approach is based on the widely-used no-tillage farming practices. When a farmer sows, the surface changes its color or spectral response, usually to a darker color depending on the soil type. These changes are mainly caused by the newly exposed soil, which was just uplifted to the surface and by the mixing of last season's crop residues with the soil. After discussion with experts, we assumed that when no-tillage farming practices are implemented, the first detectable field-scale change in the surface after the harvest is the sowing. Based on this, a robust simple change detection method to detect sowing dates was developed, which can be applied over a wide range of soil types, atmospheric conditions and sensors, regardless of crop types.

2.3.1. Study fields and regions

Our approach to directly detecting sowing dates using high spatio-temporal resolution CubeSat data was first tested using the Australian National Variety Trials (NVT) testing fields. The NVT is an Australian national program (www.nvtonline.com.au). In order to evaluate the capability of the new approach to detect sowing dates in different conditions, 16 NVT fields located at eight sites were chosen across Australia from the 2017 growing season (Figure 1). The studied field trials were selected to be located in the different main Australian wheat growing regions, and have diverse management practices, soil types, and weather conditions (Table 1). The average size of the NVT fields analysed in this study was 0.27 ha (about 70 m X 40 m).

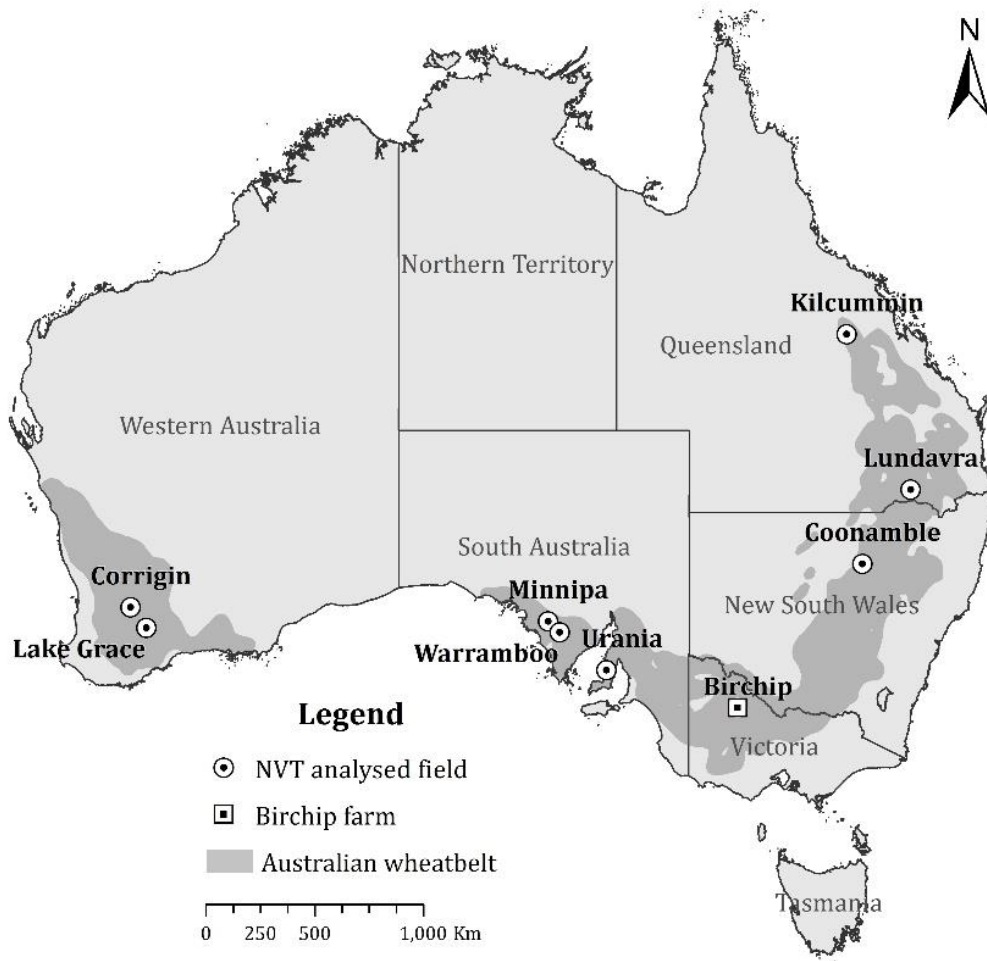


Figure 1. The spatial distribution of the National Variety Trials (NVT) testing fields used in this study and the location of the Birchip farm. The dark grey area corresponds to the Australian wheatbelt (based on Chenu et al. 2013).

In the second stage, the methodology was tested for 50 fields of a commercial farm located near Birchip, Victoria (Figure 1, Table 1) in 2017. The farm corresponds to an average Victorian farm. It is 6,400 ha in size (the average field size is 116 ha), and grows a large variety of crops including wheat, barley (*Hordeum vulgare*), canola (*Brassica napus*), lentils (*Lens culinaris*), oats (*Avena* spp.), vetch (*Vicia sativa*) and peas (*Pisum sativum*) across 55 fields. The studied sowing period in this farm spanned from late-March to mid-June (Figure 2).

Table 1. Summary of the sowing dates, soil types and annual weather statistics of the sites analysed in this study. Soil types were sourced from the Digital Atlas of Australian Soils Science (2000) and climate characteristics from the Australian Bureau of Meteorology (BOM) (2018).

Site	State	Sowing date	Soil type	Annual rainfall (mm)	Annual mean max/min temp (°C)	Mean number of days of rain (≥ 1 mm)	Annual mean number of clear days during the sowing season (Apr-Jun)
Kilcummin	Queensland	10/04/2017	Dark brown cracking clays	552	29.6 / 15.4	42	26
Lundavra	Queensland	23/05/2017	Hard alkaline brown soils	576	27.0 / 12.7	53	19
Lake Grace	Western Australia	28/05/2017	Hard alkaline yellow soils	345	23.5 / 10.4	52	25
Coonamble	New South Wales	29/05/2017	Grey clays	546	26.5 / 11.7	50	26
Corrigin	Western Australia	1/06/2017	Loamy yellow earths	373	23.8 / 10.0	59	25
Urania	South Australia	2/06/2017	Brown calcareous earths	506	21.8 / 11.3	79	24
Minnipa	South Australia	16/06/2017	Calcareous loamy earths	279	24.7 / 11.2	51	24
Warrambo	South Australia	19/06/2017	Calcareous loamy earths	313	25.2 / 9.3	60	34
Birchip	Victoria	17/3/2017 - 13/6/2017	Calcareous loamy soils	374	22.9 / 9.3	59	37

2.3.2. Imagery

Images were sourced from PlanetScope CubeSat satellites data. PlanetScope is a satellite constellation comprising multiple launches of groups of individual CubeSat 3U form factor (10 cm by 10 cm by 30 cm) satellites (Planet Team 2018). A complete PlanetScope constellation consists of about 120 satellites, which have the capability to image all of the Earth's land surface on a daily basis. The PlanetScope satellites have four spectral bands Blue (455 – 515 nm), Green (500 – 590 nm), Red (590 – 670 nm) and NIR (780 – 860 nm) with a Ground Sampling Distance (GSD) of 3-4 m at the nadir and positional accuracy of <10 m RMSE (Planet Team 2018). Planet's PlanetScope constellation operates in two different orbits, the International Space Station (ISS) orbit and in a Sun Synchronous Orbit (SSO). The PlanetScope satellites, which operate at ISS, are deployed at an orbit altitude of 400 km (51.6° inclination), have a variable equatorial crossing time and a limited lifetime of about 1 year. Those on the SSO are deployed at orbit altitude of 475 km (~98° inclination), with an equator crossing time of 9:30 – 11:30 am (at local time) and expected 2-3 years lifetime (Houborg and McCabe 2018b; Planet Team 2018). The PlanetScope analytic Ortho Scene Products (Level 3B) are used in this study, which were provided at a spatial resolution of ~3 m and with sensor-

specific conversions to at-sensor top of Atmosphere radiance based on limited pre-launch calibration coefficients (Houborg and McCabe 2018b; Planet Team 2018).

This study used a total of 205 (78 for the NVT analysis and 127 for the farm analysis) Planet CubeSat satellites scenes acquired by 78 different satellites (61 for the NVT analysis and 30 for the farm analysis). For each NVT location, Planet CubeSat images were used, which covered a period of ten days before and ten days after the reported sowing dates (depending on image availability) (Table 2). On average, each field was covered by 9.5 satellite images throughout the testing period, with an average of two days gap between the images. For the farm analysis, images from 17/04/2017 to 20/06/2017 were used (Figure 2), with a five-day median gap between cloud-free CubeSat images (Table 2). The 2017 growing season was chosen since in the previous seasons, the PlanetScope constellation was not yet complete and the revisit time of each location was insufficient to accurately detect sowing dates.

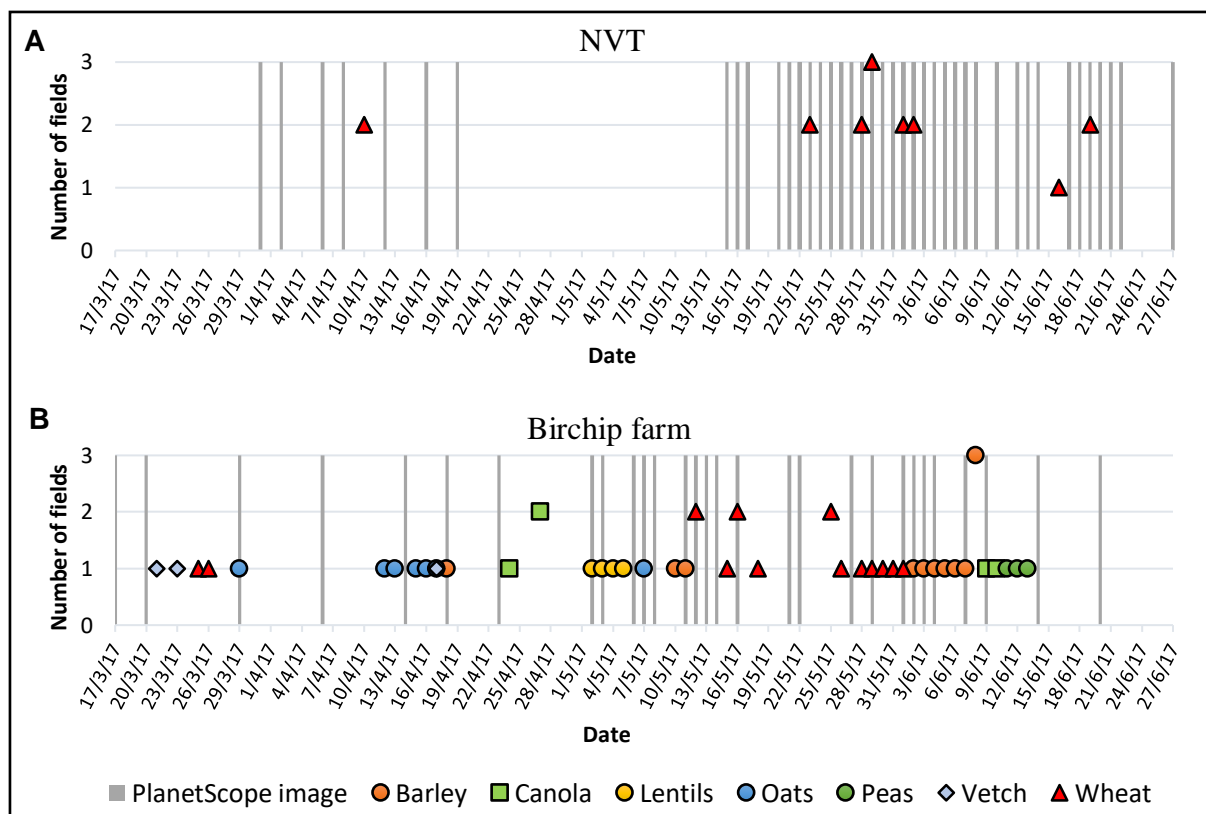


Figure 2. Number and timing of fields sown at the studied NVT (sub-figure A) trials and Birchip farm (sub-figure B) in 2017, together with the timing of available PlanetScope images (vertical grey lines). The crop type and number of fields sown each date for both NVT and Birchip farm.

Table 2. Number of images used to detect the sowing dates at each site and the median PlanetScope cover gap between the images over the period tested. The Birchip farm had 31 daily image mosaics, made out of 127 PlanetScope scenes.

Site	Number of daily satellite images	Median satellite cover gap (days)
Kilcummin	8	3
Lundavra	12	1
Lake Grace	10	2
Coonamble	13	1
Corrigin	9	2
Urania	9	1
Minnipa	10	2
Warrambo	7	3.5
Birchip	31 (mosaics)	5

2.3.3. Semi-automated sowing detection

The semi-automated method developed to detect sowing dates based on spectral changes between successive CubeSat images of a field (Figure 3) requires three inputs: **1**) two satellite images acquired on different dates ($Image_{t_1}$ and $Image_{t_2}$); **2**) the field boundaries; and **3**) a threshold value to detect changes between images, which has to be manually input for each set of images. This threshold corresponds to the percentage of values in the processed image histogram that will be used as a factor to determine a change in the image. First, the images from the same day (which may not necessarily be acquired by the same satellite) were mosaicked. Since PlanetScope satellites tend to have cross-sensor inconsistency (Houborg and McCabe 2016), the pixel values in the mosaicked image commonly suffer from the lack of uniformity. Next, each mosaicked image is clipped using the fields' boundaries; in the following step, image-processing techniques used to detect the sowing dates.

2.3.4. Change detection

Many change-detection methods have been developed to identify changes on the Earth's surface using remote sensing (e.g. Mas 1999; Bruzzone and Prieto 2000; Lu et al. 2004; Jin et al. 2013). Change detection using remote sensing images basically involves the comparison of two images of the same area, which were acquired at different times (Byrne et al. 1980).

In this study, a Principal Component Analysis (PCA) was performed separately for each image to detect the change in the fields' surface caused by no-tillage sowing (Figure 3). PCA is useful for detecting decorrelations between images (Du and Fowler 2007) and also for separating an underlying systematic data structure from noise (Wold et al. 1987). Four principal components (PC_1 , PC_2 , PC_3 and PC_4) were output from all the pixels of the mosaic

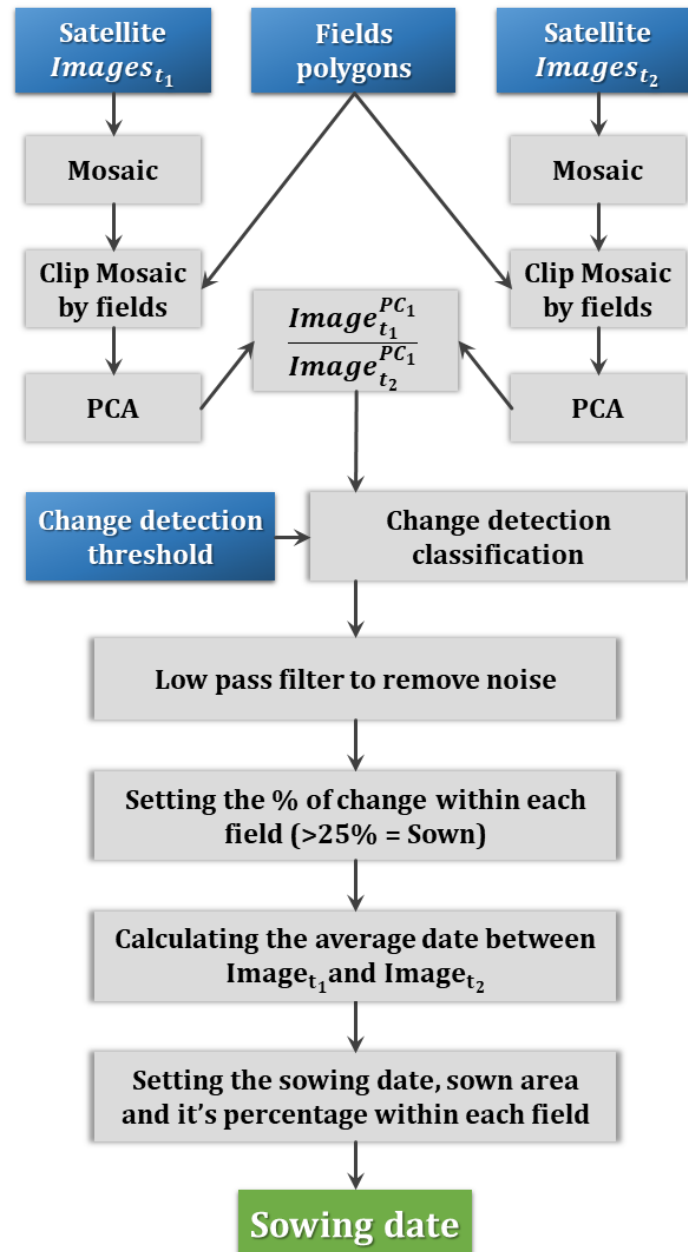


Figure 3. Explanatory diagram of the sowing-detection workflow

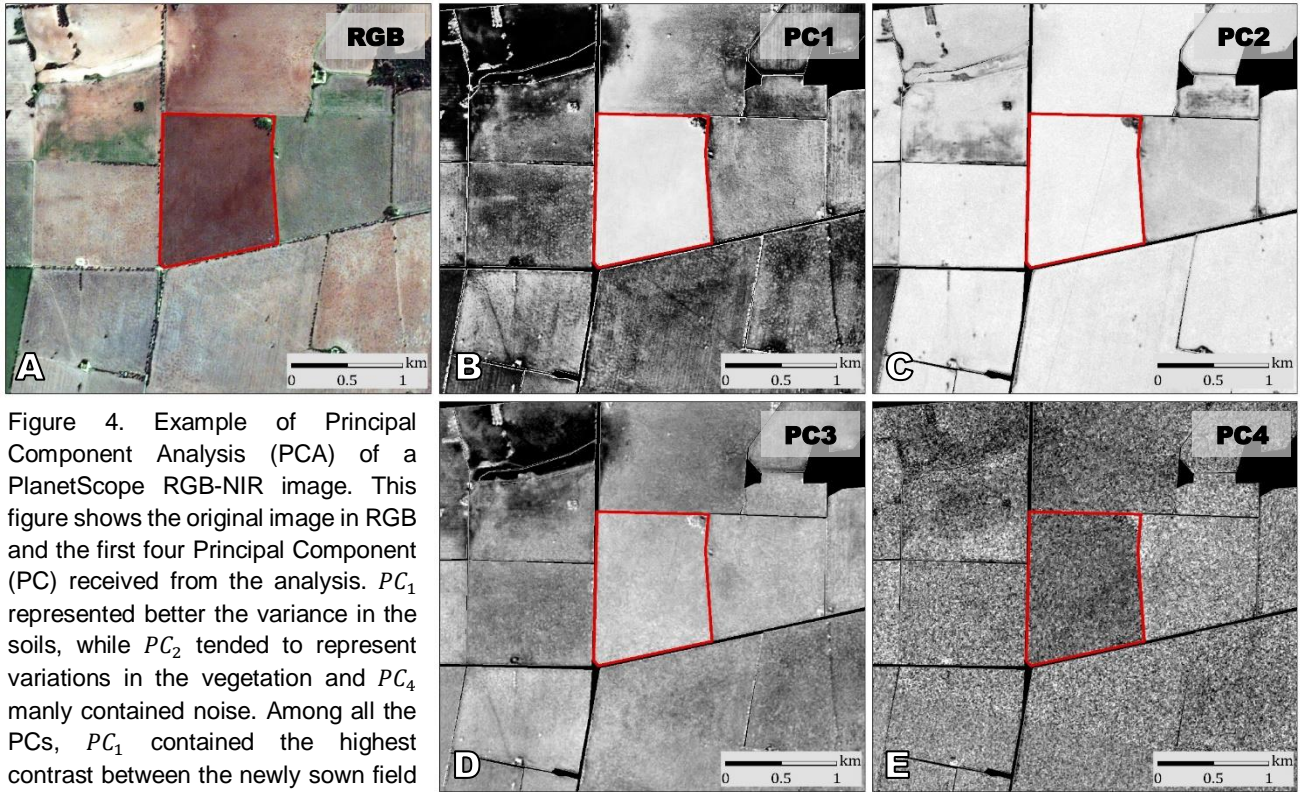


Figure 4. Example of Principal Component Analysis (PCA) of a PlanetScope RGB-NIR image. This figure shows the original image in RGB and the first four Principal Component (PC) received from the analysis. PC_1 represented better the variance in the soils, while PC_2 tended to represent variations in the vegetation and PC_4 mainly contained noise. Among all the PCs, PC_1 contained the highest contrast between the newly sown field and the other fields.

that intersected with the fields' polygons. In this study, PC_1 was systematically the best PC to represent the variance in the soils, while PC_2 tended to represent variations in the vegetation, PC_3 represented clouds and shades variance, if they were present, and PC_4 mainly contained noise (Figure 4).

Afterwards, PC_1 s from two consecutive images ($Image_{t_1}^{PC_1}$ (early image) and $Image_{t_2}^{PC_1}$ (later image)) were used to detect whether any new sowing occurred between these two dates by using the following equation:

Equation 1

$$Change = \frac{Image_{t_1}^{PC_1}}{Image_{t_2}^{PC_1}}$$

where $Image_{t_1}^{PC_1}$ is the first principal component of the earlier satellite image and $Image_{t_2}^{PC_1}$ is the first principal component of the later satellite image. The ratio of the two resulted in a one-band raster that represents the magnitude of change between the two images at the pixel level (Figure 5B). A pixel value closer to 1 indicates no change. The histogram of the ratio image generally exhibits a normal distribution (bell-shaped) around the value of 1 if there is no change across the image. When changes occur, the histogram would show a multimodal distribution, with several distributions combined, i.e. would have multiple peaks and valleys.

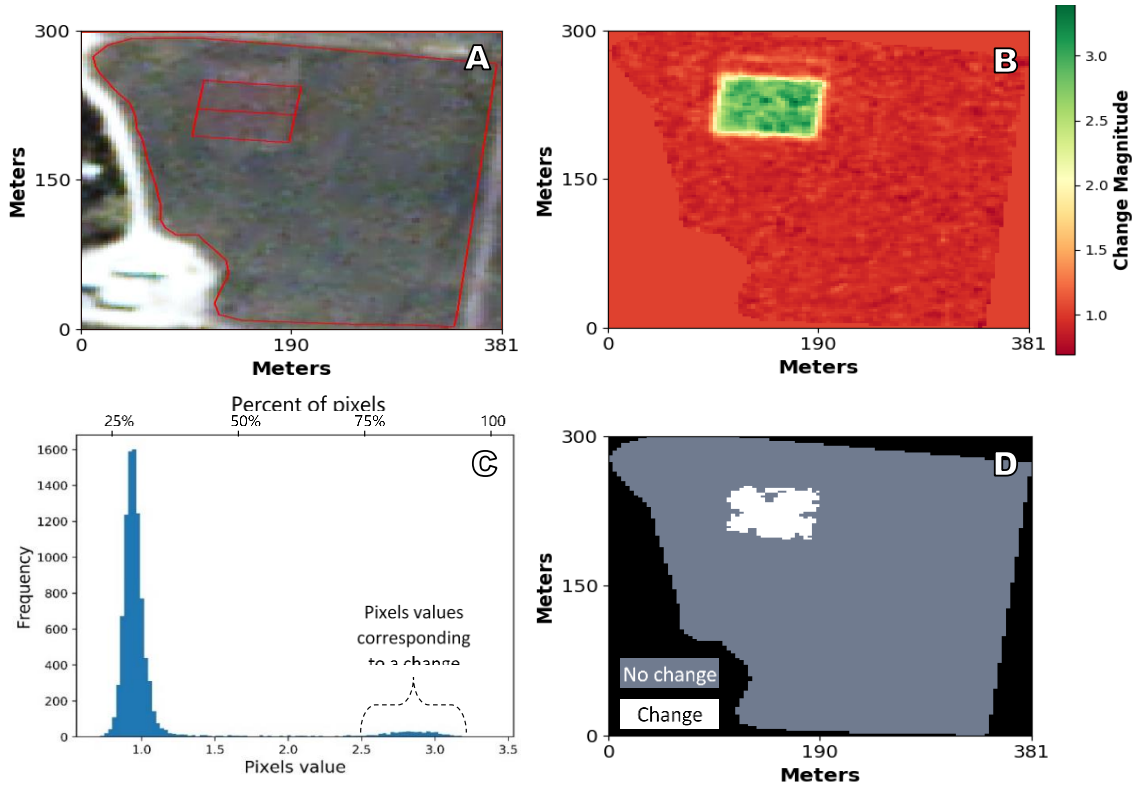


Figure 5. Use of the image histogram to choose the threshold value for change detection. **A)** True colour Image of a NVT wheat field after sowing (the red rectangles are the experimental field boundaries). **B)** Image resulting from dividing $Image_{t_1}^{PC_1}$ by $Image_{t_2}^{PC_1}$. **C)** Histogram of the image in B. **D)** Results of change detection based on a threshold set to 2.5 (as identified in C).

By examining the distributions rather than the one representing “no change”, a threshold pixel (ratio) value to identify changes were determined. For example, in Figure 5C, there are two distributions of the pixel values: one is a normal distribution around 1 and the other is a normal distribution around about 2.8. The second distribution represents the changes. The left end of the second distribution is set as the threshold value. A pixel with a value equal to or greater than this threshold is considered as “change”. The threshold values may change from scene to scene due to cross-sensor inconsistencies among the different satellites in the constellation. In our case study, threshold values range between 1.2 and 3.

The classified image at this stage (Figure 6D) contains a lot of noise, which was eliminated in two stages. First, a low-pass filtering method is used (Figure 6E) to smooth and remove speckle noise from the image (Al-Amri et al. 2010). Second, only if the area detected as changed covers more than 25% of the field area, is considered sown. This second noise cleaning stage is used to filter other changes detected on the earth surface that could have

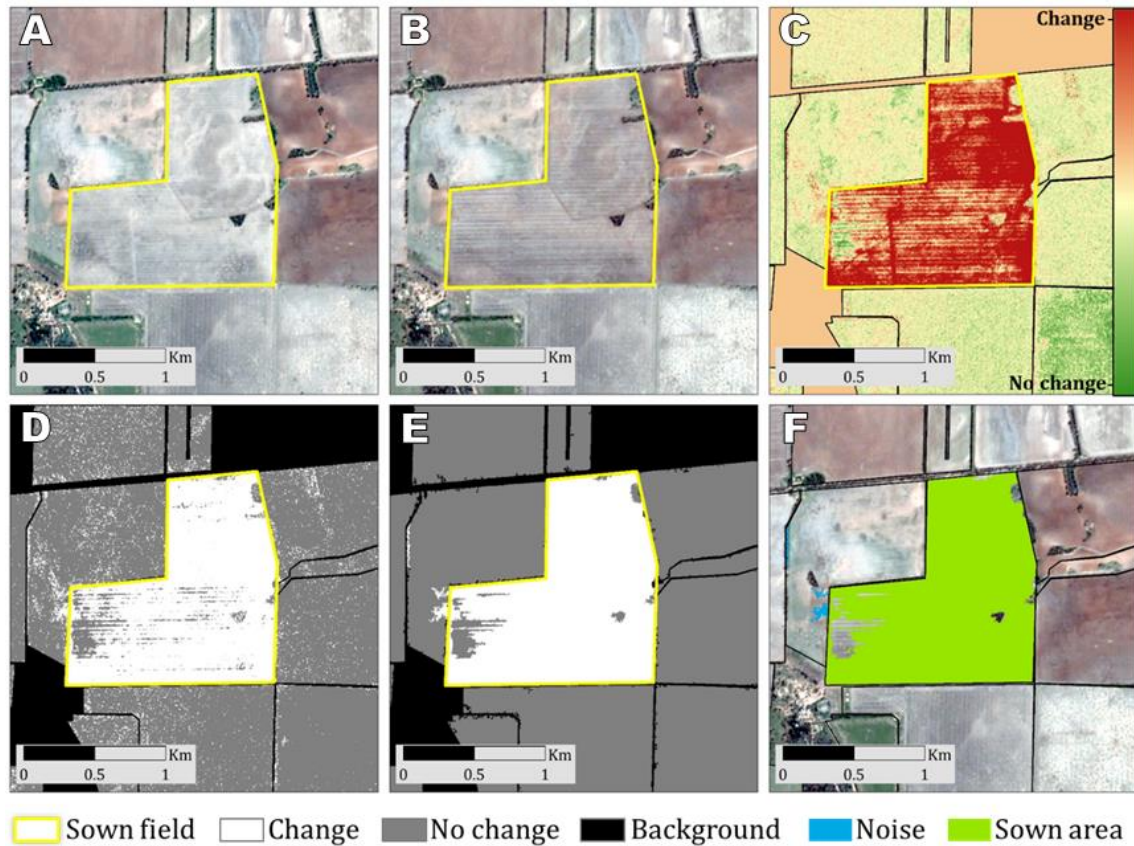


Figure 6. Example of sowing detection. This figure illustrates the sowing detection of two sown fields (boundaries in yellow) using a pair of images acquired on 6/5/2017 ($Image_{t_1}$) and 7/5/2017 ($Image_{t_2}$) (subfigures A & B). Subfigure C is the resulting image of dividing $Image_{t_1}^{PC_1}$ by $Image_{t_2}^{PC_1}$. A change between the images resulted in high values (red) and negligible changes resulted in low values (green). In subfigure D, the image is classified to identify pixels that changed (in white). A low-pass filter used to clean the image from noise (subfigure E). Finally, the semi-automated method outputs and exports the area sown between $Image_{t_1}$ and $Image_{t_2}$ (in green) (subfigure F).

changed over time, e.g. manmade (e.g. infrastructure construction) or caused by a natural phenomenon (e.g. water ponding on the surface). Note that sowing large fields (e.g. more than 250 ha) may take a few days, and it is common for farmers to not sow an entire field in a day. In such case, the latest date in which more than 25% of the field's area detected as sown, assigned as the final sowing date for that field. Since the method identified which fields were sown between $Image_{t_1}$ and $Image_{t_2}$, we know that the sowing date must be within this timeframe. Therefore, the middle date between the earlier and later images assigned as the final sowing date to each field. The final output of the semi-automated sowing detection technique is a shapefile that contains fields or parts of fields detected as sown (as shown in Figure 6), the estimated sowing date, the actual sown area and the percentage within each field that was sown in the time period between $Image_{t_1}$ and $Image_{t_2}$.

2.4. Results

A very high correlation ($R^2 = 0.99$) was found between reported and detected sowing dates in both 16 NVT fields and the Birchip farm analysis (Figure 7), which span over a wide range of soil types, atmospheric conditions, crop types and PlanetScope sensors (Table 1, Figure 2).

The proposed semi-automated methodology detected 100% of the sowings in NVT fields and 80% in the farm fields (Table 3). There was only an average -0.6 day gap (0.0 day gap for the median) between the estimated and the actual sowing dates (RMSE = 0.9 days). For the Birchip farm, the average and median gaps were -0.1 and 0.0, respectively (RMSE = 1.9 days) for the fields detected as sown (Table 3).

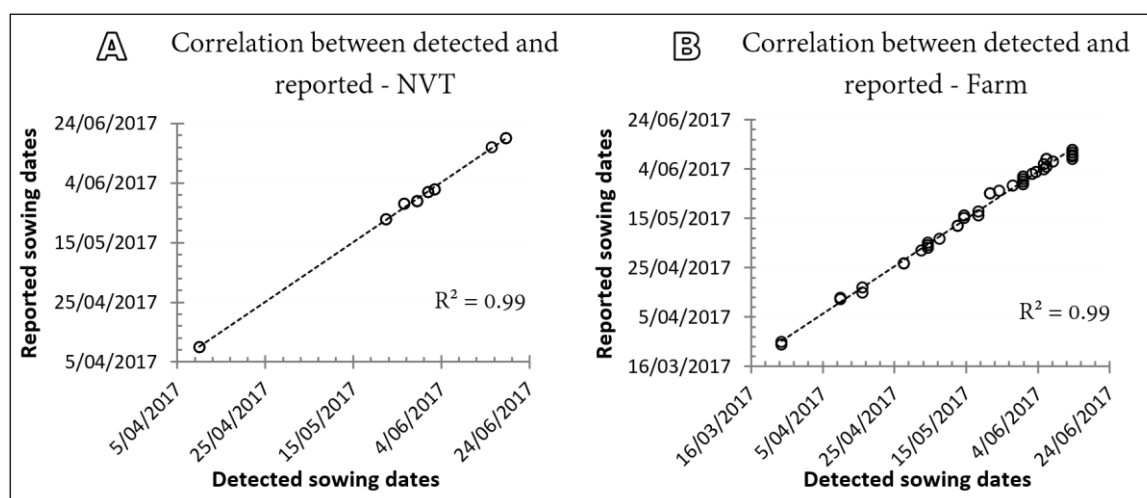


Figure 7. Correlation between the reported and detected sowing dates of the 16 NVT fields (A) and the 50 farm fields (B). The dashed line marks the trend line.

Table 3. Summary of the sowing detection analysis conducted over 16 NVT fields and 50 fields of the Birchip farm. This table shows the accuracy of the detection and the gap between the reported and the detected sowing dates.

	NVT	Birchip farm
Total number of fields	16	50
Sowing detected (No. of fields)	16 (100%)	40 (80%)
Failed to detect (No. of fields)	0 (0%)	10 (20%)
False detection (% of the detections)	1.4%	9.1%
RMSE (days)	0.9	1.9
Mean error (days)	-0.6	-0.1
Median error (days)	0.0	0.0
Standard deviation (days)	0.7	1.9

During the analysis, a few false detections were found (in only 5% of the algorithm implementations), where a field was incorrectly classified as sown by the sowing detection algorithm in the time period between $Image_{t_1}$ and $Image_{t_2}$. Most of the sowing detection analyses (95%) ended with zero false detections, however in some cases the presence of clouds in the images was translated as a change that covers most of the field and therefore was determined as sown. In the NVT analysis, only 1.4% of the detections were false detections. False detection occurred more often in the Birchip farm, when 9.1% of the detections were false detections. This higher rate of false detection was mainly due to cloud contamination in the 23/04/17 image, which caused five false detections. When excluding this image from the analysis, only 6.1% of the detections were false detections.

Detection failures, i.e. no detection of sowing, which occurred in 20% of the farm's fields, were partly due to cloud cover (50% of the cases) and the large time gap between $Image_{t_1}$ and $Image_{t_2}$ (33% of the cases). Most successful detections took place when the gap between the sowing and the satellite image date was under four days (Figure 8). In these conditions, the successful detection rate was 91% of the fields, however, when this gap is larger than five days the percentage of the successful detections drops to 62% (Figure 8). Although the distribution of the number of samples in Figure 8 is not constant along the X-axis, and the

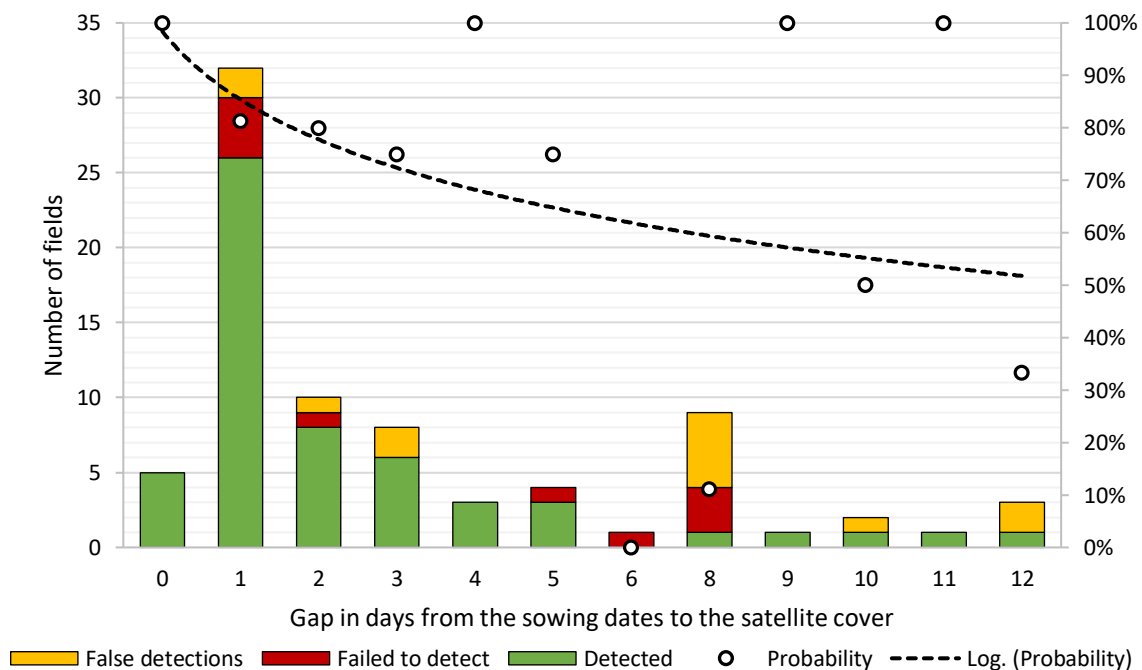


Figure 8. Distribution of the 56 successful (in green), 10 unsuccessful (in red) and 10 false (in yellow) detections of the sowing dates performed in this study (left Y-axis). This graph also show probability to achieve a successful detection (black points) relatively to the gap in days from the sowing date to the satellite cover (right Y-axis). As this gap increases, the likelihood to detect the sowing date decreases in addition to higher chances for false detections (black dashed line).

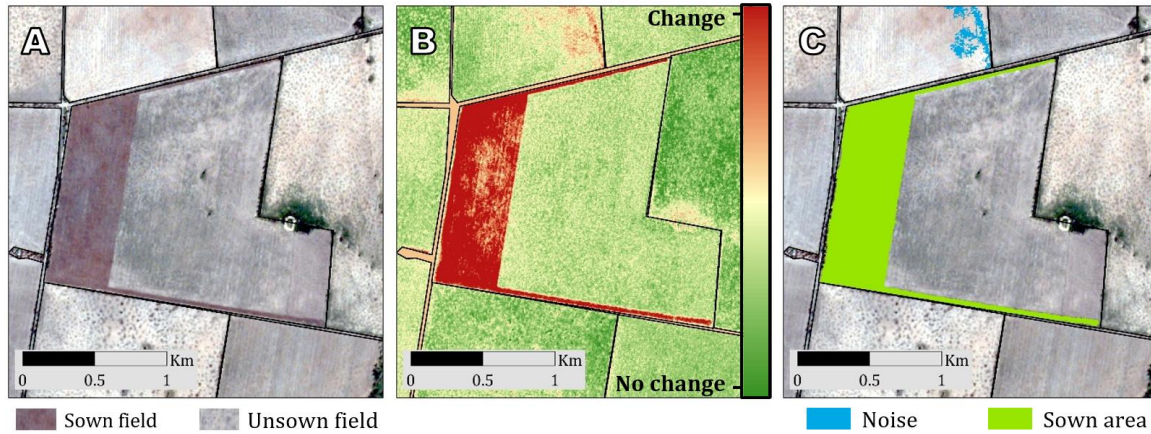


Figure 9. Sub-field sowing detection. This figure illustrates the sub-field sowing detection capabilities. Subfigure **A** shows the RGB satellite image of a partly sown field (the area in brown). Subfigure **B** is the resulting image of dividing $Image_{t_1}^{PC_1}$ by $Image_{t_2}^{PC_1}$. A change between the images resulted in high values (red). Subfigure **C** shows the field's area detected as sown (in green).

probability data points are scattered, the log probability curve shows a clear decreasing trend. The reason for a lower probability to achieve a successful detection when the gap in days from the sowing dates to the satellite cover is larger is due to the fading in contrast between the appearance of the unsown and the sown soil surfaces. These were caused mainly by the rainfall events, wind erosion and the loss of the moisture in the newly exposed soil, which makes this contrast fade with time.

Although our semi-automatic sowing detection method was designed to determine the sowing date when a field was sown in more than 25% of its area, it has the capacity to identify sub-field changes. For instance, sowing of a small proportion of the field was detected at the Birchip farm (Figure 9), as the grower did not have time to complete the sowing of the entire field at once, which is relatively common in this region.

2.5. Discussion

2.5.1. Detecting no-tillage sowing with CubeSat satellites

Remote sensing has the potential for mapping sowing dates across the globe. To our knowledge, no methodology has been published on how to detect sowing dates directly from satellite images. While, until recently, such detection could not happen due to technological limitations (low spatial resolution and revisit time of the satellites), our approach overcomes these limitations by using CubeSat images to detect no-tillage sowing, which is becoming the main sowing practice in major cropping regions worldwide (Kassam et al. 2015). When a farmer uses a no-tillage seeder without performing any pre-sowing cultivation, the surface

changes its color, usually to a darker color. These changes can sometimes appear as minor and they cannot always be seen with a naked eye (e.g. Figure 5A), given the similarity of the spectral signature of crop residues and soils in the visible and near infrared. Their spectra differ within the SWIR spectrum (Daughtry 2001), but CubeSat satellites like Planet's PlanetScope, which are attractive given their high spatio-temporal coverage, commonly do not operate in such long wavelengths. Hence, they cannot be used to detect the changes caused by no-tillage practices using SWIR-based tillage indices.

Satellite-based change detection is commonly based on the analysis of differences between two images using band-ratios or classifications derived from the original image bands. However, such approaches are very limited for CubeSats such as PlanetScope given (i) cross-sensor inconsistencies, as the PlanetScope satellites are not identical, so that each CubeSat can have a unique spectral response (McCabe et al. 2017), and (ii) different atmospheric conditions during the acquisition. Different atmospheric conditions affect the amount of radiation that reaches the sensor, causing inter-band and inter-pixel differences within the two images (Byrne et al. 1980). Unlike sensors such as Landsat, MODIS or Sentinel-2, which traditionally developed and launched by space agencies, Planet's CubeSats suffer from relatively low signal-to-noise ratio. As a result, any time-series data from those CubeSat satellites is likely to have inconsistent reflectance signals, even when applying atmospheric correction techniques (Houborg and McCabe 2018b). The main challenge when conducting change detection using these types of satellites is to separate the noise captured by the sensors, which causing dissimilarities among the different satellites in the constellation, from the actual change that happened on the Earth surface between the earlier and later images. Therefore, the reliability of detecting changes, while analysing images acquired by different CubeSat images such as PlanetScope constellation, is very limited when using the original bands (Houborg and McCabe 2016) and indices. That adds the challenges of addressing the varying atmospheric conditions to the change detection analysis. To overcome these limitations the PCA was implemented, which is a well-known statistical technique for multivariate data analysis (Wold et al. 1987; Pohl and Van Genderen 1998; Du and Fowler 2007). PCA extracts the dominant patterns in the dataset (Wold et al. 1987) and expresses this information as a set of new variables called Principal Components (Abdi and Williams 2010). The first Principal Component (PC) band in the image contains the largest fraction of data variance and the second PC band contains the second largest data variance, and so forth. The last *PC* bands seem noisy due to the very little variance they contain, which mainly represent the noise in the original spectral data (Rajendran et al. 2016). PCA is useful for detecting decorrelations between images (Du and Fowler 2007) and also for separating an underlying systematic data structure from noise (Wold et al. 1987). In remote sensing PCA is

commonly used for land-cover/land-use change detection, classification and anomaly detection (e.g. Byrne et al. 1980; Du and Fowler 2007; Celik 2009; Abdi and Williams 2010; Dronova et al. 2015; Gil-Yepes et al. 2016). Furthermore, PCA-based change-detection is effective regardless of the differences in spectral, spatial, and radiometric resolutions of the multi-sensor satellite data (Deng et al. 2008). In this study, a PCA used to extract the dominant patterns in the images, and to conduct detection of changes between the main patterns in each pair of images. This method allowed us to have a comparable measure between the images to detect changes in the fields' surface, even when images acquired by 78 different PlanetScope satellites with different atmospheric conditions were used. This study found that using PC_1 is consistently superior over the other PCs to detect changes caused by sowing, even when processing images acquired by CubeSats with a limited pre-launch sensor-specific calibration (Houborg and McCabe 2018b; Planet Team 2018).

2.5.2. Unprecedented levels of sowing detection

The semi-automated sowing detection methodology achieved high performances with 85% detection (56 out of 66) of the sown fields in both NVT and Birchip farm fields (Table 3). The estimated sowing dates were very highly correlated with the actual sowing dates ($R^2 = 0.99$; Figure 7), with a median gap of 0 days between the reported and estimated dates, and a low RMSE (0.9 and 1.9 days for the NVT fields and the Birchip farm, respectively; Table 3). The results of our semi-automated sowing detection methodology, thus exceeds the performance and accuracy of the methods used by previous studies, which used different spaceborne sensors with diverse resolutions and in various electromagnetic spectral regions (optical, fluorescence, and radar). Studies that used the green-up detection approach, succeeded in estimating sowing dates with a RMSE ranging from 6.68 to 12.41 days, and R^2 ranging between 0.15 and 0.90 between their estimations and official regional statistics (e.g. Lobell et al. 2013; Jain et al. 2016; Urban et al. 2018). Their main limitation is to use a fixed period (days to weeks) to backcast the emergence of the crops, while this period varies in particular with crops species and weather conditions (Marais Sicre et al. 2016; Urban et al. 2018). Due to the spatial and temporal variability of the environmental and crop characteristics, such a linear relationship between sowing date and crop emergence are local and would be difficult to establish without prior knowledge. Our approach on the other hand, does not depend on local conditions (soil types, rainfall or crop types (Table 1, Figure 2)), does not require the identification of a fixed number of days to backcast the emergence and is therefore more suitable for diverse environments, as long as no-tillage is practiced.

The chances for a successful sowing detection were higher in the first four days following the sowing (91% successful detections). In our study, nearly half of the detected sowing occurred

only one day after the sowing (Figure 8). Therefore, using satellites with a temporal resolution larger than four days such as Sentinel-2 or Landsat, may not achieve as high accuracy of detection as in this study. In addition, we expect that as the revisit time of CubeSat-based constellations will increase over the years, the probability for a successful detection using our approach will improve accordingly.

Detecting sowing dates at the field level requires both high spatial and temporal resolution remote sensing data. Previous studies commonly worked at a region, county or state scale and did not directly detect the sowing dates, as they often used low spatial resolution satellite data such as MODIS (250 m – 1 km) to estimate sowing dates at aggregated-level scale of multiple fields (e.g. Sakamoto et al. 2005; Lobell et al. 2013; Urban et al. 2018). These limit the accuracy of the estimations and the applications derived from these data. Additionally these were unable to distinguish crops from natural vegetation (Urban et al. 2018). In contrast to MODIS, which also has daily revisit time, CubeSat satellites such as PlanetScope have the ability to map individual farms or fields and to represent the large heterogeneity of management practices within and across fields (Jain et al. 2016). Our methodology allowed the detection of field and even sub-field level changes caused by sowing (Figure 9). This capability can thus potentially be used to monitor progress in sowing even at the sub-field scale.

Satellite-based studies on yield estimation traditionally used officially reported sowing dates (e.g. Sakamoto et al. 2005; Marinho et al. 2014; Jin et al. 2016) or sowing dates based on farmers' surveys (e.g. Ortiz-Monasterio and Lobell 2007; Jain et al. 2016). High-resolution information of sowing dates in regional datasets could be used to reduce the uncertainty of regional yield prediction using crop simulations (Mathison et al. 2017). However, the existence of such datasets is currently scarce, since only few censuses collect this information or make it publically available. Furthermore, when released, this data is often shared at aggregated regional scales (Urban et al. 2018). Establishing a large-scale database of field/farm management practices is time consuming, expensive, relies heavily on manual labour and requires some training (Jain et al. 2016). The methodology presented in this study, can be used to produce low-cost field level statistics with unprecedented spatial coverage, which can be updated on a near-daily basis. In addition, statistics at the field-level may improve our understanding of the reasons of yield gaps, and could assist to identify low-performing fields that need to be treated in order to increase their productivity (Jain et al. 2016). Our CubeSats-based change detection methodology also has the potential to be implemented to monitor rapidly changing environments in other disciplines such as for geomorphological and natural disaster/hazards studies. For example, assembling information on tillage is important for soil

erosion modelling, despite the impracticability of collecting such information using ground surveys (Mc Nairn et al. 1998).

2.5.3. Limitations and perspectives

Despite its very high accuracy, our sowing detection methodology has limitations. Although the approach presented here can overcome the presence of a few small-scattered clouds in the scene as a field is considered as sown only if the area detected as changed exceeds 25% of the field's area, most of the failed and false detections (50% of the failed detection and 60% of the false detections in the Birchip farm) were caused by the presence of clouds. The presence of cloud shadows in the images on the other hand was found to have less impact on results. This is due to the fact that the areas covered by cloud shadows produced low values in the histogram of the ratio image, which are not high enough to be considered as "changed". I believe that it is possible to overcome some of the failed and false detections caused by the presence of clouds and cloud shadows in the images (depending on the percentage of the coverage in the image) by masking the clouds and cloud shadows out of the images (Kolecka et al. 2018). There are a number of automated cloud and cloud shadow detection methods presented in the literature, e.g. Fmask (Function of mask) (Zhu and Woodcock 2012; Frantz et al. 2015). However, these methods commonly use SWIR bands as they are offering a high contrast between cloudy and cloud free pixels (Sedano et al. 2011). PlanetScope satellites do not operate in the SWIR and therefore future studies will investigate how to adapt the existing cloud and cloud shadow detection methods to these new satellites. Another approach that should be investigated is to separate cloud-based changes and sowing-based changes by analysing the shape of their appearance in the image, as clouds ought to be quite different from the mostly straight-line edges of sowed areas. Another way to potentially overcome this limitation would be to integrate the Synthetic Aperture Radar (SAR) data into the change detection analysis. In contrast to optical sensing, SAR sensors have the advantages of all-weather capabilities, and are not affected by the presence of clouds in the imagery. While SAR data may facilitate change detection in cloudy conditions, most of the spaceborne SAR systems do not have both the required temporal resolution and the spatial resolution. There are few SAR systems that operate at a high spatio-temporal resolution, but the high costs of such SAR data will probably make an operational sowing detection application impractical.

Another drawback is imposed by mosaicking of same-day images acquired by different satellites. This study noticed that when we mosaicked images from different CubeSats in the constellation, they provide a good visual coverage of the study area. However, using these mosaics to calculate the ratio-images prevented us from setting a constant value for histogram

thresholding, even when both images were acquired on the same day. Therefore, in the case of a field separated into two parts in two images, the same threshold cannot be used for both parts of the field. But it was very rare that a sown field was present in two images.

Our method has been developed for the no-tillage sowing practice. While this sowing practice is now common worldwide for rainfed crops (Kassam et al. 2015), our method is currently limited to regions where no-tillage sowing is being implemented. Yet, I believe that this approach could also be valid for other types of sowing. This potentially can be done by detecting the last field scale change on the soil surface before the seedling emergence. Once the crops can be detected by a vegetation index (e.g. NDVI) similar to the green-up approach (Lobell et al. 2013), we can estimate the date when the last field scale change was detected, which is likely to be the sowing date.

Currently our sowing detection method is still semi-automated, as users need to identify the pixel values that correspond to a change in the image histogram, in order to determine the threshold to classify pixels as 'changed' or 'not changed' (Figure 5). Additionally, our method uses the polygons to identify field locations, which were known in this study. In order to fully automate the process, the two inputs (pre-determined threshold and the polygons of the fields) need to be identified automatically. This potentially can be achieved by implementing machine-learning techniques (Toulouse et al. 2016) to identify the pixels that represents change in the histogram. It is proposed that future studies should explore the ability to use clustering algorithms such as K-Means, Fuzzy K-Means and the self-organizing map (SOM) to perform automatic classification (Kanungo et al. 2002; Zhong et al. 2006; Gonçalves et al. 2008; Kussul et al. 2017). Automation of this process is necessary in order to turn this prototype into an operational application. The location and boundaries of the fields, expressed as geo-located polygons can be also identified using remote sensing by implementing classification and segmentation techniques (Van der Sande et al. 2003; McCarty et al. 2017). Previous studies showed that time-series images could successfully be used to identify crop fields, classify between irrigated and non-irrigated crops (Jin et al. 2016; Azzari and Lobell 2017; Zhang et al. 2018), classify crop types (Van Niel and McVicar 2004) and to preform cropland area segmentation (Gallego et al. 2014; McCarty et al. 2017). Perhaps the main difficulties in the context of our method will be the separation between rainfed crops from the natural vegetation and the accurate delineation of the field boundaries. However, as rainfed crop fields in developed countries commonly tend to be large in size (> 10 ha) and in a square shape, I believe that implementing existing methods will probably achieve high accuracy.

Despite the inconsistency of spectral data, CubeSats are only starting to unleash their potential. I strongly believe that as more advanced CubeSat constellations are deployed into

space, the number of CubeSats-based applications at the field or farm level will increase and their efficacy for detecting sowing and harvesting dates will improve.

2.6. Conclusions

In this study, an innovative semi-automated sowing detection methodology was developed, based on Planet's PlanetScope data to detect changes on field surface caused by sowing. The method detected 85% of the sown fields with a very high correlation ($R^2 = 0.99$) between actual and estimated dates. Time of sowing was detected with a median gap of 0 days while achieving RMSE of 0.9 and 1.9 days in a national set of data and in a representative commercial farm, respectively. The approach may be used to produce and update near-daily low-cost field level statistics in an unprecedented spatial coverage. To fully automate sowing detection at the field scale, machine-learning techniques could be explored to identify automatically parts of a histogram corresponding to a change in spectrum (due to soil disturbance). Automation will also require the integration of classification and segmentation techniques to detect the fields' location and to extract their boundaries. Methods to account for clouds present in the images would also improve the applicability of the method. We can expect that in the future, the accuracy of this method will increase as more and more CubeSats constellations become operational.

Chapter 3 – Fusion of PlanetScope and Sentinel-2 into daily 3 m LAI

This chapter is a reproduction of the paper “Fusion of Sentinel-2 and PlanetScope time-series data into daily 3 m surface reflectance and wheat LAI monitoring” published in *International Journal of Applied Earth Observation and Geoinformation* by Sadeh et al. (2021), with section and figure numbers changed to fit the thesis structure.

3.1. Abstract

The dynamics of Leaf Area Index (LAI) from space is key to identify crop types and their phenology over large areas, and to characterize spatial variations within growers' fields. However, for years remote-sensing applications have been constrained by a trade-off between the spatial and temporal resolutions. This study resolves this limitation. Over the past decade, the number of companies and organizations developing CubeSat constellations has increased. These new satellites make it possible to acquire large image collections at high spatial and temporal resolutions at a relatively low cost. However, the images obtained from CubeSat constellations frequently suffer from inconsistency in the data calibration between the different satellites within the constellation. To overcome these inconsistencies, a new method to fuse a time series of images sourced from two different satellite constellations is proposed, combining the advantages of both (i.e., the temporal, spatial and spectral resolution). This new technique was applied to fuse PlanetScope images with Sentinel-2 images, to create spectrally-consistent daily images of wheat LAI at a 3 m resolution. The daily 3 m LAI estimations were compared with 57 *in-situ* wheat LAI measurements taken in Australia and Israel. This approach was demonstrated to successfully estimate Green LAI (LAI before the major on-set of leaf senescence) with an R^2 of 0.94 and 86% relative accuracy (RMSE of 1.37) throughout the growing season without using any ground calibration. However, both the Sentinel-2 based estimates and the fused Green LAI were underestimated at high LAI values (LAI > 3). To account for this, regression models were developed, improving the relative accuracy of the Green LAI estimations by up to a further 47% (RMSE of 0.35-0.63) in comparison with field measured LAI. The new time series fusion method is an effective tool for continuous daily monitoring of crops at high-resolution over large scales, which opens up a range of new precision agriculture applications. These high spatio-temporal resolution time-series are valuable for monitoring crop growth and health, and can improve the effectiveness of farming practices and enhance yield forecasts at the field and sub-field scales.

3.2. Introduction

Improving the spatial and temporal estimation of Leaf Area Index (LAI) and monitoring of the crop developmental stage using remotely sensed imagery can inform service providers and growers to facilitate management decisions, formulate policies, and ultimately improve profitability (Pasqualotto et al. 2019; Sun et al. 2019). LAI also plays an important role in crop monitoring and can be used in crop growth models to better predict yield (Clevers 1991; Bøgh et al. 2004; Lobell et al. 2015). One of the common applications for LAI is to provide yield estimations (e.g. Ines et al. 2013; Lobell et al. 2015; Azzari et al. 2017; Sun et al. 2017; Waldner et al. 2019).

LAI is defined as the ratio of one-sided leaf area per unit ground area (Watson 1947) and knowing the LAI of a crop has a wide range of applications. However, monitoring crop LAI by extensive *in-situ* sampling over large areas is expensive, time consuming and consequently impractical (Houborg and McCabe 2018c). Therefore, for decades scientists around the world have attempted to estimate LAI from space (e.g. Pollock and Kanemasu 1979; Wiegand et al. 1979; Chen et al. 2002; Gitelson et al. 2003; Viña et al. 2011; Nguy-Robertson et al. 2014). However, the trade-off between the spatial and temporal resolution typically restricted the use of high spatial and temporal time-series of images for agricultural applications (Waldner et al. 2019).

As crop canopy reflectance is affected by the LAI, as well as by the chlorophyll distribution, canopy structure and the background soil (Gitelson et al. 2005), methods which rely on optical remote sensing to convert surface reflectance data into LAI estimations were developed (Delegido et al. 2015). These methods are commonly classified into two groups (Delegido et al. 2015; Fang et al. 2019; Pasqualotto et al. 2019; Kimm et al. 2020): (i) physically-based retrieval methods, which are based on radiative transfer models (RTM), when the LAI is estimated based on the inversion of these models (e.g. Houborg and McCabe 2018c), and (ii) an empirical approach using either linear or nonlinear regressions with vegetation indices (VIs) as independent variables (e.g. Herrmann et al. 2011; Nguy-Robertson et al. 2014). These two groups of methods have both advantages and disadvantages. The physically-based retrieval methods are more generally applicable, but they are often limited by the nature of canopy structure and thus suffer from the ill-posed problem that may end in an unstable solution and require an a priori knowledge of targeted canopies (Bsaiibes et al. 2009; Delegido et al. 2015). The empirical methods are commonly based on pre-trained relationships between field measured LAI and VIs; they are simple and do not require intensive computation. However, these empirical relationships could only be useful in regions that are similar to those used for

calibration (Bsaiibes et al. 2009; Kimm et al. 2020) and are less reliably applied for multiple vegetation types (Pasqualotto et al. 2019).

The majority of these methods have been developed for retrieving LAI from green vegetation only (Delegido et al. 2015), which is often called the Green Leaf Area Index (Green LAI or LAI-green). The Green LAI represents the leaves which are photosynthetically active (Daughtry et al. 1992). In contrast to the brown or senescing LAI (Delegido et al. 2015), remotely sensed Green LAI is more useful for agro-ecosystem monitoring (Pasqualotto et al. 2019), assessment of water logging damage in agriculture (Liu et al. 2018), estimating vegetation phenology (Verger et al. 2016), monitoring of deforestation (Valderrama-Landeros et al. 2016), crop modelling (El Hajj et al. 2016) and yield prediction (Lobell et al. 2015).

Previous studies showed that LAI can be estimated using spaceborne sensors such as AVHRR (Franch et al. 2017), MODIS (Huang et al. 2015), Landsat (Gao et al. 2012), WorldView-2 (Psomiadis et al. 2017) and Sentinel-2 (S2) (Verrelst et al. 2015; Djamai et al. 2019; Pasqualotto et al. 2019). Each of these sensors has their pros and cons, which mainly arise from their spatial, temporal and spectral resolutions or costs. Over the last decade, the number of companies developing CubeSats has increased. These new satellites, such as Planet Labs' PlanetScope (PS) CubeSat, can be the size of a milk carton, are relatively inexpensive to build and launch to a low Earth orbit, thereby making it possible to acquire large image collections at high spatial and temporal resolutions at a relatively low cost. However, one of the major challenges working with time series CubeSat imagery is the fact that unlike large and expensive satellites such as S2 or Landsat, the images obtained from CubeSat constellations, such as Planet's PS, frequently suffer from radiometric inconsistencies in the data collected by the different satellites within the constellation, due to inter-calibration challenges and their low signal-to-noise ratio (Houborg and McCabe 2016; Houborg and McCabe 2018b; Leach et al. 2019; Sadeh et al. 2019).

The lack of suitable combinations of both high spatial and temporal resolution time series from well calibrated satellite images (Waldner et al. 2019) motivated a few attempts to fuse CubeSat imagery with these other types of imagery into high spatio-temporal LAI datasets. For example, Houborg and McCabe (2018b) created Landsat-consistent LAI of an irrigated alfalfa field in Saudi Arabia by fusing PS, Landsat and MODIS images coupled with *in-situ* measurements, to spatially and temporally enhance Landsat-based LAI to the PlanetScope resolution. Li et al. (2019) generated red-edge bands at 3 m spatial resolution by fusing S2 and PS images, by using the weight-and-unmixing algorithm as well as the SUPer-REsolution for multi-spectral Multi-resolution Estimation (Wu-SupReME) approach. However, their fusion method was tested with only a few individual images acquired on selected dates, and their

relationship between *in-situ* wheat LAI measurements and the VIs from fused images only applies to Jiangsu Province, China, where it was established. Kimm et al. (2020) used the Moderate Resolution Imaging Spectroradiometer (MODIS)-Landsat STAIR (SaTellite dAta IntegRation) fusion product (Luo et al. 2018) and fused it with PS data to produce daily LAI estimation of corn and soybean in the U.S. Corn Belt. The STAIR method uses an adaptive-average correction that takes into account different land cover types through an automatic segmentation of the image (Luo et al. 2018).

Motivated by the inconsistency issues of the data acquired by the different satellites within the constellation, this study: 1) proposed a new method to fuse time series of images sourced from two different satellites to overcome the inconsistencies between the different sensors within the CubeSat constellation, and combines the advantages of both data sources in terms of their temporal, spatial and spectral resolutions. In contrast to some other fusion methods (e.g. Gao et al. 2006; Li et al. 2019), which can take only one or two pairs of images as input at a time, this new method can process a time series from an unlimited number of images; 2) applied this new technique to fuse PS images (with a spatial resolution of ~3 m, and a daily revisit time) and S2 images (resolution of 10 m and five-day revisit time) to create daily, S2-consistent surface reflectance blue, green, red (visible) and near-infrared (NIR) and crop Green LAI at a 3 m resolution; 3) tested the approach for improved wheat LAI estimation over wheat fields in Australia and Israel, so as to provide an assessment over different soil types, farm management, climates and crop varieties. The guideline for the development of the method was that the method should be simple, so it could be easily be replicated and applied elsewhere. Therefore, the reliable and well-studied Sentinel-2 LAI product was selected, which offer a global coverage of LAI estimates in a relatively high spatial resolution. Recently developed methods for LAI estimation tend to use sophisticated computing techniques such as machine learning, but they typically involve the use of ground-based training data specific to the study area. This paper contends that a practical and robust method for LAI estimation should be simple, effective, repeatable and universal. Therefore, the Sentinel-2 LAI product was selected as the reference, having global coverage of LAI estimates at relatively high spatial resolution. By converting the fused VIs into Sentinel-2-like LAI estimates (as described in section 2.4), the need of having ground LAI data is unnecessary. The resulting daily 3 m LAI estimations were compared with *in-situ* wheat LAI measurements made using ground-based methods. This new time series fusion method facilitates continuous daily high-resolution monitoring of crops over large scales.

3.3. Methodology

3.3.1. Field trials and in-situ LAI measurements

3.3.1.1. Cora Lynn trial

Winter-wheat variety RGT Accroc was grown in a 76 x 74 m field 80 kilometres South-East of Melbourne, Victoria, Australia, at Cora Lynn (38.1336° S, 145.6324° W, average annual rainfall of 857 mm (Australian Bureau of Meteorology 2020)). The crop was sown in the silty loam at a 5 cm depth on Aug 7, 2018, with 100 kg ha⁻¹ MAP (mono-ammonium phosphate) applied at sowing. The crop was grown under rainfed conditions, with only one irrigation (Nov 16, 2018) of 50 m³ ha⁻¹ with a linear shift irrigator, to avoid plant death. Four sets of above-ground plant parts were collected from the four sides of the field (at least 2 m from the edge) 16 times during the growing season, in a 0.5 x 0.5 m sampling area. For one or two of those four sets, leaf blades, stems and sheaths, and heads were portioned to measure their dry biomass and calculate the proportion of leaf material (i.e., dry weight of the blades divided by the total above-ground biomass). A subset of approximately 100 blades from the sampled leaves were then scanned using a Canon imageRUNNER ADVANCE C3330 scanner (Canon Inc) and weighed, after oven dry at 60°C for at least 48 hours, to measure the specific leaf area. LAI was calculated by multiplying specific leaf area, the proportion of leaf in the subsample and the average biomass of four samples, and by dividing by the sampling area.

3.3.1.2. Birchip fields

Five rainfed spring-wheat fields were studied in a farm located near Birchip, Victoria, Australia (Apr-Nov, 2018; 35.982° S, 142.916° W), representing an average Victorian farm in the Australian wheat belt. The Birchip farm, which, is 6,400 ha in size (average field size is 116 ha), was chosen for this study as it represents a typical Australian rainfed crop farm. This site is located on fine sandy clay loam texture soil and receives an average annual rainfall of 374 mm (Australian Bureau of Meteorology 2020). The LAI measurements were conducted on the 17th and 18th of September 2018 using a LAI-2000 Plant Canopy Analyzer (LI-COR). Ten 20 m X 20 m plots located in five wheat fields were selected (two in each field). All plots were located at least 20 m away from the edges of the field and were representative of the crops' conditions in their area. Field data were acquired following the measurement guidelines suggested by the instrument manual (LI-COR 1992). In total, 240 individual LAI measurements were sampled for the ten plots, with each plot containing 24 LAI measurements.

3.3.1.3. Saad and Yavne fields

LAI was measured over six rainfed spring-wheat fields, located in two commercial farms near Saad (four fields of ~39 ha in total; Feb-Apr, 2018; 31.477° N, 34.538° W) and Yavne (two

fields of ~13 ha in total; Jan-Apr, 2019; 31.809° S, 34.716° W) in Israel. The Saad fields are located over a clay soil and receive an average annual rainfall of 415 mm (Israel Meteorological Service 2020). The Yavne fields are located over a sandy loam soil and receive an average annual rainfall of 515 mm (Israel Meteorological Service 2020). LAI in these two sites was measured using the SunScan Canopy Analysis System (SS1-COM-R4 Complete System with Radio Link developed by Delta-T Company, Cambridge, United Kingdom). The four fields in Saad were close to each other with different sowing dates and irrigation regimes: Kitain cv sown on 20/11/2017 and grown under both rainfed (1) and with some irrigation (2); Amit cv was sown on 29/11/2017 and grown under rainfed conditions (3), and durum wheat cultivar C9 was sown on 19/11/2017 and grown under rainfed conditions. The two fields in Yavne were adjacent and sown on 16/11/2018. The measurements performed six times for the Saad farm fields and seven times for the Yavne farm fields during the growing season. Each LAI value used for the analyses was the average of LAI measured at 2 to 4 points, separated to each other by a distance of ~50 m. At each point, around 30 field measurements were taken every ~20 cm from each other, regardless of whether plants were present or not.

3.3.2. Imagery

3.3.2.1. Sentinel-2 (surface reflectance and LAI)

The European Space Agency (ESA) Copernicus Sentinel-2 (S2) includes a constellation of two polar-orbiting satellites positioned in the same sun-synchronous orbit, but phased at 180° to each other. S2 carries an optical sensor payload that samples 13 spectral bands: four bands at 10 m, six bands at 20 m and three bands at 60 m spatial resolution. It provides a revisit frequency of 5 days (at the Equator) with a 290 km swath width (Drusch et al. 2012; SUHET 2015). S2 images can be freely downloaded at the Copernicus Open Access Hub website (<https://scihub.copernicus.eu/dhus/#/home>). In this study, clear-sky images downloaded via ESA's application programming interface (API), using the field's polygon to determine the region of interest (ROI) to be downloaded. S2 Level-2A Bottom Of Atmosphere (BOA) products were only available (and used) for Israel during the study period. For Australia, we thus used the Sen2Cor module (Louis et al. 2016) within ESA's Sentinel Application Platform (SNAP) software (version 7.0) to convert the Level-1C product (Top Of Atmosphere (TOA) reflectance) images from TOA to BOA, in order to minimize the influence of the atmospheric conditions present at the time of acquisition. Next, S2-based LAI data (from Israel and Australia) were generated using the Biophysical Processor module embedded in SNAP, which computes biophysical products from S2 BOA reflectance. This processor uses the top-of-canopy reflectance data to estimate a number of biophysical variables including LAI (Weiss and Baret 2016).

3.3.2.2. PlanetScope

PlanetScope (PS) is a CubeSat 3U form factor (10 cm x 10 cm x 30 cm) satellite constellation operated by Planet Labs, Inc. The PlanetScope constellation consists of about 120 satellites, with the capability to image all of the Earth's land surface on a daily basis. The PlanetScope satellites have four spectral bands; Blue (455 – 515 nm), Green (500 – 590 nm), Red (590 – 670 nm) and NIR (780 – 860 nm). These have a Ground Sampling Distance (GSD) of 3-4 m at nadir and a positional accuracy of <10 m RMSE (Planet Team 2018). This study used the Planet Surface Reflectance Product provided at a spatial resolution of ~3 m. These images are atmospherically corrected to BOA reflectance, which provides more consistency across time and location localized atmospheric conditions while minimizing uncertainty in the spectral response (Planet Team 2020). Despite the fact that both PS and S2 provide imagery in the visible and NIR regions, their bandwidths and spectral response are very different as shown in Figure 10. For each analysed field, cloud-free PlanetScope images were downloaded using Planet's API, according to the field's domain.

3.3.3. Data fusion of reflectance

In order to fuse images acquired by the PS CubeSats constellation and S2, we have developed a simple fusion method (Figure 11). The data fusion process required four inputs: (1) High spatio-temporal resolution images (e.g. PS); (2) lower spatial resolution, but with higher spectral resolution images (e.g. S2); (3) an index or product produced by input number 2 (e.g. LAI); and 4) the ROI (i.e., a polygon of field's domain). The outputs of this fusion method are daily fused surface reflectance images and daily images of the desired index or product, in the original pixel size of the high spatial resolution input. In this study, we tested our fusion method to produce S2-like visible-NIR bands and LAI images at the spatial and temporal resolution of PS (i.e., daily images in 3 m).

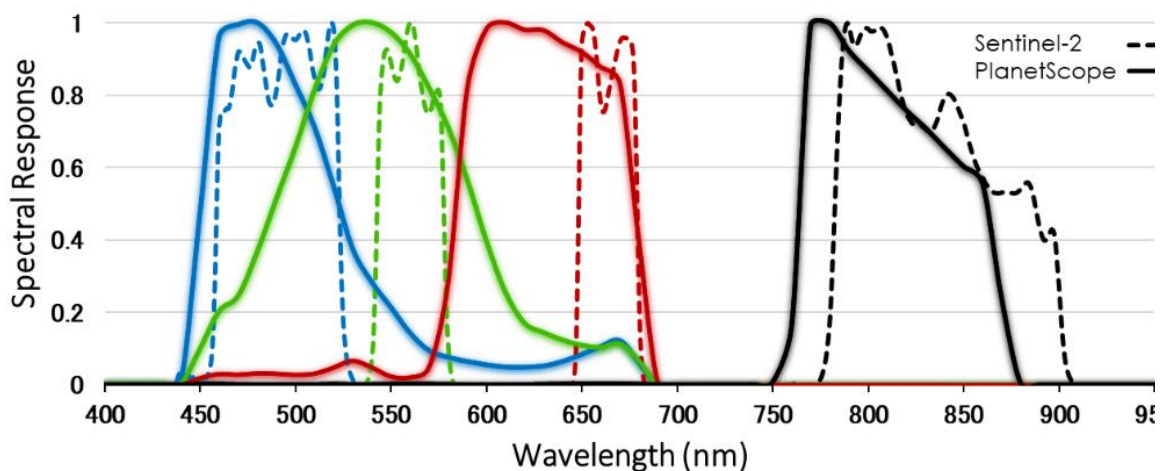


Figure 10. The spectral response of Sentinel-2 and PlanetScope in the Blue, Green, Red and Infrared bands.

First, PS, S2 (bands 2, 3, 4 & 8) and S2-based LAI images were extracted by the field's domain. Then, each type of consecutive pair of images acquired at two different dates are linearly interpolated to create a daily time-series of images. This results in three separate time-series, (1) PS BOA (3 m), (2) S2 BOA (10 m), and (3) S2 LAI (10 m) made from both real as well as synthetic (interpolated) images. Next, the S2-based datasets (2 + 3) were resampled (using cubic interpolation) from 10 m to 3 m pixel size. Then, same-day PS and the resized S2 images are separated into their individual bands (Blue, Green, Red and NIR), and fused by averaging the pixel values between each pair of bands as follow:

Equation 2

$$Fused\ Band_i = (PlanetScope_{BOA}\ Band_i + Sentinel2_{BOA}\ Band_i) / 2$$

Where, $Band_i$ correspond to each of RGB visible and the NIR bands (Figure 10). After the fusion, the bands are recombined to form an RGB-NIR image. This stage yields in daily 3 m fused surface reflectance images.

3.3.4. Fused LAI in 3 m

The new dataset is then used to calculate 13 selected vegetation indices shown in the literature to have a high correlation with LAI: SR, EVI2, NDVI, GCVI, MTVI2, MSAVI, WDRVI, Green-WDRVI, OSAVI, GSR, GNDVI, RDVI and TVI. Finally, the vegetation indices from the fused image (daily, 3 m) are converted to LAI using a linear regression between the different vegetation indices to the resized LAI time-series (daily, 3 m) from S2. A four-day moving window ($t_0, t_{-1}, t_{-2}, \& t_{-3}$) with same day (t_0) pairs of S2-LAI and a fused vegetation index image, was used to calculate the average slopes and intercepts between all four pairs (i.e., S2-LAI and one of the fused-based vegetation indices at the time). In this process, the slope and intercept of LAI and VI for all pixels from the field was calculated for each day and each studied vegetation index, and then averaged across the four days of the moving-window. Hence for each vegetation index, the correction of each image is done using a different slope and intercept, calculated for each four-day window, thus operating as a moving average. This was done in order to minimise the signal inconsistency created by the PS sensors (daily measurements), which causes variability in the vegetation indices. The algorithm was designed to use the last four days so it can work in operational near real-time mode, as the future images are still not available. However, when processing an existing time series of images, the algorithm can be easily be modified to include the next images in the time series (e.g., $t_{-2}, t_{-1}, t_0, t_1, t_2$).

The following equation is then used to correct and generate the daily high spatio-temporal LAI dataset:

Equation 3

$$\text{Corrected image} = \text{Fused vegetation index} * \text{four-day window slopes average} + \text{four-day window intercepts average}$$

This method (Figure 11) enables daily 3 m LAI images to be generated at the same quality as the S2 LAI product.

The remotely sensed LAI estimates were validated against the *in-situ* measurements. As the LAI was measured using different approaches over different study areas, the accuracy of the remotely sensed Green LAI was evaluated using a 23X23 square metre plot and a 65X65 square metre plot for the Birchip and Cora Lynn sites respectively. The fields in Israel were compared at the field level as each LAI measurement point was about 50 m apart from each other, and around 30 field measurements were taken at each of these points. The fields near Saad (four fields of ~39 ha in total) and Yavne (two fields of ~13 ha in total) are much smaller than the Australian commercial fields analysed in this study (the average field size is 116 ha) and the development of the crops in these fields has been far more homogeneous.

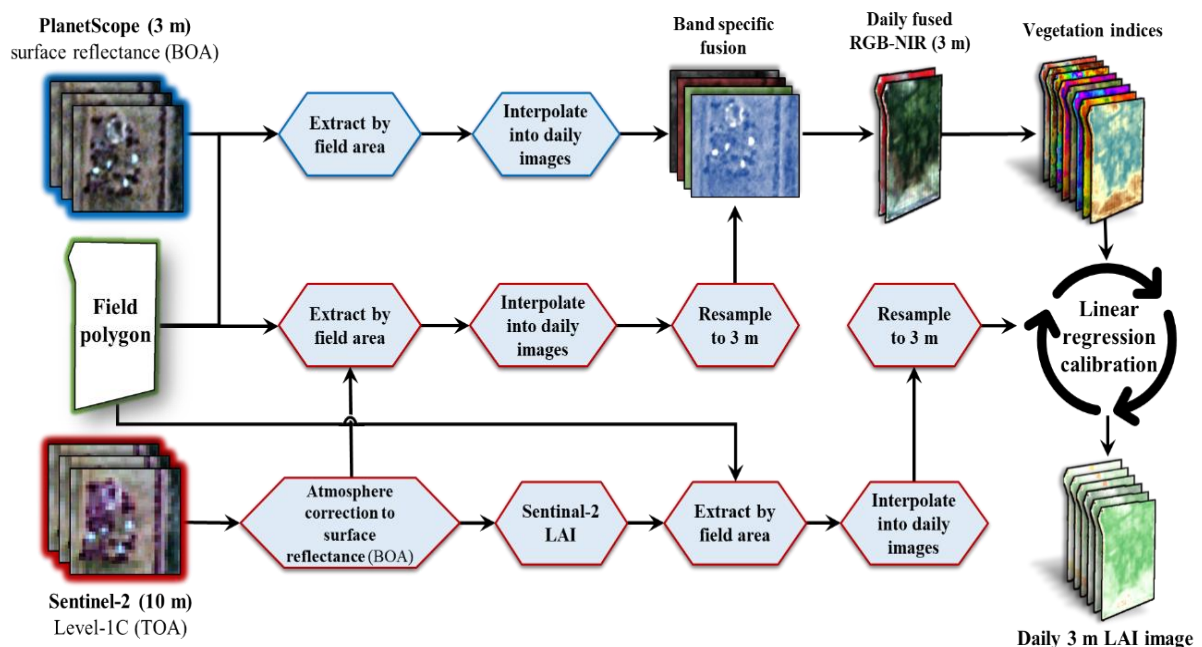


Figure 11. The data fusion workflow of PlanetScope (with a spatial resolution of ~3 m) and Sentinel-2 (10 m) imagery into daily surface reflectance images, vegetation indices and LAI maps with a 3 m resolution.

Table 4. Definition of the multispectral vegetation indices investigated in this study.

Vegetation index	Equation	Reference
Simple Ratio (SR)	$\frac{NIR}{Red}$	Jordan (1969)
Enhanced Vegetation Index 2 (EVI2)	$\frac{2.5(NIR - Red)}{(NIR + 2.4Red + 1)}$	Jiang et al. (2008); Nguy-Robertson et al. (2012)
Green Chlorophyll Vegetation Index (GCVI)	$(NIR - Green) - 1$	Gitelson et al. (2003); Gitelson et al. (2005)
Normalized Difference Vegetation Index (NDVI)	$\frac{NIR - Red}{NIR + Red}$	Rouse et al. (1974)
Modified Triangular Vegetation Index 2 (MTVI2)	$\frac{1.5 [1.2(NIR - Green) - 2.5(Red - Green)]}{\sqrt{(2NIR + 1)^2 - (6NIR - 5\sqrt{Red})} - 0.5}$	Haboudane et al. (2004)
Modified Soil-Adjusted Vegetation Index (MSAVI)	$0.5 [2NIR + 1 - \sqrt{(2NIR + 1)^2 - 8(NIR - Red)}]$	Qi et al. (1994); Haboudane et al. (2004)
Wide Dynamic Range Vegetation Index (WDRVI)	$\frac{\alpha \cdot NIR - Red}{\alpha \cdot NIR + Red} + \frac{1 - \alpha}{1 + \alpha}$	Gitelson (2004); Nguy-Robertson et al. (2014)
Green Wide Dynamic Range Vegetation Index (Green-WDRVI)	$\frac{\alpha \cdot NIR - Green}{\alpha \cdot NIR + Green} + \frac{1 - \alpha}{1 + \alpha}$	Peng and Gitelson (2011); Nguy-Robertson et al. (2014)
Optimized Soil-Adjusted Vegetation Index (OSAVI)	$\frac{NIR - Red}{NIR + Red + 0.16}$	Rondeaux et al. (1996)
Green Simple Ratio (GSR)	$\frac{NIR}{Green}$	Sripada et al. (2006)
Green NDVI (GNDVI)	$\frac{NIR - Green}{NIR + Green}$	Gitelson and Merzlyak (1994)
Renormalized Difference Vegetation Index (RDVI)	$\frac{NIR - Red}{\sqrt{NIR + Red}}$	Roujean and Breon (1995)
Transformed Vegetative Index (TVI)	$\sqrt{\frac{NIR - Red}{NIR + Red} + 0.5}$	Rouse et al. (1974) Haas et al. (1975)

* α in WDRVI and Green-WDRVI = 0.1 following Nguy-Robertson et al. (2014)

3.3.5. Adjustment of remotely-sensed Green LAI

The Green LAI (representing a canopy mostly photosynthetically active) are difficult to measure from space when leaves shade each other. To account for this, the Green LAI estimations were tested for improvement by adjusting the generic S2-LAI product estimations, to better estimates wheat Green LAI. This was done by 'fine-tuning' the results received in the

previous stage, using second-order polynomial regressions (Table 7), which was found to best represent the correlation between the *in-situ* and remotely estimated Green LAI for a considered vegetation index. The performance of this correction approach was tested using an independent dataset of Green LAI for two ~30 ha, rainfed-wheat fields located near Yanco, NSW, Australia (Apr-Nov, 2019; 34.716° S, 146.088° W). In this site, located more than 350 km from the nearest field used for training (i.e., further than the distance between Paris and London), the LAI was measured weekly using LI-COR LAI-2200 sensor during one month around the peak LAI (Oct, 2019). The LAI measurements were taken at least 10 m away from the edges of the field and were representative of the crops' conditions in their area. Owing to the high spatial variability of the vegetation development in these two fields, the validation between the *in-situ* and the remotely sensed LAI performed on the crops located around the actual location of the LAI measurements (using a 0.3 and 0.4 ha plots). The results of the corrected Green LAI were compared with the *in-situ* Green LAI measurements and the non-corrected estimations, in order to validate the proposed correction method to better estimate wheat Green LAI from space.

3.4. Results

3.4.1. Fused surface reflectance accuracy

The implementation of the new fusion method to generate time series of images resulted in a new dataset, which maintained both the high spatial and temporal resolution of PS and the spectral quality of S2 (Figure 12 and Figure 13). Figure 12 illustrates how a 10 m image from S2 fails to provide information about objects smaller than 10 m such as buildings, trees, and roads. By contrast, the fused image enabled easy identification of objects that could not be

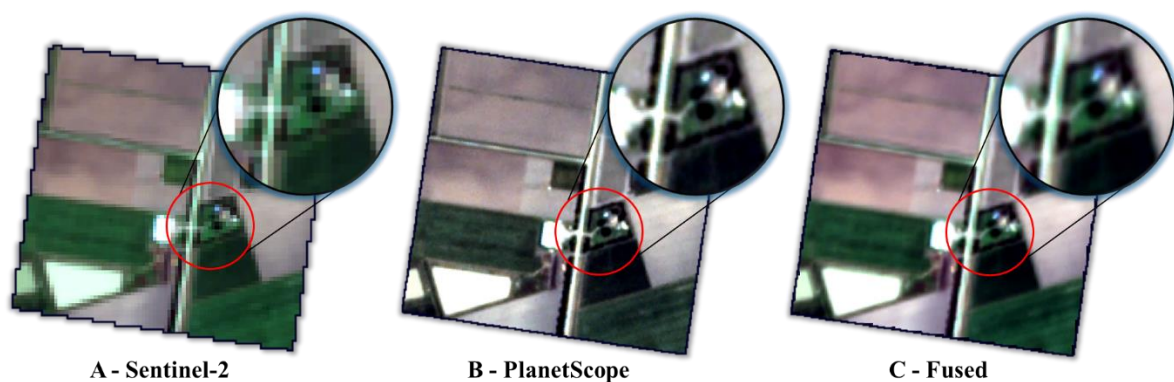


Figure 12. An example of the Sentinel-2 and PlanetScope fusion outcome (in natural color composite image) for the Cora Lynn experimental field area. Both source images were acquired on the 29/9/18 and are in BOA reflectance values. (A) The original Sentinel-2 image (10 m), (B) the original PlanetScope image (3 m), and (C) the fused image (3m).

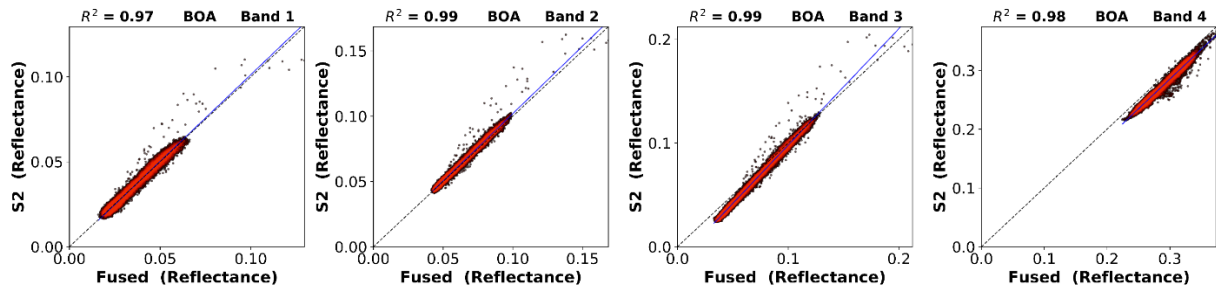


Figure 13. An example of the correlation between a same day pair of an original S2 image and a fused image (image date 23/7/18) for the field near Birchip, Victoria, Australia. Each scatterplot represents the comparison of a different spectral band, where band 1, 2, 3 and 4 represents the Blue, Green, Red and NIR wavelengths, respectively and the pixels values are in surface reflectance (BOA). The blue line in the figures is the trend line. The dotted line is the 1:1 line.

recognized in the S2 image while preserving the S2 reflectance information as shown in Figure 13.

The correlation between the BOA surface reflectance of S2, PS and the fused images was compared and evaluated at the pixel level, while excluding 15 m from the fields' edges to avoid having mixed pixels with the surrounding objects. The mean and median R^2 across all bands for the studied PS and S2 images were only 0.6 and 0.7, respectively, (Table 5). The highest correlations were found in the NIR wavelength (0.75 and 0.81 mean and median, respectively), while the lowest correlation was found in the blue wavelength (0.46 and 0.53 mean and median, respectively). The fused images were found to be highly correlated with the S2 images in all four bands (0.88 and 0.94 mean and median, respectively), being slightly higher than the correlation found between the fused images and PS images (0.84 and 0.9 mean and median, respectively). Sometimes when small objects such as scattered trees were located within the field, the analysis showed scattered pixels with lower correlation (Figure 13). This

Table 5. Comparison of the median and mean correlation (R^2) that was found between S2, PS and the fused images of all the images analysed across all sites (2,463 images in each dataset), for each band and all four bands together.

Datasets	Blue (band 1)		Green (band 2)		Red (band 3)		NIR (band 4)		All bands	
	median	mean	median	mean	median	mean	median	mean	median	mean
PS - S2	0.53	0.46	0.71	0.58	0.75	0.6	0.81	0.75	0.7	0.6
Fused - PS	0.86	0.79	0.9	0.82	0.91	0.84	0.94	0.91	0.9	0.84
Fused - S2	0.9	0.81	0.95	0.88	0.96	0.88	0.96	0.93	0.94	0.88

is often because the lower resolution of S2 tends to represent these objects as mixed pixels, while they can be clearly identified in the fused image.

3.4.2. Fused vegetation indices

The ability of the fused images to produce daily vegetation indices (VIs) time-series in values similar to S2-based time-series was tested and compared to both VIs derived from S2 and PS. Overall, it was found that PS-based VIs tended to have lower values than S2-based VIs except in the early stages of the growing season (e.g., on low VIs values) where PS-based VIs were slightly higher (e.g., Figure 14). Furthermore, a time-series of VIs generated based on PS images were noisier than the one generated based on S2 images or fused images (e.g., Figure 14).

After fusing the S2 and PS images and calculating the VIs, daily 3 m LAI maps were generated (Figure 15). These maps were compared to S2-based LAI maps at the pixel level and at the

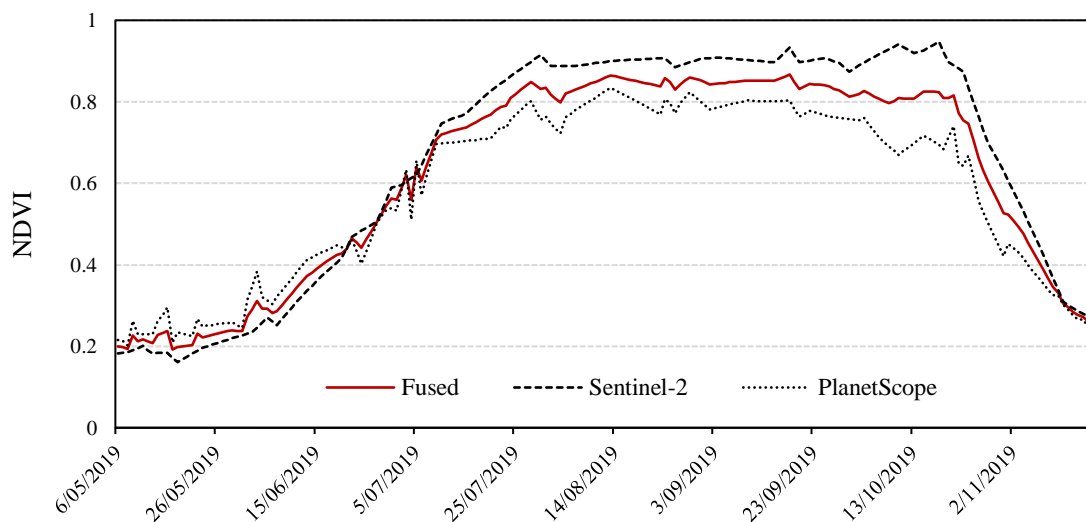


Figure 14. Six-month time series of Sentinel-2 (black dashed line), PlanetScope (dot line) and the fused images (red line) NDVI calculated from atmospherically corrected bottom of atmosphere (BOA) reflectances. The data represent the daily mean NDVI for a 131-ha wheat field located near Birchip, Victoria, Australia.

field level (Figure 16). Next, remotely-sensed LAI estimations from both S2 and fused images were compared with *in-situ* LAI measurements conducted in the field (Figure 16).

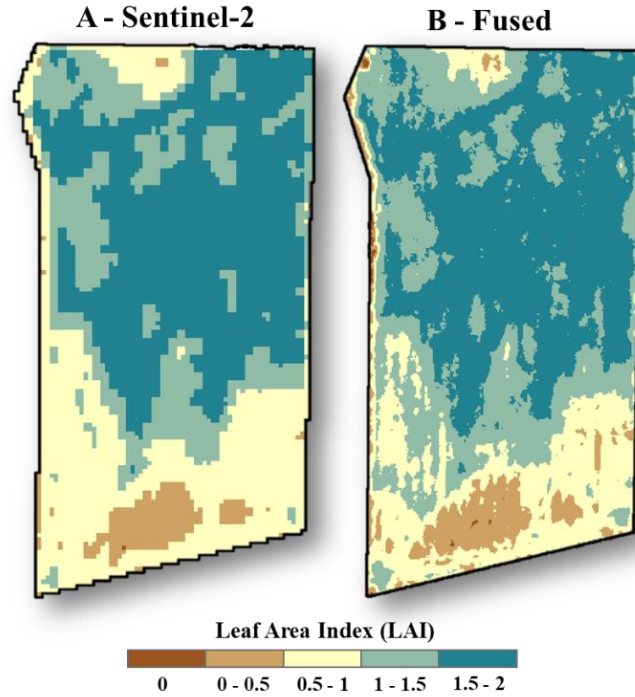


Figure 15. Comparison between (A) S2-based LAI map (10 m) and (B) S2-PS fusion-based LAI (3 m) map (image date – 7/8/18) of an 88-ha wheat field near Birchip, Victoria, Australia.

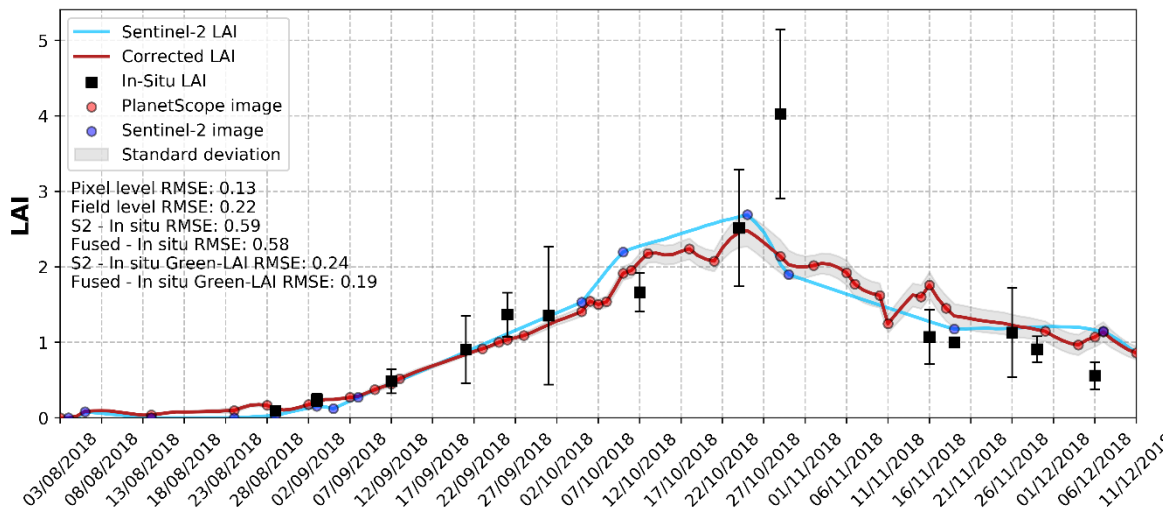


Figure 16. Comparison of the changes in LAI estimated over the growing season through in-situ point measurements (black dots, where the error bars represent the measurements standard deviation), S2 images (blue line) and the fused method (red line) of the 2018 Cora Lynn wheat trial. In this example, the fused-based Renormalized Difference Vegetation Index (RDVI) was used to calculate LAI using a linear regression model. RMSE is presented (i) between the S2-based LAI to the fused-based LAI estimations at the pixel ('Pixel level RMSE') and the field levels (median daily error at the pixel level; 'Field level RMSE'), (ii) between the S2-LAI and in-situ LAI measurements ('S2 - In-situ RMSE'), (iii) between the Fused images-LAI and in-situ LAI measurements ('Fused - In-situ RMSE'), (iv) between the S2-LAI and the Green LAI in-situ measurements only ('S2 - In-situ Green LAI RMSE'), and between the Fused images-LAI and the Green LAI in-situ measurements only ('Fused - In-situ Green LAI RMSE'). In this example, in-situ LAI measurement performed on the 29/10/18 were unusually high, and in any case, much higher than the estimated remotely-sensed LAI. This could be partly due to measurement error (the error bar was also big) combined with the known underestimation of S2-LAI for high LAI values. The grey color is the daily field-scale LAI standard deviation of the fused LAI dataset.

3.4.3. LAI estimations

Overall, 57 *in-situ* LAI measurements conducted across 12 wheat fields, half located in Australia and the other half in Israel, were available to evaluate the accuracy of the remotely-sensed LAI estimations. S2-based LAI was found to have an RMSE of 1.60 ($R^2 = 0.84$), while the RMSE for fused-based LAI estimations ranged from 1.73-1.78 depending on the VI considered ($R^2 = 0.82-0.84$;). In general, both S2 and fused-based LAI estimations tended to underestimate the *in-situ* LAI values, especially in LAI values larger than 3 (Figure 17). In comparison with the *in-situ* measurements, S2-based LAI estimations generally had slightly better accuracy compared to VI-based LAI estimations. As optical remote sensing is mainly estimating the Green LAI (Haboudane et al. 2004), the results have been separated into two groups, the *in-situ* measurements conducted when the crops were at the Green LAI stage (-B) and the measurements that were performed during the senescing stage (-C). The peak of the fusion-based LAI was used as the threshold to define these two groups. When analysing the Green LAI separately, the accuracy of the remotely sensed LAI estimations was found to be much higher (). Overall, remotely-sensed LAI estimations during the Green LAI phase were found to have an RMSE of 1.08 ($R^2 = 0.95$) for the S2-based LAI, and an RMSE of 1.37-1.4 ($R^2 = 0.92-0.94$) for the fused-based LAI estimated from the different VIs (-B). S2-based LAI median error for the Green LAI was only -0.38 and the fused-based LAI from best performing VI, i.e., RDVI, had a median error of -0.73 (-B). Hence, estimating Green LAI using the new method, or with S2-based LAI estimations, was highly correlated to *in-situ* measurement when the crops were still mostly photosynthetically active. However, underestimations of high LAI values (> 3) was observed in all studied fields (Figure 17), probably due to increasing overlap of leaves with higher LAI. One of the main disadvantages of using normalized difference VIs (e.g., NDVI) to remotely estimate LAI is the fact that they tend to saturate asymptotically under conditions of medium-to-high aboveground biomass density (Gitelson 2004). To minimise these underestimations, a correction equation was sought for the different VI-based fused dataset and the S2-based LAI. Polynomial order two regression was found as the most suitable to fit the fused-based Green LAI estimates against *in-situ* LAI measurements (Figure 17 and Table 7), with an R^2 of 0.95 (RMSE of 0.62) for the S2-based LAI , and an R^2 ranging between 0.92 and 0.94 depending on the VI considered (RMSE of 0.67 - 0.78). Overall, the SR, MTVI2 and RDVI indices showed the best fits.

Table 6. Performance of remotely-sensed LAI estimations for both S2-based and the fused-based LAI compared to the *in-situ* LAI measurements. The best performing index from the VIs-based fusion is coloured in red. Indices of the S2-based LAI estimations are coloured in green when they performed better than the fused data. Table A presents the results for all of the LAI measurements conducted in the field ($n = 57$), table B refers only to *in-situ* measurements conducted when the crops were at the Green LAI stage ($n = 25$) and table C presents the results only for LAI measurements conducted in the field during the senescing stage of the crops ($n = 32$). Overall, the best accuracy was achieved during the Green LAI stage (table B).

	(A) All samples ($n = 57$)													
	S2 LAI	EVI2	GCVI	GNDVI	Green-WDRVI	GSR	MSAVI	MTVI2	NDVI	OSAVI	RDVI	TVI	WDRVI	SR
RMSE	1.60	1.77	1.77	1.78	1.77	1.77	1.77	1.74	1.77	1.77	1.77	1.77	1.76	1.73
R^2	0.84	0.83	0.84	0.83	0.83	0.84	0.82	0.83	0.82	0.82	0.83	0.83	0.83	0.84
Mean error	-1.08	-1.24	-1.26	-1.26	-1.26	-1.26	-1.24	-1.22	-1.24	-1.24	-1.24	-1.25	-1.24	-1.23
Median error	-0.55	-0.85	-0.93	-0.91	-0.93	-0.93	-0.83	-0.86	-0.84	-0.84	-0.91	-0.84	-0.88	-0.88
Mean accuracy %	79.52	76.13	75.02	75.26	75.11	75.03	76.25	76.88	76.18	76.18	77.11	76.16	75.98	75.96
Median accuracy %	76.62	74.41	73.24	73.64	73.38	73.27	74.46	75.37	74.39	74.39	74.20	75.45	73.84	72.41

	(B) Green LAI only ($n = 25$)													
	S2 LAI	EVI2	GCVI	GNDVI	Green-WDRVI	GSR	MSAVI	MTVI2	NDVI	OSAVI	RDVI	TVI	WDRVI	SR
RMSE	1.08	1.39	1.39	1.38	1.39	1.39	1.38	1.37	1.39	1.39	1.38	1.38	1.40	1.37
R^2	0.95	0.93	0.93	0.92	0.92	0.93	0.93	0.94	0.93	0.93	0.94	0.93	0.93	0.94
Mean error	-0.68	-0.95	-0.97	-0.95	-0.96	-0.97	-0.94	-0.93	-0.94	-0.94	-0.96	-0.94	-0.96	-0.95
Median error	-0.38	-0.83	-0.87	-0.80	-0.84	-0.87	-0.81	-0.74	-0.82	-0.82	-0.73	-0.82	-0.85	-0.87
Mean accuracy %	90.85	84.84	83.74	84.12	83.86	83.76	85.08	85.58	84.92	84.92	86.49	85.81	84.60	84.65
Median accuracy %	86.95	79.73	76.34	78.33	76.69	76.34	80.57	80.44	80.15	80.15	81.37	80.47	79.63	79.33

	(C) LAI of senescing canopy ($n = 32$)													
	S2 LAI	EVI2	GCVI	GNDVI	Green-WDRVI	GSR	MSAVI	MTVI2	NDVI	OSAVI	RDVI	TVI	WDRVI	SR
RMSE	1.90	2.02	2.01	2.03	2.02	2.01	2.02	1.99	2.02	2.02	2.02	2.02	2.00	1.97
R^2	0.83	0.79	0.81	0.80	0.81	0.81	0.79	0.79	0.79	0.79	0.79	0.80	0.80	0.80
Mean error	-1.41	-1.51	-1.51	-1.52	-1.52	-1.51	-1.51	-1.49	-1.51	-1.51	-1.51	-1.52	-1.50	-1.48
Median error	-1.08	-1.16	-1.23	-1.25	-1.24	-1.23	-1.16	-1.11	-1.16	-1.16	-1.14	-1.16	-1.16	-1.13
Mean accuracy %	70.66	69.32	68.21	68.33	68.27	68.22	69.35	70.09	69.35	69.35	69.78	68.63	69.24	69.18
Median accuracy %	62.58	62.73	58.58	59.80	59.07	58.58	60.83	61.06	62.01	62.01	59.78	61.11	61.74	59.95

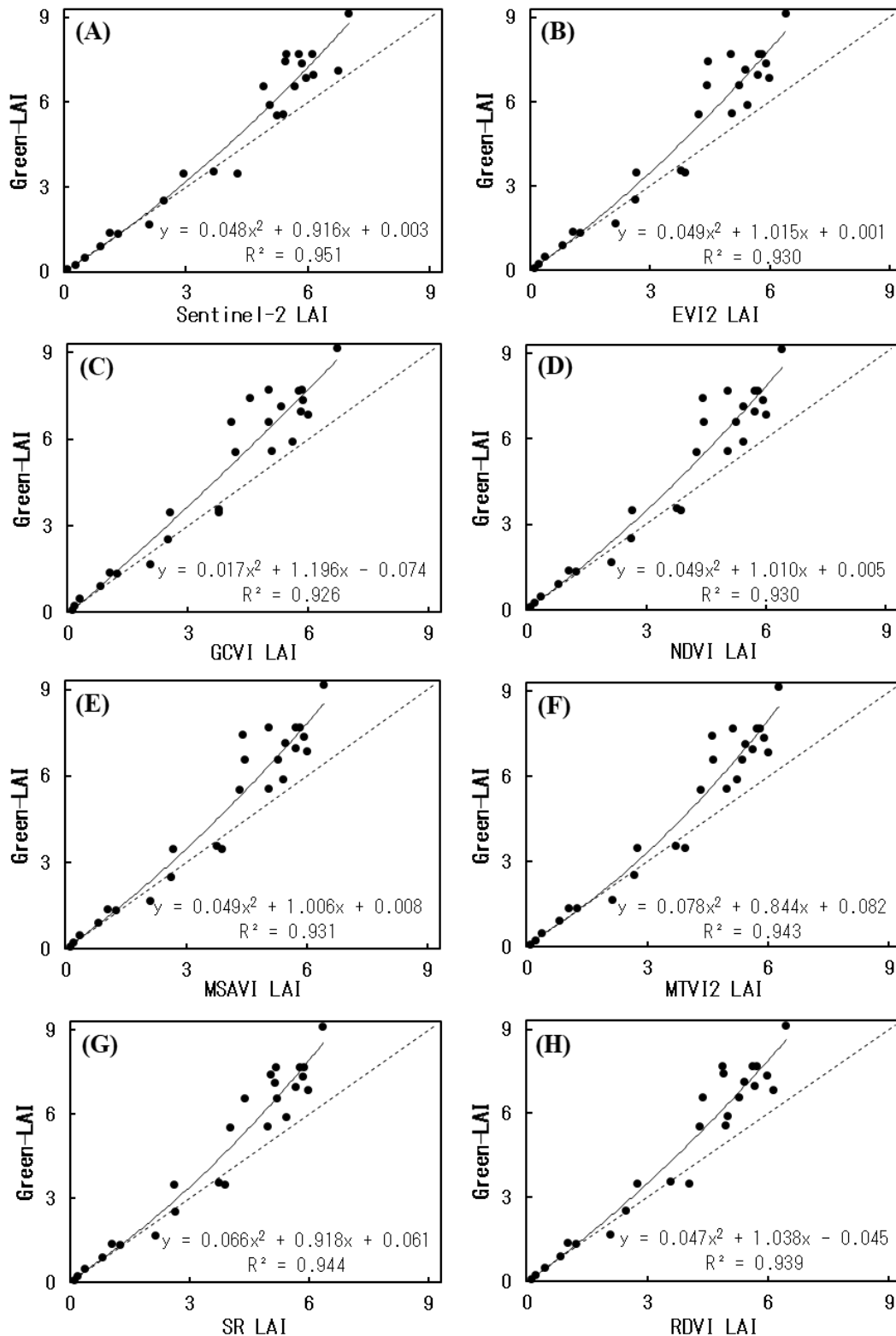


Figure 17. Example of in-situ vs. remotely sensed Green LAI: (A) Sentinel-2 based LAI (B) Enhanced Vegetation Index 2 (EVI2), (C) Green Chlorophyll Vegetation Index (GCVI), (D) Normalized Difference Vegetation Index (NDVI), (E) Modified Soil-Adjusted Vegetation Index (MSAVI), (F) Modified Triangular Vegetation Index 2 (MTVI2), (G) Simple Ratio (SR), (H) Renormalized Difference Vegetation Index (RDVI). In all plots, solid line is best-fit function, dashed line is the one-to-one line. Statistics relative to these fits are presented in Table 4.

Table 7. Best-fit functions of the relationships between Green LAI and VI's obtained using a cross-validation procedure for wheat, when $x = VI$, $y = \text{Green LAI}$, and the RMSE is the root mean squared error of the Green LAI estimation. Fits were performed for data of all the studied trials and are presented in Figure 17.

VI	Equation	R²	RMSE
Sentinel-2 LAI	$y = 0.0482x^2 + 0.9161x + 0.0026$	0.95	0.62
SR	$y = 0.0658x^2 + 0.9179x + 0.0614$	0.94	0.67
MTVI2	$y = 0.0784x^2 + 0.8443x + 0.0823$	0.94	0.68
RDVI	$y = 0.0475x^2 + 1.0382x - 0.0452$	0.94	0.70
WDRVI	$y = 0.0502x^2 + 1.0139x + 0.0005$	0.93	0.74
MSAVI	$y = 0.0493x^2 + 1.006x + 0.0085$	0.93	0.74
TVI	$y = 0.0464x^2 + 1.0279x - 0.0131$	0.93	0.75
OSAVI	$y = 0.0492x^2 + 1.0102x + 0.0052$	0.93	0.75
NDVI	$y = 0.0492x^2 + 1.0102x + 0.0051$	0.93	0.75
EVIZ	$y = 0.0489x^2 + 1.015x + 0.0012$	0.93	0.75
GSR	$y = 0.0171x^2 + 1.196x - 0.0743$	0.93	0.77
GCVI	$y = 0.0171x^2 + 1.1961x - 0.074$	0.93	0.77
Green WDRVI	$y = 0.0164x^2 + 1.1967x - 0.0726$	0.92	0.78
GNDVI	$y = 0.0183x^2 + 1.1786x - 0.0599$	0.92	0.78

3.4.4. Adjustment of S2 LAI to estimate Green LAI in wheat

The equations presented in Table 7 were used to correct, for the different fusion-based LAI and S2-based LAI, the underestimation of the remotely sensed Green LAI estimations (as illustrated in Error! Reference source not found.). The performance of the proposed correction method was tested using an independent LAI dataset that was collected in two wheat fields located near Yanco, NSW, Australia. The performance of this method was evaluated for estimating Green LAI with and without the proposed correction.

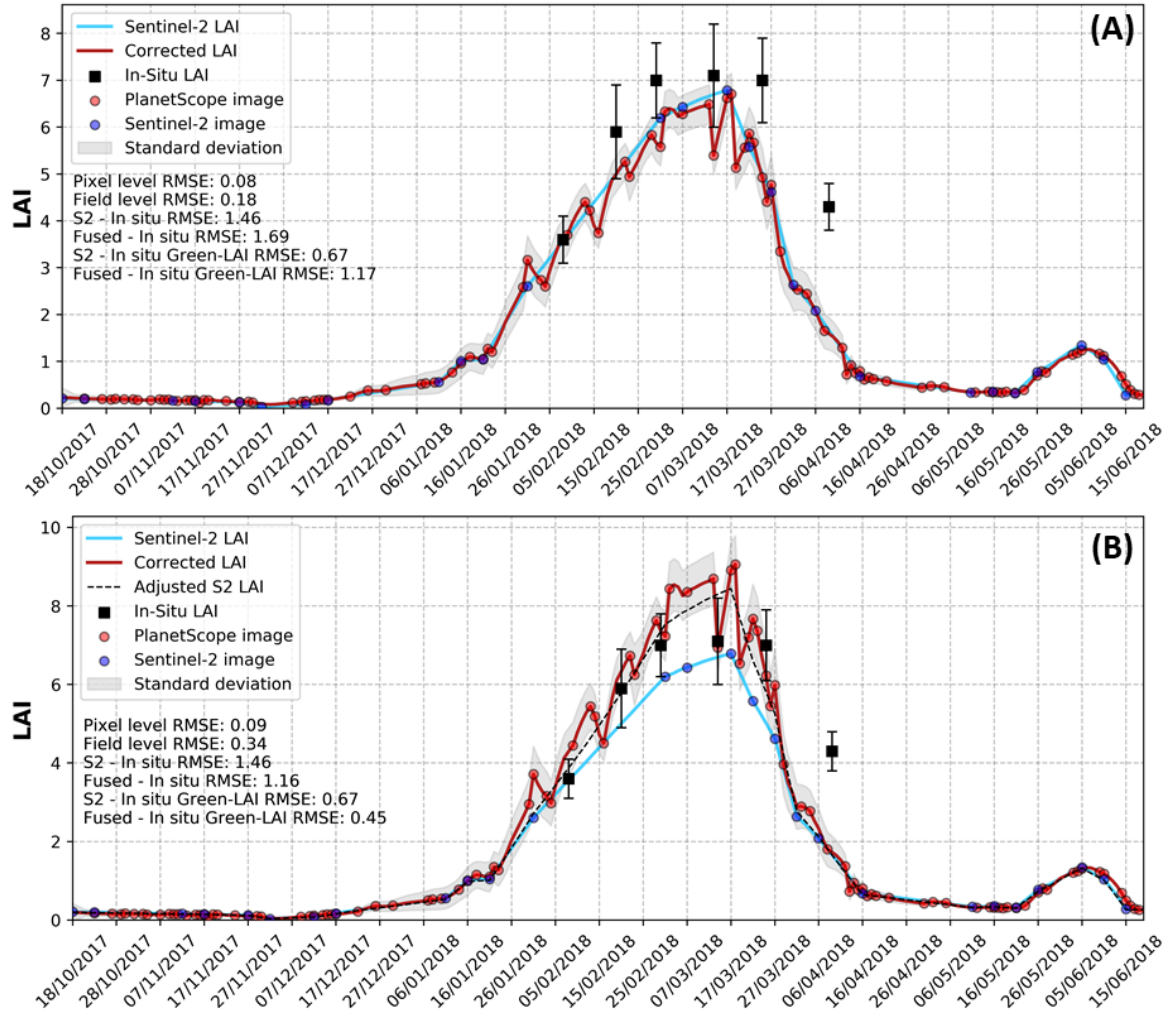


Figure 18. Change over time in LAI for pre- (A) and post- (B) corrected LAI values in a wheat field in Saad in 2018. See caption of Figure 16 for details of the legend. (The field which presented in this figure was not used as part of the independent LAI dataset, it is used here as it illustrates best the adjustment concept).

The results of the Green LAI correction analysis showed that the estimations of the fused-based LAI were improved by up to 47% compared with non-corrected Green LAI estimations (Table 8). The RMSE between the *in-situ* LAI and the fused-based LAI before the correction ranged between 0.53 and 0.87 (among the different indices), while the correction achieved higher accuracy with RMSE ranging between 0.35 and 0.63, as shown in Table 8. Even though the proposed correction aimed to adjust the Green LAI phase of the wheat, the results show that this method also adjusts LAI estimations at the senescing phase. In the conditions tested, the best pre-correction performing indices were for MSAVI, MTVI2 and GNDVI (all < 0.6 RMSE) and the best post-correction performing indices were for GNDVI, Green WDRVI, GDVI, GSR, TVI, OSAVI and NDVI (all < 0.45 RMSE). Overall, the RMSE of Green LAI estimations improved by more than 25% for 10 out of the 13 indices analysed. As for the S2-LAI estimations, in these fields, the S2-LAI underperformed the fused-based LAI estimations

Table 8. Performance of remotely-sensed Green LAI estimates compared to *in-situ* LAI measurement for pre and post Green LAI correction. This table presents the RMSE between the *in-situ* Green LAI measurements of an independent dataset (fields of Yanco) and the Green LAI estimated from the fused data and S2 images. Best performing index are coloured in red.

Green LAI	S2 LAI	EVI2	GCVI	GNDVI	Green-WDRVI	GSR	MSAVI	MTVI2	NDVI	OSAVI	RDVI	TVI	WDRVI	SR
RMSE (pre-correction)	1.38	0.75	0.76	0.58	0.76	0.76	0.53	0.56	0.65	0.65	0.68	0.61	0.87	0.84
RMSE (post-correction)	1.38	0.50	0.43	0.35	0.40	0.43	0.45	0.57	0.44	0.44	0.54	0.44	0.62	0.63
Improvement in RMSE	0%	34%	43%	39%	47%	43%	14%	-1%	31%	32%	21%	28%	29%	25%
R ² (pre-correction)	0.412	0.767	0.921	0.584	0.761	0.921	0.272	0.024	0.439	0.481	0.519	0.365	0.906	0.748
R ² (post-correction)	0.411	0.764	0.920	0.584	0.760	0.920	0.270	0.024	0.433	0.475	0.527	0.359	0.898	0.746

having an RMSE of 1.38 ($R^2 = 0.41$). This stems from the fact that the *in-situ* LAI was measured along the edges of these fields. Therefore, the 10 m spatial resolution of S2 suffered from mixed pixels, representing not only the wheat LAI but also the road around the fields, while the 3 m fused LAI images overcame this limitation.

3.5. Discussion

Monitoring crop performance is essential to guarantee a high quality and profitable yield; however, it has always been a challenge, especially with the large fields that are common to modern agriculture. The large cultivated areas and the frequent monitoring requirement (Waldner et al. 2019), makes remote sensing a valuable tool for farmers and agronomists to achieve maximum yield (Raun et al. 2002).

3.5.1. The advantages of the proposed fusion approach

CubeSats, such as Planet's PS, are relatively cheap to build and can offer high spatio-temporal imagery at lower costs than traditional satellites. However, CubeSat constellations tend to suffer from cross-sensor inconsistencies in radiometric quality and dissimilarity of their spectral responses among satellites in the constellation, contributing to the noise observed in time-series data acquired from these sensors (Houborg and McCabe 2016, 2018a). Such inconsistencies limit the accuracy of surface reflectance-based applications such as estimation of vegetation indices (Figure 14), hindering the use of CubeSat satellites to monitor changes on the Earth surface (Sadeh et al. 2019) and for land surface characterization (Houborg and McCabe 2018a). Fusion of CubeSat imagery with a consistent and reliable dataset, such as S2, can overcome this limitation and eliminate the noise that exists in

CubeSat data (Houborg and McCabe 2018a; Kimm et al. 2020). PS bands have different bandwidths and spectral responses from S2 (Figure 10). While S2 RGB-NIR bands have a clear spectral separation between the bands, PS's RGB bands overlap each other. Such overlaps can cause contamination of the signals acquired by each band, with radiation belonging to neighbouring bands, limiting the accuracy of different applications such as VI calculations and classification.

The fusion approach proposed in this study helps to resolve the challenges posed by PS spectral responses, producing time-series of images that preserve both the high spatial and temporal resolution of PS and the spectral quality of S2 (as shown in Figure 12 and Figure 13). In practical terms this means that S2-consistent, surface reflectance RGB-NIR images and crop Green LAI could be generated at a 3 m resolution on a daily basis.

The fusion method proposed here can process a time series from an unlimited number of images to generate a time series of images that covers the whole growing season. This makes the method robust and flexible, and the user can theoretically fuse images acquired by a number of sensors. The fleet of Earth observing satellites is increasing every year, offering unprecedented imagery in a range of spectral resolutions acquired across various bands. However, some previous image fusion methods are limited to a maximum of three input bands of a lower resolution at a time (Gašparović and Jogun 2018). This method is not limited by the number of bands to be fused, so long as the higher spatial-resolution bands covers the spectral range of the lower resolution bands. In order to reduce sampling gaps between images sourced from one dataset such as S2, the method can also use Landsat images for example, to increase the temporal resolution of these coarse images. In the same way, if a new CubeSat constellation that can provide data complimentary to PS becomes operational, it can be integrated as an additional high spatio-temporal imagery in the fusion process.

The potential of spaceborne remote sensing to provide relevant information by monitoring crop performance has long been recognized. However, despite the technological and methodological progress over the past decades, remotely sensed data are still not as broadly and operationally used by farmers as they should be. This could be because of the cost of images with both high temporal (<5 days) and spatial resolutions (<5 m) (e.g. DigitalGlobe's WorldView-2 and 3), which limit the profitability of the farm. Or it could be that agronomists and farmers do not have the knowledge and skills required to process and analyse the satellite data. This study has addressed the issue of availability for affordable high spatio-temporal data for crop monitoring at a field and sub-field scale. In addition, such new datasets can be used for precision agriculture applications, which until now couldn't be implemented owing to the temporal or spatial limitation of the existing publicly available sources of satellite data.

Near real-time estimation of Green LAI can provide farmers with the tool to monitor the crop health and growth status, which may support farm management actions such as irrigation and fertilization (Pasqualotto et al. 2019). However, waiting for a cloud free image from the publicly available satellite imagery (e.g. S2 and Landsat), often results in an image that is too late to act in the field or will result in an incorrect interpretation (Khan et al. 2018). For example, the peak of the crop LAI, which has been found to be as an important parameter to provide early estimates of grain yield (Waldner et al. 2019), can be easily missed by the 16 and 5 day revisit times of Landsat (Jin et al. 2017b) and S2 (Clevers et al. 2017), respectively. Consequently, methods developed for yield estimation based on the peak of VIs (e.g. Franch et al. 2015) or LAI (e.g. Lobell et al. 2015) are not able to provide accurate yield estimates, or to target small farm holders fields. The proposed method allows time-series gaps due to clouds to be filled and improves the probability of identifying the peak LAI. However, it still faces some limitations in near real-time monitoring on cloudy days.

3.5.2. Estimating wheat Green LAI

Similar to previous studies (e.g. Djamai and Fernandes 2018; Dhakar et al. 2019; Djamai et al. 2019; Pasqualotto et al. 2019), this study found S2 LAI products (created using the Biophysical Processor within ESA's SNAP software (Weiss and Baret 2016)) are capable of estimating wheat Green LAI ($R^2 = 0.95$ and 1.08 RMSE). However, the S2-LAI product was found to be less suitable for estimating wheat senescence-LAI ($R^2 = 0.83$ and 1.9 RMSE). One of the strengths of this new method for estimating LAI is the fact that it combines both methods, i.e., the physically-based retrieval method (e.g. RTM) and the empirical approach (e.g. using VIs), to convert surface reflectance data into LAI estimates. First, the method uses the RTM-based S2-LAI product as a benchmark and then uses the fused VI image pixel values within the defined region of interest, as an automatic "field-based" calibration to convert the fused VIs into S2-like LAI estimates, through a series of linear regression models. These regressions, which are automatically generated for each day in the time-series, are uniquely fitted to the area of interest (e.g. the analysed field). This approach enables the method to be more robust and valid across different soil types, crop types and varieties, farm managements and environmental conditions.

In this study, 13 different VIs were tested (), having been indicated in the literature to be highly correlated with LAI, and evaluated their performance to estimate LAI in the new fusion method. It was found that overall this new method is not sensitive to a specific VI, with the fused Green LAI estimates from the various indices ranging from RMSE 1.37-1.4 ($R^2 = 0.92-0.94$). The MTVI2, SR and the RDVI were found to be the best performing VIs in this study. However, all

of these VI fuse-based estimates slightly underperformed the accuracy of S2 Green LAI estimations (Table 7) while also providing daily estimates at 3 m resolution.

The S2-LAI product has a few limitations. As noted by previous studies, similar to other remotely sensed LAI estimation (e.g. Houborg et al. 2016; Djamai et al. 2019), S2-based Green LAI estimates also tend to underestimate high Green LAI values ($LAI > \sim 3$) (Herrmann et al. 2011; Verrelst et al. 2015; Dhakar et al. 2019; Djamai et al. 2019; Pasqualotto et al. 2019). These underestimations are probably produced by the asymptotically saturation of the surface reflectance data caused by the high biomass density (Gitelson 2004). These underestimations become even more significant from $LAI = 6$ and higher (as shown in Figure 17), which causes the uncertainties in estimating high LAI values ($LAI > 6$) using SNAP's Biophysical Processor, as reported by Weiss and Baret (2016). The analysis of the new fused LAI time series showed similar underestimations as the S2-LAI, which is not surprising considering the fact that the fused-LAI was created using S2-LAI data.

Similar to other LAI products such as MODIS LAI (Myneni and Park 2015) or Visible Infrared Imaging Radiometer Suite (VIIRS) LAI products (Knyazikhin and Myneni 2018), the S2-LAI product uses a generic method to estimate LAI for any type of vegetation (Weiss and Baret 2016). Therefore, in order to have a better match for a specific crop type, a correction should be applied to calibrate the data (Weiss and Baret 2016). This study has developed regression models (Table 7) to adjust both the S2-based LAI and the fused-LAI estimates to provide more accurate wheat LAI estimates. The performance of the proposed correction, which was tested using an independent wheat LAI dataset measured in NSW, Australia, showed a clear improvement in the accuracy of the method to estimate wheat Green LAI. In 10 out of the 13 indices tested, the RMSE of Green LAI estimates improved by more than 25%, while four improved by more than 39% (Green WDRVI improved in 47%) in comparison with non-corrected Green LAI estimations (Table 8), with RMSE ranging between 0.35-0.63. The S2-LAI data was also tested, however the comparison between the pre-correction and the post-correction Green LAI estimates showed that S2-LAI underperformed the fused-based LAI estimates with 1.38 RMSE ($R^2 = 0.41$). The reason that the S2-LAI data was not able to reproduce the accuracy achieved by the fused data to estimate the Green LAI is due to the fact that the *in-situ* LAI in these fields was measured along the fields' boundaries, and therefore, the S2 data suffered from mixed pixels that lowered the Green LAI estimates. The mixed pixel effect has long been recognised as a main drawback to monitor crop performance and characterization from space, especially when using low and medium spatial resolution data such as the imagery acquired by MODIS and Landsat (Gao et al. 2006; Gao et al. 2012; Jain et al. 2016; Khan et al. 2018; Li et al. 2019). While LAI information driven from low to medium resolution satellite images may be lost for certain surface types that appear only at

smaller spatial scales (Gao et al. 2012), the 3 m fused LAI data presented in this study overcome these limitations.

Despite the fact that the proposed correction aimed to adjust the Green LAI phase of the wheat, the results shows that the accuracy of the LAI estimations of the senescing phase has also improved after applying the correction. Future study should attempt to develop an adjustment method that will target the crop's senescing phase only. As demonstrated in Figure 15, the new high-resolution dataset was able to better describe the spatial patterns of the crops and to identify vegetation with less active growth within the sub-field scale. Moreover, the new high spatio-temporal LAI estimates can be potentially used to monitor crops grown on small holder (<2 ha) farms in developing countries (Jain et al. 2016). Nevertheless, it is expected that the regression models proposed here to correct S2 Green LAI estimations will exceed those of the non-corrected wheat Green LAI estimates, when overcoming the mixed pixel effect. However, this should be further tested in future studies.

3.5.3. Limitations and prospects

Even with the promising results presented here, there are some limitations that should be noted. The fused daily 3 m LAI data was evaluated across two countries in four different geographic locations, over 12 wheat fields with diverse farm management practices, soil types, climates and varieties. However, the correction method presented in this study to adjust the S2 and fused Green LAI was tested in one geo-location only. Therefore, future studies should explore the performance of the proposed correction over a larger number of fields and environments.

While some previous studies have suggested that red-edge based VIs may help to mitigate the saturation problem encountered when estimating high LAI values using traditional VIs based on visible reflectance, such as NDVI (Nguy-Robertson et al. 2012; Dong et al. 2019). Nguy-Robertson et al. (2014) have shown that this is not universally true. Moreover, current PlanetScope imagery does not provide red-edge data and so was not applied in this study.

Even though it is one of the cheapest commercial high-resolution image products currently available in the market, Planet's PS images are not free like S2. Farmers and others, who may want to use such high spatio-temporal LAI and VIs data, should consider the cost effectiveness of this data. It is expected that high-performing farmers will find it very beneficial while individual farmers in developing countries may find the costs too high. As more CubeSat constellations come on line in coming years, prices of their imagery will likely be reduced.

The uncertainty of the PlanetScope data in terms of its geo-location accuracy has been reported as less than 10 m RMSE (Planet Team 2018). This uncertainty may affect the utility of the PlanetScope data when carrying out detailed time-series analyses (Houborg and McCabe 2018b). Implementing a co-registration practice to reduce the cross-scene co-registration error, similar to the co-registration technique proposed by Houborg and McCabe (2018b), is likely to increase the spatial correlation between consecutive scenes.

This study used a simple linear interpolation to fill data gaps between the cloud free images for both PS and S2 images to create evenly spaced time series. Despite the simplicity of this approach, previous studies showed that linear interpolation is an effective way to interpolate between periods with valid data to assign values to the periods of missing satellite observations with considerable accuracy (Sakamoto et al. 2010; Zhu et al. 2011; Pan et al. 2015; Maynard et al. 2016). Importantly, this study implemented a fusion method to monitor field crops, which commonly do not change over a single day. This method can therefore be used for other disciplines that also have a slow temporal evolution, such as forestry, land cover classification (Gašparović et al. 2018), geomorphological and environmental studies, or to monitor urban development over time. Nevertheless, the utility of its implementation for monitoring rapidly changing environments or phenomena such as flash floods or fires, should be further investigated and evaluated compared to change detection techniques (e.g. Sadeh et al. 2019).

As a prospect for future improvements and research directions, it is suggested that future studies should test the proposed fusion method over other crops types, explore the possibility of adding more sensors in the fusion process (e.g. Landsat) and examine the suitability of this fusion method to fuse other sensor data (other than S2 and PS). Although this study attempted to generate high resolution LAI, this method can potentially be useful to produce high spatio-temporal time series of Fraction of Absorbed Photosynthetically Active Radiation (FAPAR), Fraction of vegetation cover (FCOVER), Chlorophyll content in the leaf (Cab) and Canopy Water Content (CWC). Furthermore, future studies should explore the suitability of the proposed fusion method for improving the spatial and temporal data obtained from sensors operated in the shortwave infrared (SWIR) and the thermal wavelengths, such as those on-board Landsat and S2.

3.6. Conclusions

With the increasing number of CubeSat constellations expected to become operational in the coming years, a new era of Earth observing satellite-based applications has begun. This paper presents the first study to fuse time series imagery sourced from Sentinel-2 (S2) and Planets'

PlanetScope (PS) CubeSat constellation. The fusion method proposed in this study enabled S2-consistent, cloud free, surface reflectance RGB-NIR images and crop Green LAI to be generated at a 3 m resolution. Overall, the results from the study demonstrated that the new fused time-series data combined the spatial, temporal and spectral advantages of both sensors, allowing wheat Green LAI to be monitored on a daily basis with an RMSE of 1.37 and R^2 of 0.94 in wheat.

Furthermore, this study proposed a correction method to compensate the underestimations in high LAI values (> 3) between the remotely sensed LAI estimations and the *in-situ* measurements. With the implementation of the correction method, the accuracy of the Green LAI estimations improved by up to 47% (RMSE = 0.35-0.63).

Although tested to fuse S2 and PS data for LAI estimations, this new time series fusion method can be used to fuse other sources of imagery with different spectral, spatial and temporal resolutions. Furthermore, it may be used to estimate indices or parameters other than LAI. The proposed method is not limited to a specific number of bands, wavelengths or images, and can integrate numerous sources of imagery. This new time series fusion method can be used for continuous daily high-resolution monitoring of crops over large scales, and can potentially be used for a range of new precision agriculture applications. Such time-series are critical for crop health and growth status monitoring, and will improve the effectiveness of farming practices such as water management and fertilization, as well as improve yield forecasts.

Chapter 4 – The Versatile Crop Yield Estimator

4.1. Abstract

One of the major challenges in monitoring and managing food security is to provide reliable, consistent and scalable crop yield projections. Accurate production forecasts, as early as possible prior to the harvest, are critical for market stability, as well as for farmers, grains companies and governments. For decades, methods have been developed for using Earth observing satellite data to monitor crop conditions and their production across different spatial and temporal scales. These methods typically rely heavily on detailed official crop statistics or used ground-based data to develop empirical forecasting models, which limit their application to the regions where they were calibrated. Accordingly, this study proposed a new method named the VeRsatile Crop Yield Estimator (VeRCYe), which aimed to overcome the above limitation for wheat yield estimation at the field and pixel scales, by combining the advantages of both high spatio-temporal resolution remote sensing and crop model simulations. In this process, the sowing and harvest dates of each field were detected (RMSE = 2.6 – 2.7 days) using PlanetScope imagery. In addition, Sentinel-2 and PlanetScope data were fused into a daily 3 m LAI dataset. Finally, the detected sowing dates and the LAI datasets were coupled with the APSIM-Wheat crop model to estimate wheat yield at the field and pixel scales. This study tested the method over multiple wheat fields located in the Australian wheat-belt, covering a large range of pedo-climatic conditions and farm management practices across three growing seasons (2017 - 2019). VeRCYe estimated field-scale yield with $R^2 = 0.88$ (RMSE of 757 kg/ha, 15% error), and produced yield maps at 3 m resolution up to four months before crop harvest ($R^2 = 0.32$, RMSE of 1,213 kg/ha). The advantages of VeRCYe are that (1) it can be used to estimate yield without the need for ground calibration, (2) it can theoretically be applied to other crop types, and (3) it can be used with any remotely sensed LAI. Furthermore, VeRCYe can help to identify yield gaps, understand yield variability and its causes from the pixel-level to a regional-level.

4.2. Introduction

One of the major challenges in monitoring and managing food security is to provide reliable, consistent and scalable crop yield projections (Nakalembe et al. 2021). Therefore, accurate production forecasts as early as possible prior to harvest are critical for market stability, as well as for farmers, grains companies and governments (Hammer et al. 2001; Becker-Reshef

et al. 2020; Benami et al. 2021). Importantly, climate variability and extreme weather events are projected to increasingly affect future crop yields, potentially leading to severe food crises, risks and uncertainties within the global food system (Hammer et al. 2001; Ray et al. 2015; Feng et al. 2020). Spaceborne remote sensing is considered a reliable, affordable, large-scale, and timely source to improve crop yield prediction (Becker-Reshef et al. 2020), therefore many yield prediction methods using satellite data have been developed in the last few decades (e.g. Idso et al. 1977; Ferencz et al. 2004; Prasad et al. 2006; Franch et al. 2015). Traditionally, these methods are based on the correlation between Vegetation Indices (VIs) and crop yield (e.g. Raun et al. 2001; Labus et al. 2002; Becker-Reshef et al. 2010; Bognár et al. 2017). However, reliance upon a unique and local relationship is not ideal, especially when crops experience highly variable environmental conditions through space and time.

In the last decade, methods which combine satellite images and machine learning techniques have become very popular (e.g. Cai et al. 2019; Jeffries et al. 2019; Feng et al. 2020; Kamir et al. 2020). However, these methods often require a large amount of ground data from different sources, including yield, sowing dates, soil properties, cultivars, farm management practices and weather for training and calibrating the model. Such data is rarely available for yield estimation over large scales such as at the district or country level (Feng et al. 2020), and therefore these models are calibrated locally. While locally calibrated yield estimation methods may achieve good accuracy of yield estimation (Donohue et al. 2018; Chen et al. 2020), the use of these methods is usually limited to the area in which they were calibrated.

Despite the growing availability of Earth observing data to monitor crop development and yield estimation, use of spaceborne sensors is limited by the type of data they can retrieve. Optical remote sensing cannot see through the crop canopy or the soil surface but it can, for example, provide valid information about canopy chlorophyll content (Gitelson et al. 2005). Synthetic Aperture Radar (SAR) may provide complementary data on the surface's roughness, slope, geometry and the soil moisture (Walker et al. 2004; Sadeh et al. 2018). Conversely, crop growth models can be used to simulate key physiological processes including phenology, organ (such as leaf and grain) development, water and nutrient uptake, biomass, and response to abiotic stresses (Holzworth et al. 2014; Huang et al. 2019). Therefore, merging the capabilities of remote sensing with crop model simulations has a great potential for improving capabilities in monitoring crop development and yield estimation through space and time.

One of the ways to blend the abilities of crop models and remotely sensed data is by using data assimilation techniques, and many of these studies have focused on the assimilation of Leaf Area Index (LAI) into the models (Ines et al. 2013; Huang et al. 2015; Huang et al. 2019;

Pan et al. 2019). LAI has been found to be a good indicator of crop status and leaf abundance, as well as phenological stage, and can be used as an indicator of different farm management practices, or the impact of biotic and abiotic stresses (Huang et al. 2019). However, often these data assimilation techniques required local calibration through field measurements (Pan et al. 2019; Beyene et al. 2021; Manivasagam et al. 2021), which limit their ability to estimate crop yield over large areas or in environments different from where the calibration data was collected.

In order to bypass the need for ground calibration data, Lobell et al. (2015) developed a new approach named the scalable satellite-based crop yield mapper (SCYM), which uses crop model simulations to train a regression that relates final crop yield to observed values of VIs for available satellite images during the growing season (Lobell et al. 2015; Azzari et al. 2017). In the process, Lobell et al. (2015) used the Agricultural Production Systems sIMulator (APSIM) (Holzworth et al. 2014) to generate a large number of crop model simulations that span a realistic range of soil, climate, and management settings for a specified region. The advantage of using APSIM is that each individual simulation provides output on daily crop attributes, such as LAI and yield (Lobell et al. 2015). SCYM uses published equations to convert the LAI to optical based VIs, providing per pixel yield predictions by applying a regression to Landsat-based VI and gridded weather data (Lobell et al. 2015; Azzari et al. 2017). Despite the potential of this innovative approach, the accuracy of the method was limited. Moreover, its operational capability in predicting crop yields was also reported at limited, particularly in smallholder farms due to limitations in the spatial resolution and temporal frequency of satellite images (Azzari et al. 2017; Nakalembe et al. 2021).

For years, the trade-off between high spatial and temporal resolution coverage has limited remotely-sensed applications such as crop yield estimation at the field and sub-field scales (Waldner et al. 2019). In order to fill that gap, a number of companies developed and launched Earth observing CubeSats to a low Earth orbit. These new satellites, such as Planet Labs' PlanetScope (PS) CubeSat, are relatively inexpensive to build, thereby making it possible to acquire large image collections at high spatial and temporal resolutions at a relatively low cost. However, images obtained from CubeSat constellations, such as Planet's PS, frequently suffer from radiometric inconsistencies in the data collected by the different satellites within the constellation, due to inter-calibration challenges and their low signal-to-noise ratio. (Houborg and McCabe 2016; Houborg and McCabe 2018b; Leach et al. 2019; Sadeh et al. 2019). To overcome this issues, some methods have been developed to fuse CubeSat data with other satellite images such as Sentinel-2 (S2) or Landsat to produce consistent high spatio-temporal LAI datasets (Houborg and McCabe 2018a; Kimm et al. 2020; Sadeh et al. 2021). While recent studies have shown that improvement in crop yield estimations can be

expected by using more frequent high-spatial and high-temporal satellite images per growing season (Jain et al. 2016; Jin et al. 2017a; Waldner et al. 2019; Manivasagam et al. 2021), the potential of using such a unique high spatio-temporal LAI dataset to improve yield estimations has yet to be fully evaluated.

While most studies attempted to estimate crop yield at regional, state or national scales (e.g. Ines et al. 2013; Huang et al. 2015; Azzari et al. 2017; Jin et al. 2017b; Cai et al. 2019; Jin et al. 2019), few studies have attempted to estimate yields without calibration through ground data (e.g. Becker-Reshef et al. 2010; Franch et al. 2015; Lobell et al. 2015; Azzari et al. 2017; Jin et al. 2019). In addition, a relatively small number of studies tried to estimate yield at the pixel and field scales (e.g. Donohue et al. 2018; Lai et al. 2018; Chen et al. 2020; Manivasagam et al. 2021; Sagan et al. 2021), but only very few attempted doing so without any ground-based data for calibration (e.g. Jain et al. 2016; Burke and Lobell 2017; Dado et al. 2020; Deines et al. 2021), achieving limited success. To overcome these limitations, it appears promising to develop new methods that combine earth observation data with data generated from crop growth models, and thus eliminate the need for *in-situ* yield measurement and to preform global yield monitoring (Lobell et al. 2015; Waldner et al. 2019).

The objective of this study was therefore to develop a new approach for estimating crop yield at the field and pixel scales, without relying on ground data for calibration. The proposed method was tested on estimating wheat yield, as wheat is an agricultural commodity which has an important place in the global food production. However, since it is typically grown in relatively arid regions of the world, it is particularly vulnerable to climate variability (Hammer et al. 2001), hence the importance of the success of the method. First, a CubeSat-based sowing date detection method (Chapter 2) was used to identify cultivated fields and the date when they were sown. Second, PS images and S2 images were fused to create daily LAI datasets at 3 m resolution (Chapter 3). Finally, the detected sowing dates and the field's LAI datasets were coupled with the APSIM-Wheat crop model (Holzworth et al. 2014) to estimate wheat yield 2-3 months before the harvest at the field scale and at 3 m pixel size within the field.

Lastly to identify where and when the estimated yield has been harvested and become available to be trade in the market, this study tested the ability of the sowing date detection method to also detect the timing of the harvest for the studied fields.

4.3. Methodology

This study developed a new method named VeRCYe – VeRsatile Crop Yield Estimator (pronounced as “versi”) for estimating crop yield at the field and pixel scales. VeRCYe uses the LAI as the linking parameter between the remotely sensed (RS) data and APSIM’s plant development and yield estimation. It includes four main steps:

1. A CubeSat-based sowing date detection method developed to identify cultivated fields and the date when they were sown (Chapter 2).
2. Fusion of PlanetScope (PS) images and Sentinel-2 (S2) images to create daily LAI datasets at 3 m resolution (Chapter 3).
3. Coupling of the sowing dates and Leaf Area Index (LAI) datasets with The Agricultural Production Systems sIMulator (APSIM)-Wheat crop model to estimate wheat yield at the field scale.
4. Detection of the harvest date for each field using the same methodology as for the sowing date detection in step 1 but at the end of the season.

The accuracy of the estimated yield was evaluated against the reported yield from 27 farmers’ fields (Table 9), including 21 fields sourced from the National Paddock Survey (Lawes et al. 2018). Fields were located across five Australian states. Reported data spanned three growing seasons (2017 – 2019) and corresponded to various weather conditions, soils, farm management practices and wheat cultivars. It is important to note that at no stage ground-based data (included yield) was used in this study for training or calibration purposes.

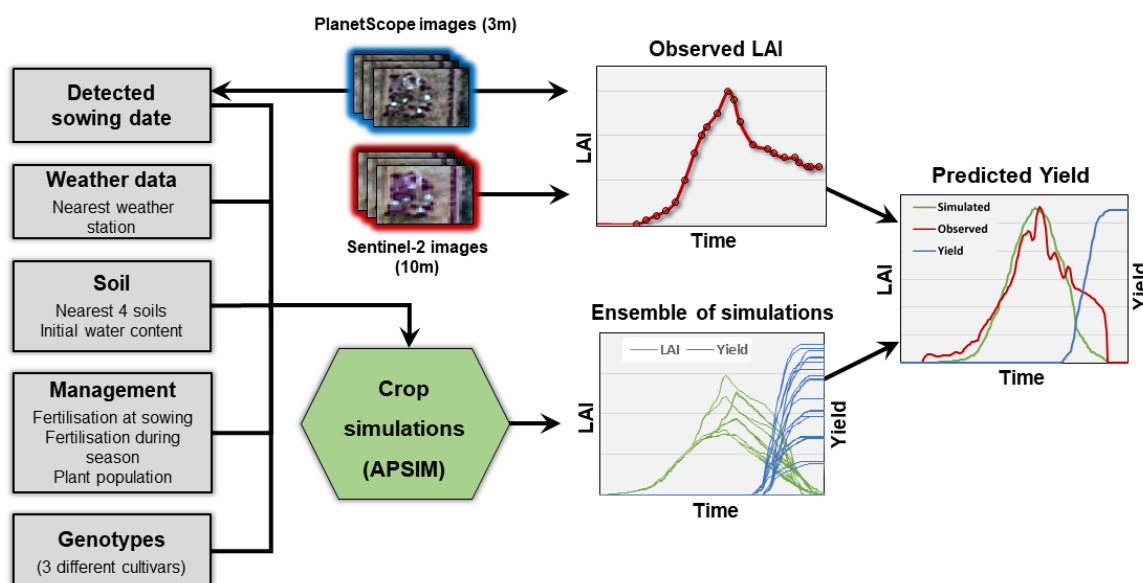


Figure 19. VeRCYe’s field-scale yield prediction workflow

Table 9. This table summarizes the field-scale reference data used in this study to evaluate the performance of the proposed method. Overall, 27 fields were analysed, however, the sowing and harvest dates were not reported for all fields and for some the harvester data was missing.

	Field-scale yield data	Yield map (harvester data)	Sowing dates	Harvest dates
Number of fields	27	22	22	20

4.3.1. Sowing and harvest date detection

In order to extract the sowing dates for each field, the semi-automated sowing date detection proposed in Chapter 2 was implemented. By contrast to Sadeh et al. (2019) who used Equation 1 ($\text{Change} = \frac{\text{Image}_{t_1}^{PC_1}}{\text{Image}_{t_2}^{PC_1}}$, Page 17) to identify the change between two consecutive

images, the current chapter used the following as it provided improved performance:

$$\text{Change} = \text{Image}_{t_2}^{PC_1} - \text{Image}_{t_1}^{PC_1} \quad \text{Equation 4}$$

where $\text{Image}_{t_1}^{PC_1}$ is the first principal component of the earlier satellite image and $\text{Image}_{t_2}^{PC_1}$ is the first principal component of the later satellite image.

This study also tested the ability of the sowing date detection method to detect the harvest dates and the field area that had been harvested between two consecutive images. It was found that the sowing date detection method was effective in detecting the harvested area of the field, after modifying the “change” equation (Equation 4) to:

$$\text{Change} = \text{Image}_{t_1}^{PC_1} - \text{Image}_{t_2}^{PC_1} \quad \text{Equation 5}$$

where here the $\text{Image}_{t_1}^{PC_1}$ is the first principal component of the earlier satellite image and $\text{Image}_{t_2}^{PC_1}$ is the first principal component of the later satellite image. This modification was required, as sowing often corresponds to a change in color from bright to dark, while at harvest, the field changes from dark brown to the light brown-yellow color of the crop residue.

The accuracy of the detection of both sowing and harvest dates was evaluated against the dates reported by farmers. Out of all 27 fields used in this Chapter, only 22 had reported sowing dates and 20 had reported harvest dates.

4.3.2. APSIM model simulations

APSIM Next Generation crop model (Holzworth et al. 2018) allows the running of numerous possible scenarios that represent a realistic range of environmental conditions and farm

management practices. This helps to overcome the gaps in knowledge of the farm management practices used in specific fields. As part of this process, ~2,000 simulations of APSIM are generated for each field (Figure 19). The weather data were taken from the nearest weather station (www.bom.gov.au/) and the soil properties from the four nearest soils available in the APSoil database (www.apsim.info/apsim-model/apsoil/). Each APSIM simulation outputs daily crop characteristics including LAI as well as a grain yield estimation (kg/ha). Simulations that best reflect the LAI evolution selected using an automatic rule-based algorithm (to be described below) to estimate the likely final yield.

To simulate the range of plausible scenarios for each studied crop, the following information was used (see summary in Table 10):

Weather data - The weather records were downloaded from the nearest weather station to the field from the SILO database. SILO is a database of historical climate data for Australia (Jeffrey et al. 2001). It is maintained by the Queensland Government and the datasets are constructed using observed data provided by the Australian Bureau of Meteorology (SILO 2018). The weather data file includes daily radiation (MJ/m²), min and max temperature, rain and evapotranspiration (mm).

Sowing date – the sowing dates used in APSIM were the dates detected using the adapted sowing date detection method from Chapter 2, as described above. These sowing dates were also used as the start date for APSIM simulations.

Soil characteristics – Australian wheat crops are mostly grown under rain-fed conditions and so heavily rely on water available in the soil profile. Therefore, choosing an appropriate soil to represent the field is important to obtain reliable results. Soil characteristics were sourced from the APSoil database, which is a repository of soils developed for use by the APSIM cropping systems model (Holzworth et al. 2014). APSoil is focused on the physical and chemical soil characteristics that drive crop production, particularly soil water and crop nutrition (Dalglish et al. 2016). In this study, the nearest 4 soil types were used alternatively to generate the simulations. This increased the likelihood to consider a soil with similar properties as the soil from the field of interest.

Cultivars – For each field, three cultivars were chosen to represent early, mid, and late maturing cultivars from each region, as proposed by Zheng et al. (2015) and by analysing the popular cultivars used by farmers in each region as reported in the National Paddock Survey. The cultivars used for each region in the Australian wheat-belt were: South-East: Axe, Mace, Gregory, South-West: Mace, Scepter, Yitpi and East: Suntop, Baxter, Lancer.

Initial soil water – The soil water at the start of the simulation was set as 20%, 50% or 100% of the plant available soil water capacity (PAWC) of the soil, with the soil profile filled from the top (typically due to rainfall refilling the soil during the intercropping season).

Fertilization at sowing – Farmers commonly fertilize the soil at sowing, however the rate of fertilizers used typically varies across regions. Three different rates were used to cover the range from current common practice in Australia (Chenu et al. 2013). Those were 30, 50 or 100 kg/ha of nitrogen, in the form of NO_3 .

Fertilization during the season – Similar to the fertilization at sowing, the rate of fertilizers used by farmers during the growing season typically vary across regions of the wheat-belt but also depending on seasonal rainfall. As the actual dates of the fertilization are unknown for the studied fields, different optional rates were applied in the simulations to cover common practice in Australia following Chenu et al. (2013). These rates were 0, 30 or 60 kg/ha of NO_3 at the stage ‘stem elongation’ (also referred to as Zadoks’ growth stage 31 (Zadoks et al. 1974)) and 0 or 30 kg/ha of NO_3 at the ‘booting’ stage (i.e. Zadoks’ growth stage 40).

Plant population – The number of plants per m^2 , or plant ‘density’, was set at 50, 100 & 150 plants per m^2 , based on common practices reported in Chenu et al. (2013) and in the National Paddock Survey.

4.3.3. Remotely sensed LAI

The newly developed method to fuse PlanetScope and Sentinel-2 imagery into daily 3 m LAI (Sadeh et al. 2021) was used to generate a times-series of LAI for each of the fields analysed. This Chapter tested the ability of using both the original remotely sensed LAI time-series, being equivalent to the generic S2-LAI product (but in 3 m daily datasets) and the modified remotely sensed LAI dataset, which adjusted the generic S2-LAI product estimations to better estimate wheat Green LAI (Sadeh et al. 2021).

4.3.4. Coupling APSIM model simulation with remotely sensed LAI for field scale yield estimations

For each field, ~2,000 different simulations of APSIM were generated spanning a realistic range of possible environmental and on-farm variables (as described above). Simulations

Table 10. Inputs used to run APSIM simulations at each studied field.

Parameters	Inputs / Rules
Constant parameters	
Sowing date	The detected sowing date
Sowing depth (mm)	30
Row spacing (cm)	25
Duration of rainfall accumulation (days)	7
Changing parameters (factorials)	
Weather data	From the nearest weather station
Cultivars	Early, mid, and late maturing local cultivars
Plant population (plants per m ²)	50, 100, 150
Soil characteristics	The nearest 4 APSOIL soils
Initial soil water	20%, 50%, 100% of the plant available water capacity from the soil
Fertilization at sowing (NO ₃ kg/ha)	30, 50, 100
Fertilization at Zadok Stage – 31 (Stem elongation stage) (NO ₃ kg/ha)	0, 30, 60
Fertilization at Zadok Stage – 40 (Booting stage) (NO ₃ kg/ha)	0, 30

most likely to accurately predict the yield of the field of interest were selected using the

following steps. The selection steps all focused on LAI data around the field’s peak LAI, within a target time window when the Remotely Sensed (RS) LAI values were $\geq 60\%$ of the RS max LAI (LAI_{max}) (Figure 20).

- Step 1.** Find the highest 40% of the remotely sensed LAI values during the season and their timing. The threshold 40% was chosen following sensitivity tests which showed that there is no need of using the entire season data to produce and accurate yield estimation. This criterion is commonly reached about two months before harvest on average.
- Step 2.** For each APSIM simulation, extract only the simulated LAI that fall within the duration of the highest 40% of the remotely sensed LAI values.
- Step 3.** Calculate the following variables for each APSIM simulation:
- The gap in LAI between the max simulated and RS LAI (e.g. on the Y axis).
 - The gap in days between the timing of max simulated LAI and max RS LAI (e.g. on the X axis).
 - RMSE between simulated and RS LAI (for the highest 40% of the remotely sensed LAI values), assuming that low RMSE represents a good match between the RS and the simulated LAI.
 - RMSE between simulated and RS Green LAI, representing the stages when the leaves are photosynthetically active (Daughtry et al. 1992), in the range of 1 to maximum RS LAI.
 - RMSE between simulated and RS Senescence LAI, representing the stages when the leaves are not photosynthetically active (Delegido et al. 2015), in the range between the maximum RS LAI and 1.

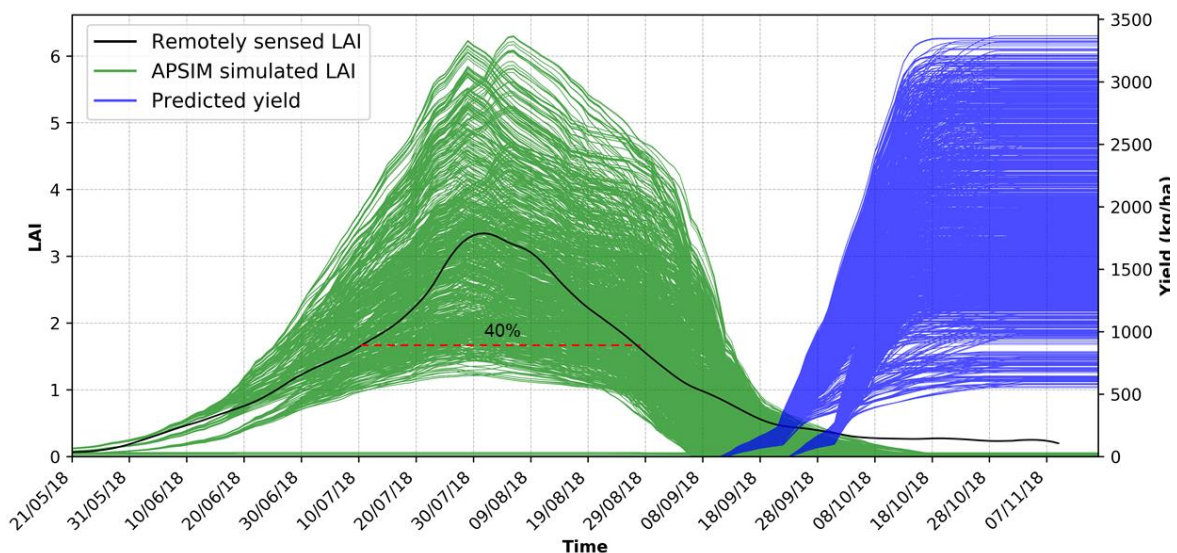


Figure 20. An example of ~2,000 different simulations of APSIM that were generated for a field, spanning a realistic range of possible environmental and on-farm variables. The green lines represent the simulated LAI and the blue lines represent the associated yield predictions. The remotely sensed LAI in black and the dashed red line illustrates the field’s highset 40% of the remotely sensed LAI values during the growing season. These were used to select APSIM simulation with most-similar LAI patterns, and estimate the range of plausible yields.

Step 4. Selecting only the simulations with the lowest 20% gap in LAI between the maximum (peak) simulated and RS (Figure 22).

Step 5. From the simulations selected in Step 4, select only the simulations with a gap in days between the timing of the max simulated and RS LAI that is within a range of ± 5 days. If none of the simulation answers this rule, then the selection range will increase to ± 10 days gap between the timing of the max simulated and RS LAI. If still none of the simulations meets this rule then it will increase to ± 15 , 20, 25 and eventually 30 days (Figure 21).

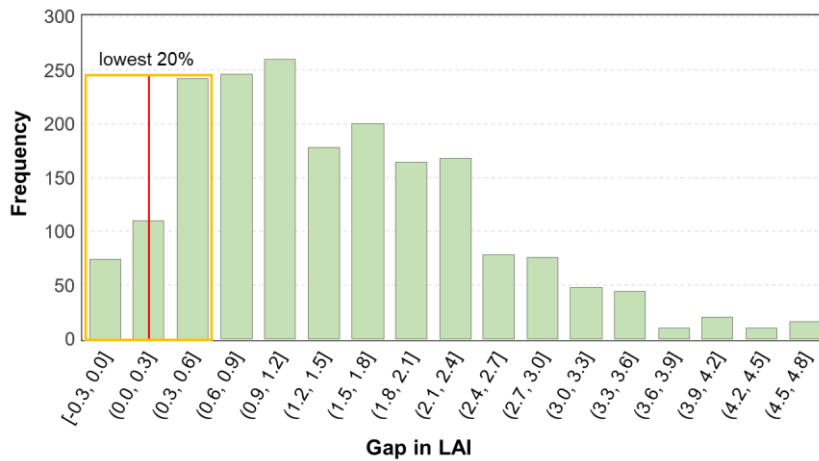


Figure 22. Illustration of the selection process of the lowest 20% gaps between the season’s maximum (peak) of simulated LAI and the maximum RS LAI. This figure shows the histogram of all the gaps calculated for a specific field, when the yellow box represents the simulations which has the lowest 20% values, while the red line represents a perfect match between the maximum simulated LAI and the maximum RS LAI (i.e. gap = 0 LAI).

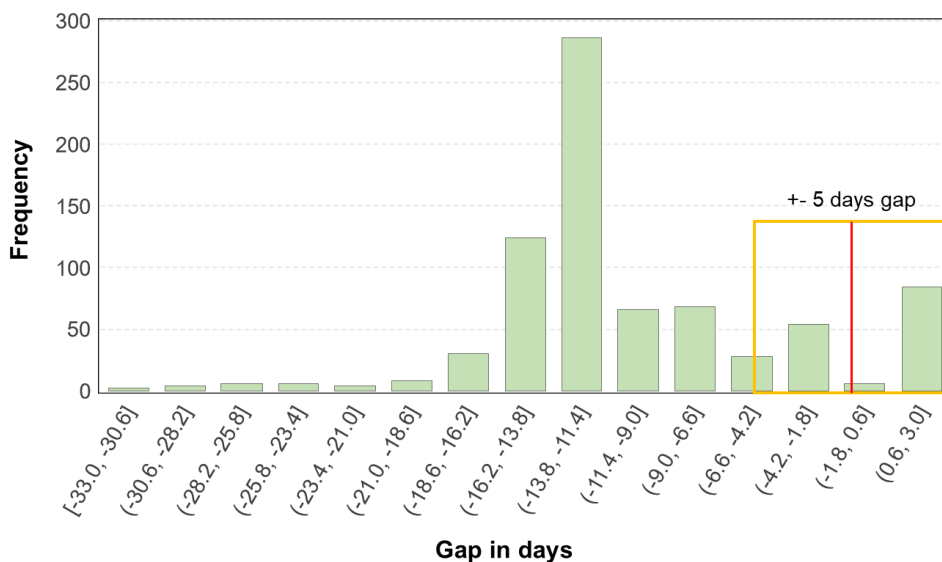


Figure 21. This figure shows the histogram of the gap in days between simulated and RS LAI of all simulations selected in the previous step. This histogram illustrates the selection rule of the simulations in which their gap in days between the timing of the max simulated and RS LAI is within the range of ± 5 days gap (the yellow box). The red line represents a perfect match between the timing of the maximum simulated LAI and the maximum RS LAI (i.e. gap = 0 days).

Step 6. An attempt to estimate the yield of a field by identifying the best fit simulated LAI to the remotely sensed LAI constantly resulted in underestimation of the forecasted yield. This study found that the simulations which ended with high accuracy of yield estimations (in comparison to the reported yield) frequently had higher simulated LAI during the senescence period than the RS Senescence LAI. This aligned with the finding of Chapter 3, in which the S2-LAI product was found less suitable for estimating wheat Senescence LAI than Green LAI. An illustration of the underestimation of the RS Senescence LAI is shown in Figure 23.

In order to overcome the underestimation of the remotely sensed Senescence LAI, the simulations that will continue for the next step must be simulations with the highest 20% of the average Senescence LAI (of the simulations selected in the last step). The threshold of 20% resulted from sensitivity tests conducted to evaluate which percentage would best perform in this process. The sensitivity tests aimed to identify the smallest possible percentage in order to minimize the sample size of the data analysed to save processing time. A breakdown of the simulated LAI and their associated estimated yield is shown in Figure 24. In this figure the remaining 389 APSIM simulations (out of

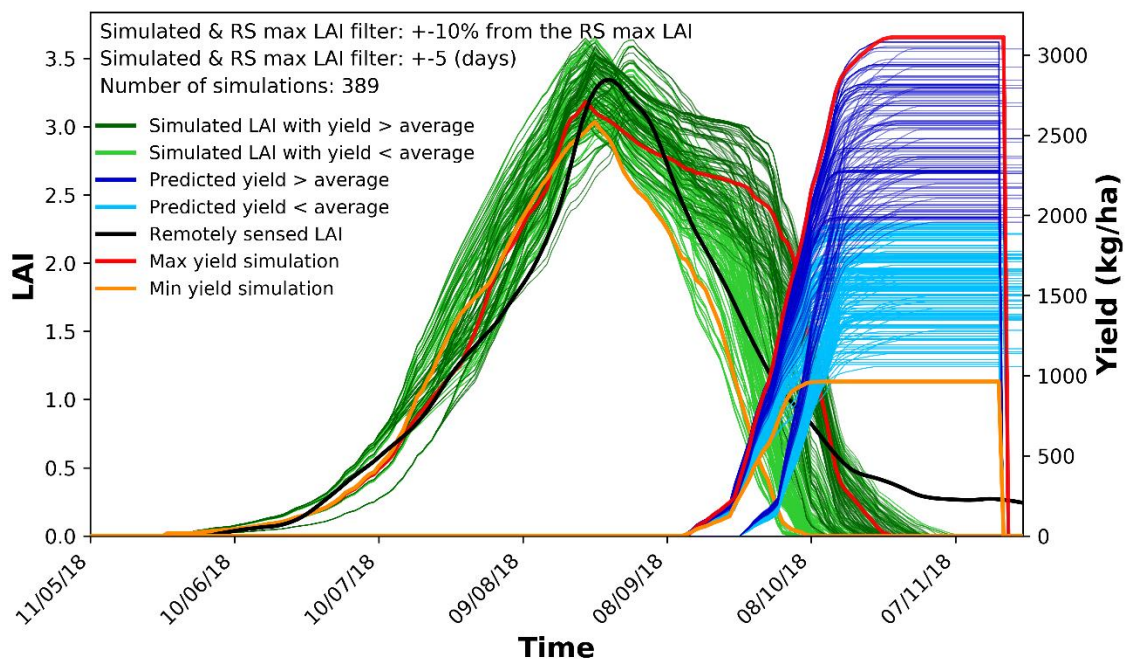


Figure 24. A breakdown of the simulated LAI and their associated estimated yield after applying the gap in max LAI value and timing filters, which resulted in 389 selected simulations (out of ~2,000). The simulations whose estimated yield ended above the average of all ~2,000 simulations are colored in a darker color (dark green for simulated LAI and dark blue for its associated estimated yield) then the simulations which their estimated yield ended to be below the average (light green for simulated LAI and light blue for its associated estimated yield). The simulation that resulted in the lowest yield is highlighted in orange while the simulation that resulted in the highest yield is highlighted in red while the remotely sensed LAI is in black.

~2,000 initial simulations), which resulted from the gap in max LAI value and timing filters (steps 4 & 5 above) are plotted. In Figure 24, the simulations for which their estimated yield ended to be above the average of all ~2,000 simulations are colored in a darker color (dark green for simulated LAI and dark blue for its associated estimated yield) than the simulations which their estimated yield ended to be below the average (light green for simulated LAI and light blue for its associated estimated yield). The simulation that resulted in the lowest yield is highlighted in orange while the simulation that resulted in the highest yield is highlighted in red. Figure 24 shows that there is no clear trend which simulation will result in a higher, and therefore more accurate yield estimation, when looking only at the Green LAI stage. However, at the senescence stage a clear trend was found with the simulations having a low Senescence LAI likely to result in lower yield estimation, while no clear correlation could be seen for Green LAI and yield. The result of this rule (Step 6) is presented in Figure 25.

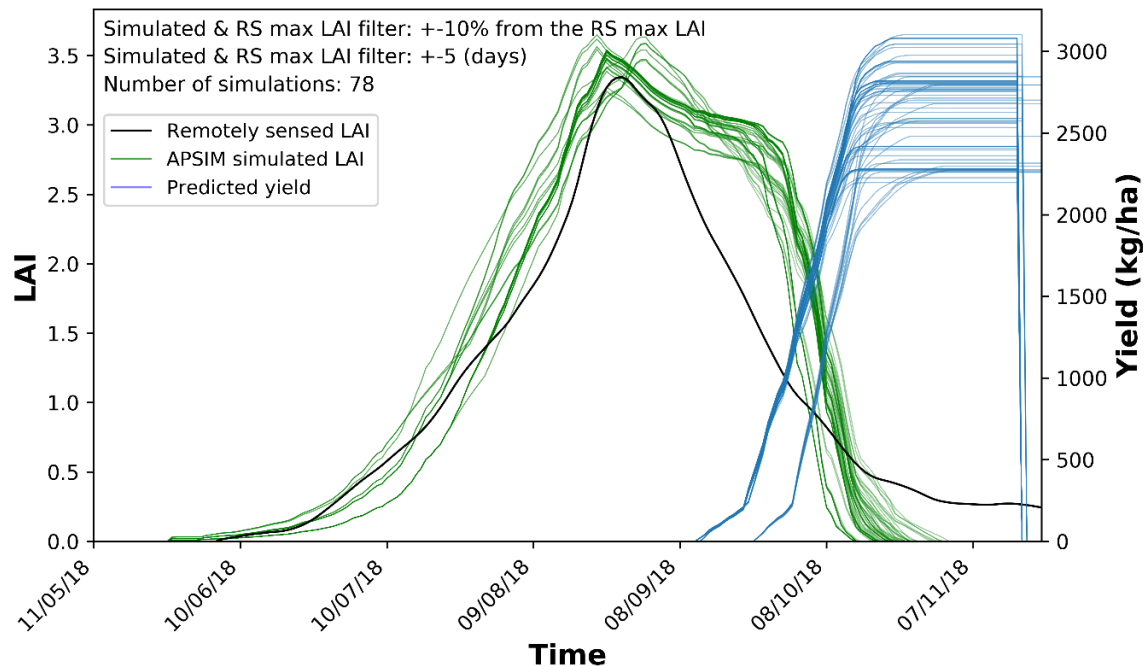


Figure 25. This figure shows the simulations who met all the conditions in the previous steps. In this example 78 simulations were selected out of 389 in the previous step.

Step 7. Finally, the estimated field-scale yield is set to be the average of the simulations with the lowest 20% RMSE between simulated and RS Green LAI (low RMSE represents a good match between the simulated and RS LAI). An example of the output of the field-scale yield estimation is shown in Figure 26. In order to cover different scenarios, if step 6 results with less than 10 simulations, then the estimated yield is set to be the average

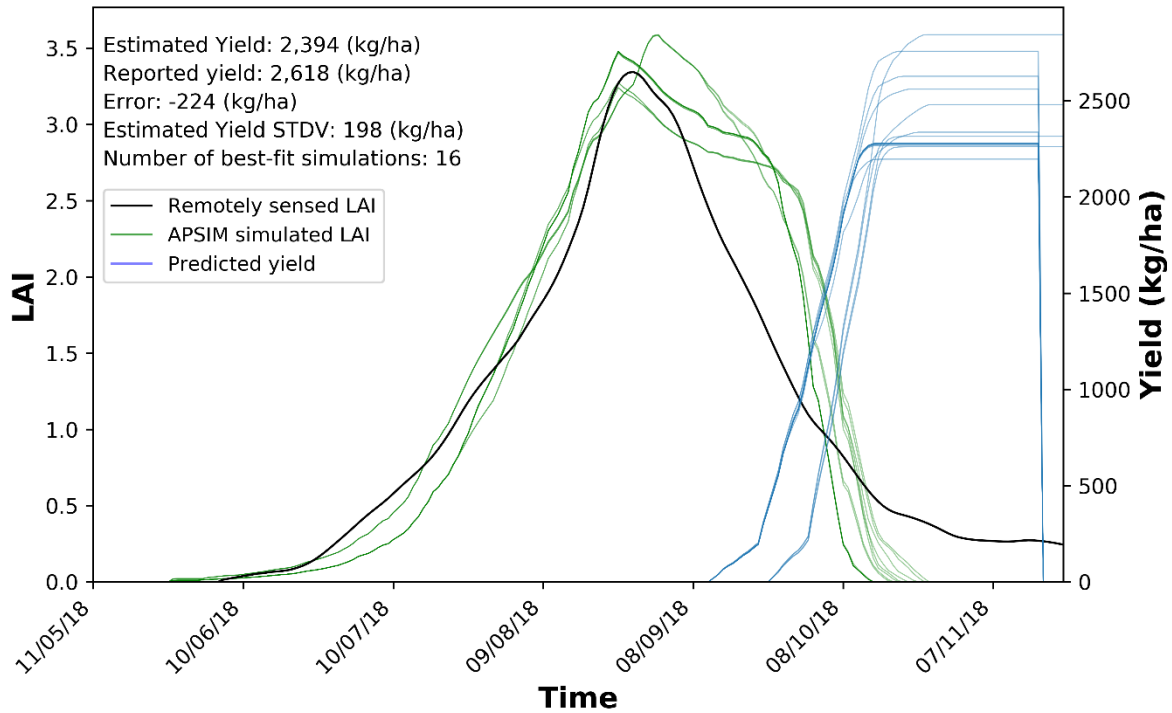


Figure 26. An example of the output of field-scale yield prediction of a wheat field located near Mallala in South Australia.

of all these simulations, i.e. without applying Step 7. In the scenario of an extremely low yield, such as during a severe drought, the ability to accurately estimate LAI using satellites is very limited. While crop models will still simulate crops with very low LAI in such scenarios, the extreme under-developed crop surrounded by bare soil is typically associated with a reduced RS LAI dramatically owing to the mixed pixel effect (Gao et al. 2012). Such crops typically have a very low yield and therefore should be addressed as a worst-case-scenario. Consequently, in case that the maximum RS LAI was lower than 0.9, the estimated yield is set to be the average of the three simulations with lowest yield estimation.

4.3.5. Generating yield maps at 3 m spatial resolution

This study also predicted yield at the pixel scale two months before the harvest. This was done using the 3 m daily LAI maps produced from the fusion between PS and S2 (Sadeh et al. 2021), which was converted to yield at the pixel level. In this process, a Conversion Factor (CF) (Equation 7) was used to convert LAI maps to yield maps (kg/ha). The CF was calculated as:

$$\text{Conversion Factor} = \frac{\text{Estimated Yield}}{\text{Remotely Sensed LAI}_{\text{max}}} \quad \text{Equation 6}$$

where Estimated Yield is the estimated field-scale yield, as described in section 4.3.4 and Remotely Sensed LAI_{max} corresponds to the season's maximum field-scale median LAI value from the RS LAI map, for the day when RS LAI was detected as being the maximum, within that field, during the growing season. Next, each pixel of the LAI map (of the Remotely Sensed LAI_{max}) was multiplied with the CF, which converted the LAI values into yield (kg/ha) at the pixel level. This process resulted in a yield map at a spatial resolution of 3 m.

Combine harvesters equipped with yield monitors collect geolocated point yield data during the harvest (Fulton et al. 2018). In this study the harvesters' raw point measurements (commonly provided at a density of 10 m) were interpolated to a grid using the Inverse Distance Weighting (IDW) interpolation (Bartier and Keller 1996) into standardized 5 m yield maps, after removing outlier measurements of less than 100 kg/ha or above 10,000 kg/ha, as well as data points located within 5 m of the field boundaries. Finally, the generated yield maps were smoothed by using a low pass filter (3 by 3 pixels kernel) and used to assess the accuracy of yield maps estimated based on the RS LAI map and APSIM simulations.

4.4. Results

4.4.1. Sowing and harvest date detection accuracy

Implementing the sowing date detection method resulted in the accurate detection of sowing in 20 out of the 22 (90.9%) fields analysed (Table 11). There was only an average 0.95-day gap (0.5-day gap for the median) between the detected and reported sowing dates (RMSE = 2.7 days).

When using this method to detect harvest dates, this study showed that after making the small adjustments for harvest detection the method was also suitable for detecting harvested area and its timing (Figure 27). Implementation of the method resulted in the detection of harvest dates for all 20 analysed fields. Furthermore, there was only an average -0.1-day gap (0-day gap for the median) between the detected and the reported harvest dates (RMSE = 2.6 days). A summary of these results is presented in Table 11.

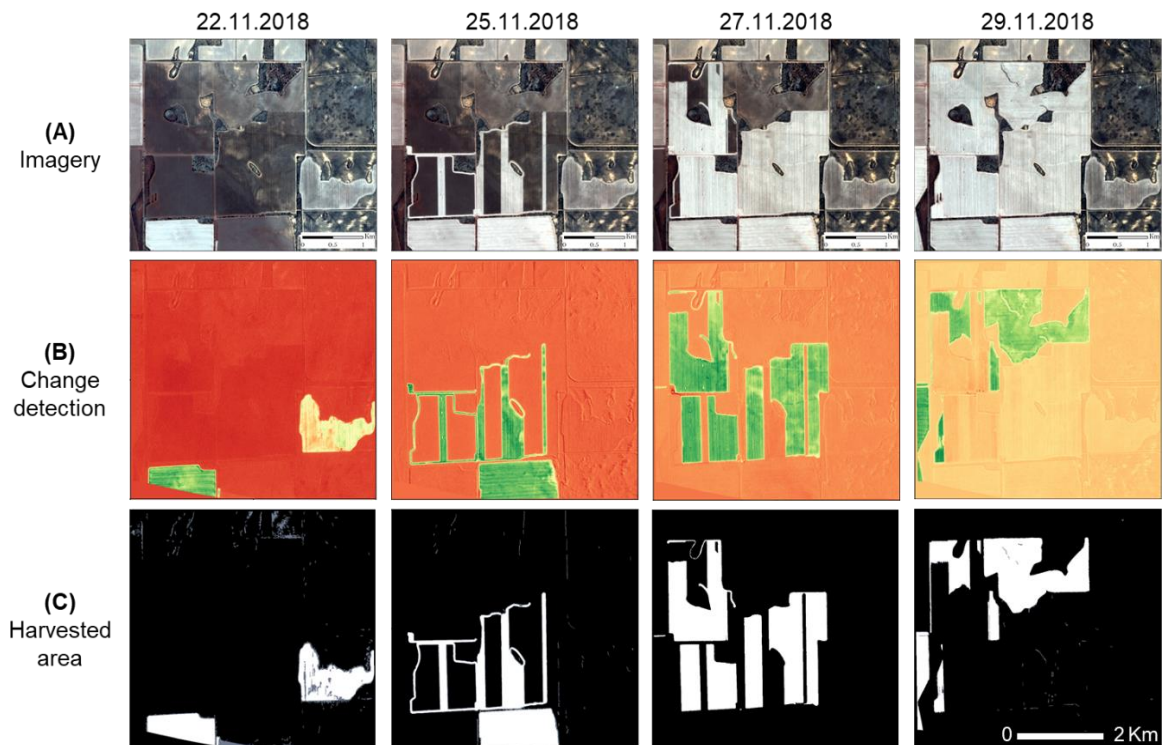


Figure 27. Example of harvest detection. This figure illustrates the harvest detection of a farm land 1,400 ha in size, near Mullewa, Western Australia, using four PS images taken over eight days. In (A) the pre-harvested wheat can be seen in a dark brown color, while the harvested area has bright yellow/grey colors. In (B), resulting images of subtracting $Image_{t_1}^{PC_1}$ from $Image_{t_2}^{PC_1}$, where a change between the images resulted in high values (green) and negligible changes resulted in low values (red). The area classified as harvested is shown in white (C) and the grey areas classified as noise.

Table 11. The performance of the sowing date detection method (Sadeh et al. 2019) in detecting the sowing and harvest dates of fields analysed.

	Sowing date detection	Harvest date detection
Average error (days)	0.95	-0.1
Median error (days)	0.5	0
RMSE (days)	2.7	2.6
Detected fields	90.9% (20/22)	100% (20/20)

4.4.2. Field-scale yield estimations accuracy

The ability of VeRCYe to estimate yield at the field-scale was tested over 27 fields, using (i) the fused 3 m daily LAI time-series which is equivalent to the generic S2-LAI (original) as well as (ii) the fused LAI time-series which was adjusted to better estimate wheat Green LAI (corrected). In addition, for each of these datasets, this study investigated which of the 13 different vegetation indices (VIs) tested by Sadeh et al. (2021) to fuse PS and S2 into high spatio-temporal resolution LAI, resulted in best performances for the VeRCYe approach.

The results, as shown in Table 12, indicate that when using the fused-LAI equivalent to the original generic S2-LAI, VeRCYe was able to estimate field-scale yield with an RMSE of 971 kg/ha, and an average and median error of -740 kg/ha and -573 kg/ha respectively (for the best performing VI). The R^2 between the yield estimates using this dataset and the reported

Table 12. Performance of VeRCYe field-scale yield estimations (n=27) using either (i) the fused-based LAI dataset equivalent to the generic S2-LAI (original) or (ii) the adjusted fused-based LAI that correct for underestimation of high LAI values (LAI > 3). This table shows which of the 13 different vegetation indices (VI's) used in Sadeh et al. (2021) to fuse PS and S2 into high spatio-temporal resolution LAI, resulted with the most accurate yield estimation. The best performances in each performance metric is coloured in red.

VI	NDVI	EVI2	MTVI2	MSAVI	WDRVI	GREEN WDRVI	GCVI	OSAVI	GSR	GNDVI	RDVI	TVI	SR
Original LAI													
Average error (kg/ha)	-845	-856	-740	-807	-847	-833	-827	-851	-826	-835	-774	-819	-835
Median error (kg/ha)	-868	-868	-573	-573	-868	-898	-870	-868	-870	-868	-653	-675	-833
RMSE (kg/ha)	1049	1059	971	1031	1044	1006	993	1053	993	1017	1002	1025	1038
R^2	0.85	0.85	0.85	0.84	0.86	0.88	0.89	0.85	0.89	0.87	0.84	0.86	0.86
Adjusted LAI													
Average error (kg/ha)	-545	-558	-550	-572	-594	-563	-561	-538	-575	-556	-519	-579	-627
Median error (kg/ha)	-378	-378	-511	-488	-488	-525	-380	-378	-488	-525	-438	-488	-554
RMSE (kg/ha)	817	850	817	834	887	845	854	809	866	832	757	829	913
R^2	0.86	0.84	0.86	0.86	0.83	0.85	0.84	0.86	0.84	0.85	0.88	0.86	0.83

yield ranged between 0.84 and 0.89 for all VIs tested, while overall the MTVI2-based fused LAI outperformed the other VIs for most of the performance metrics.

Using the adjusted LAI improved the accuracy of the field-scale yield estimation substantially with an RMSE of 757 kg/ha, and an average and median error of -519 kg/ha and -438 kg/ha respectively (for the best performing VI). The R^2 between the estimated and the reported yield ranged between 0.83 and 0.88 for all VIs tested, while overall the RDVI-based fused LAI outperformed the other VIs for most of the performance metrics. Therefore, RDVI was chosen as the preference VI to be used in VeRCYe.

Overall, this study found that VeRCYe was not very sensitive to the VI used to generate the fused LAI as shown in Table 12. However, overall the results highlight that using the fused-LAI equivalent to the original generic S2-LAI underperformed the field-scale yield estimations, resulting from using the adjusted LAI dataset. This study found that the adjusted RS LAI based on the RDVI resulted in the best yield estimation accuracy with $R^2 = 0.88$ and RMSE = 757 kg/ha (average of -15%) between the reported and estimated yield (Figure 28). In addition, this method was able to estimate both the lowest (under 1,050 kg/ha) and highest yields (above 6,500 kg/ha) with satisfying accuracy, with a RMSE of 178 kg/ha (average error of 1 kg/ha, -1%) and 522 kg/ha (average error of 468 kg/ha, -7%), respectively. Despite these satisfying results, VeRCYe tended to underestimate the reported yield in the tested conditions as shown in Figure 28. When using the 40% of the sessions' highest LAI values, the yield

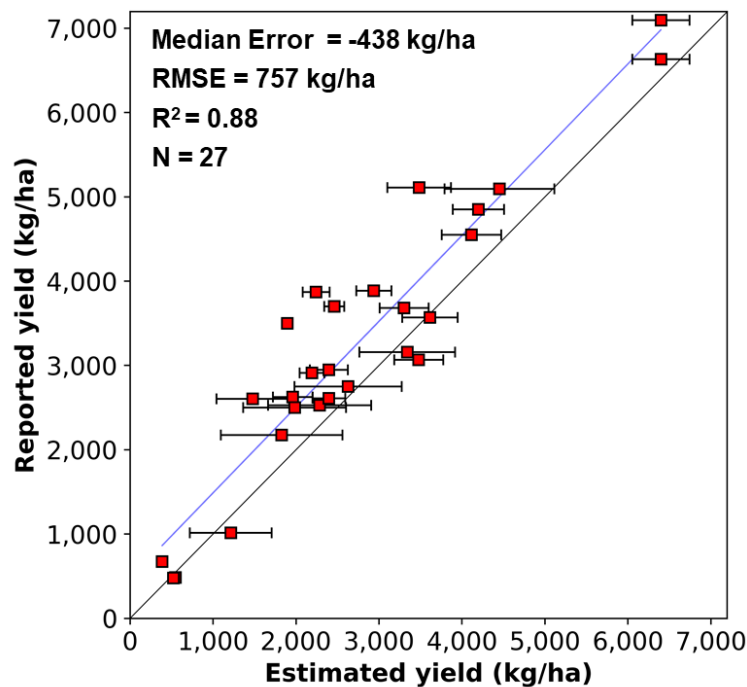


Figure 28. A comparison between wheat yield reported by farmers and estimated yield at the field-scale estimated by VeRCYe, when using the adjusted RS LAI based on the RDVI. In the figure each red square represents of the in 27 fields for which yield was reported by farmers; the whiskers represent the standard deviations of the estimated yield; the black line represents the 1:1 line and the blue line represent the trendline.

estimation was conducted 2 months before the harvest on average, and therefore can be considered as the forecasting period tested in this study (Figure 29).

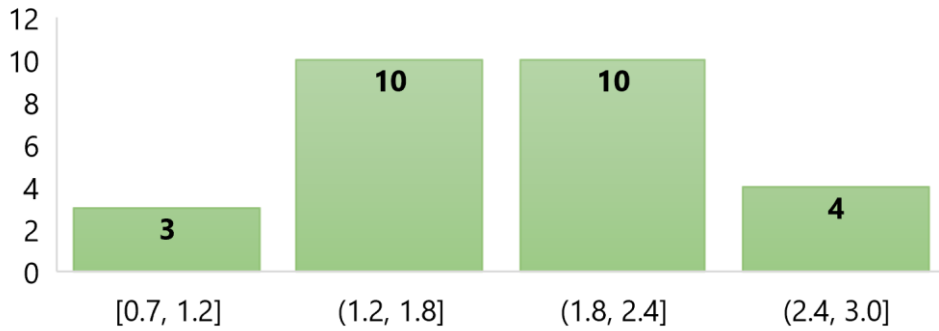


Figure 29. A histogram that how long before the harvest the forecasting conducted. This was determent on the data what the highest 40% of the sessions' LAI was achieved.

4.4.3. Yield map accuracy

The ability of VeRCYe to estimate yield at the pixel level was tested for 22 fields. The results, as shown in Table 13, indicate that when using the fused-LAI which was equivalent to the original generic S2-LAI, the proposed yield estimation method was able to produce estimated yield maps with an RMSE of 1,108 kg/ha, and an average and median error of -467 kg/ha and -534 kg/ha respectively (for the best performing VIs) at the pixel level (for all pixels of all maps). The R² between the yield estimates using this dataset and the reported yield ranged between 0.28 and 0.32 for all VIs tested, while overall the RDVI-based fused LAI slightly outperformed the other VIs. In contrast to the improvement achieved by using the adjusted LAI dataset in

Table 13. The performance of VeRCYe to accurately generate sub-field scale yield estimations by creating 3 m yield maps (n=22). This table shows a comparison of the accuracy of the yield maps which were based either on (i) the fused-based LAI dataset equivalent to the generic S2-LAI (original) or (ii) the adjusted fused-based LAI. This table shows which of the 13 different vegetation indices (VI's) tested by Sadeh et al. (2021) to fused PS and S2 into high spatio-temporal resolution LAI, resulted with the most accurate yield maps estimation. The best performances in each performance metric is coloured in red.

VI	NDVI	EVI2	MTVI2	MSAVI	WDRVI	GREEN WDRVI	GCVI	OSAVI	GSR	GNDVI	RDVI	TVI	SR
Original LAI													
Average error (kg/ha)	-604	-625	-572	-652	-665	-644	-606	-605	-605	-660	-467	-660	-652
Median error (kg/ha)	-627	-607	-677	-681	-690	-646	-600	-627	-601	-654	-534	-683	-672
RMSE (kg/ha)	1108	1133	1215	1145	1184	1165	1148	1109	1147	1156	1199	1146	1183
R²	0.30	0.31	0.28	0.30	0.32	0.31	0.32	0.30	0.32	0.31	0.30	0.30	0.32
Adjusted LAI													
Average error (kg/ha)	-894	-911	-687	-839	-886	-845	-813	-895	-812	-872	-668	-852	-856
Median error (kg/ha)	-999	-999	-855	-845	-966	-944	-926	-999	-927	-966	-819	-851	-976
RMSE (kg/ha)	1288	1303	1281	1258	1285	1235	1214	1289	1213	1250	1299	1261	1272
R²	0.30	0.31	0.27	0.29	0.32	0.31	0.32	0.30	0.32	0.31	0.30	0.30	0.32

estimating field-scale yield, using it to generate yield maps did not result in improved accuracy. Using the adjusted LAI resulted with an RMSE of 1,213 kg/ha, and an average and median error of -668 kg/ha and -819 kg/ha respectively (for the best performing VIs) at the pixel level. The R^2 between the estimated yield maps and the harvesters' yield maps ranged between 0.27 and 0.32 on average for all VIs tested, while overall the RDVI and the GSR-based fused LAI outperformed the other VIs in most of the parameters. It is important to note that in some cases the correlation at the pixel level between the harvester and the estimated yield maps was higher than $R^2 = 0.81$ (RMSE > 525 kg/ha) as shown in Figure 30. As VerCYe uses the peak of the field's LAI (the day with the highest field-scale median LAI of the season) to generate the yield maps, it enabled the creation of yield map estimation at 3 m pixel size on the day the crops reached their sessional peak of LAI, which was observed in this study to be between 2 and 4 months before the harvest. Yield maps produced in this study enabled the estimation of yield at the pixel level on average of 2 months before the reported harvest.

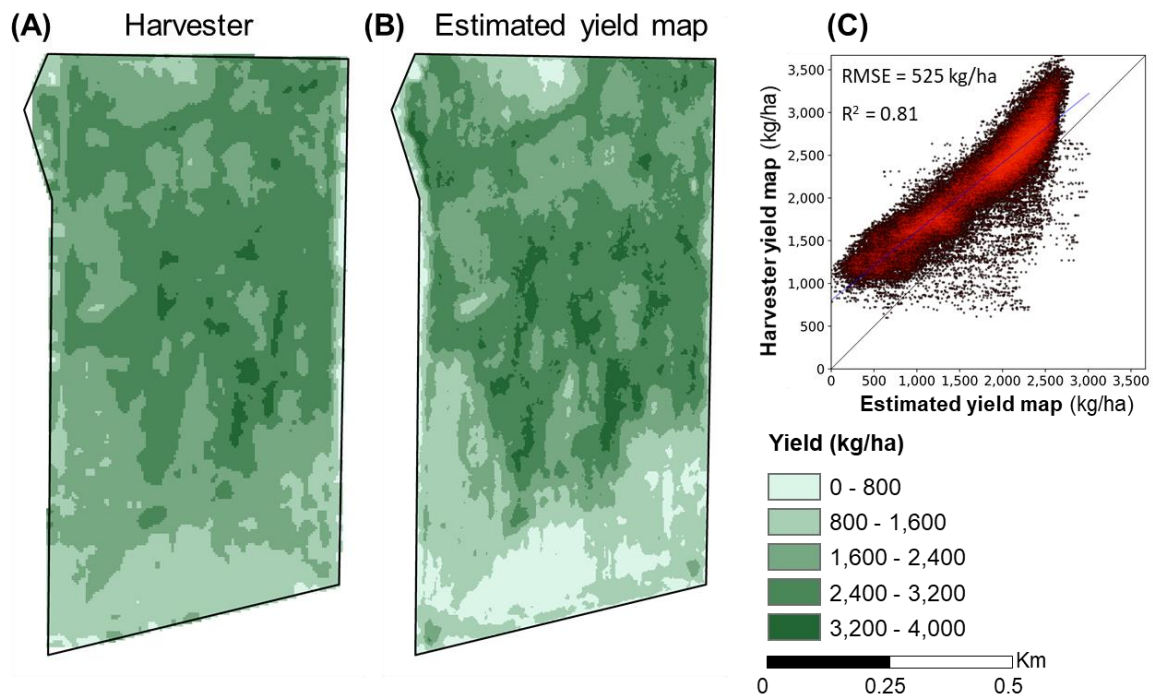


Figure 30. Yield map generated by the harvester (A) and yield map generated three months before harvest using the proposed methodology (B) and their comparison (C) for a wheat field near Birchip, Victoria, Australia. In (B) the estimated 3 m pixel size yield map was generated on the 18.8.2018 which was three months before the harvest (17.11.2018), while the crop was at the Flag Leaf growth stage (Zadoks growth stage 36.68). In (C) the correlation between the two yield maps is presented in the form of a scatterplot, where the black line represents the 1:1 line and the blue line represents the trend line. The correlation analysis between these maps found a RMSE = 525 kg/ha and $R^2 = 0.81$.

4.5. Discussion

This study presented the new VeRCYe method to estimate crop yield at the field and pixel scales which do not rely on detailed official crop statistics or in-situ measurements, by coupling the capabilities of remote sensing and crop models. Many studies have shown a linear relationship between the photosynthetic capacity estimated from spectral responses and the crop phenology, which can be used to predict wheat yields using satellite remote sensing (e.g. Becker-Reshef et al. 2010; Franch et al. 2015; Bognár et al. 2017). However, reliance upon a unique linear relationship is not ideal, especially when crops experience highly variable environmental conditions, as in Australia, where crops are frequently stressed by heat waves, frosts and droughts (Chenu et al. 2013; Zheng et al. 2015; Ababaei and Chenu 2020).

4.5.1. Wheat yield estimation at the field-scale

Despite the extensive research on yield estimation via remote sensing, it is still difficult to directly compare the results of this study with other studies that have attempted to estimate field-scale yield, mainly owing to differences in crop type and the spatial and temporal extent of these estimations (Dado et al. 2020).

In this study, VeRCYe was able to estimate wheat yield on average two months before the harvest with a satisfactory accuracy. The results suggest that using the wheat adjusted LAI (Chapter 3) improves the accuracy of field-scale yield estimations substantially in comparison to using the fused-LAI, which is equivalent to the original generic S2-LAI. Implementing VeRCYe using the adjusted LAI resulted in an RMSE of 757 kg/ha ($R^2 = 0.88$) across the 27 studied fields, while implementation with the original LAI achieved an RMSE of 971 kg/ha ($R^2 = 0.89$). The best performing VI fusion-based LAI for the original LAI was MTVI2 and for the adjusted LAI, RDVI was found to achieve the best accuracy with an RMSE of 757 kg/ha ($R^2 = 0.88$). These two VIs were reported by Sadeh et al. (2021) to be among the best performing VIs to estimate wheat Green-LAI using their proposed fusion method.

VeRCYe was found as a scalable approach for estimating wheat yield without the need for calibration, performing almost as well (and sometimes even better) than approaches that use field data for calibration. For example, Feng et al. (2020) estimated plot-scale wheat yield in the south-eastern Australian wheat belt, using daily MODIS NDVI data (500 m spatial resolution), one month prior to harvest ($R^2 = 0.72$, RMSE = 700 kg/ha), and at two months before harvest ($R^2 = 0.38$, RMSE = 1,101 kg/ha). Donohue et al. (2018) developed a field-scale regional crop yield model called C-Crop, which has been locally calibrated using yield data collected by the farmers' harvesters. They estimated wheat yield for fields across

Australia with an RMSE of 730 kg/ha and mean $R^2 = 0.68$. Chen et al. (2020) developed a semi-empirical model called Crop-SI to estimate the yield of the three major crops in the dryland Australian wheat belt, achieving an RMSE of 540 kg/ha and $R^2 = 0.74$ from the observed field-scale wheat yield. Both the C-Crop and Crop-SI methods used MODIS time-series NDVI at 250 m resolution (provided at 16-day intervals). Filippi et al. (2019) reported RMSE = 360 kg/ha in their study to predict wheat, barley, and canola crop yield for several large farms in Western Australia using only on-farm data and the MODIS 16-day Enhanced Vegetation Index (EVI) at 250 m resolution through machine learning techniques. Zhao et al. (2020) linked S2-based VI time-series with field-scale wheat yield using a linear regression model in Northern New South Wales, Australia. In their study they managed to achieve an RMSE of 640 kg/ha and $R^2 = 0.93$ when validating their models on an independent set of fields.

Cai et al. (2019) compared the performance of one linear regression method and three machine learning models to estimate wheat yield in Australia, at the statistical division level, using the MODIS 16-day EVI dataset and climate data, achieving $R^2 = 0.73$ up to two months before the harvest. However, all of these studies required the collection of an extensive and unique dataset measured in-situ to train or calibrate their models. These kinds of datasets are rare, expensive to obtain and very time consuming to perform. Furthermore, these methods, which require ground calibration data, are typically limited in applicability to the regions from which the *in-situ* data were derived. In addition, most methods provide yield estimations at a low resolution, and often cannot be used for field and pixel scales yield predictions. One of the reasons why many of these studies have estimated yield at the regional scale is the difficulty to predict yields at a smaller scale, owing to the variability of the environmental conditions and farm practices within even the same region (Feng et al. 2020). The development of VeRCYe was motivated to overcome these limitations, with its great advantage being that it uses agro-physiological knowledge embedded in a crop model (APSIM) which can be directly related to crop performance monitored by satellite through space and time. In addition, VeRCYe can theoretically be applied to different crop types across different regions, without the need for local calibration, but also applicable for a rapid changing environment (e.g. Ababaei and Chenu 2020) to which farmers are already adapting (e.g. Flohr et al. 2018).

The SCYM approach, which provides yield estimates at the pixel level by applying regression to satellite images and gridded weather data, had some promising results when first tested in estimating yield at the field-scale for maize ($R^2 = 0.35$) and soybean ($R^2 = 0.32$) in the Midwestern United States (Lobell et al. 2015). However, later studies that tested the method for wheat yield estimation achieved limited success. Jain et al. (2016) mapped yields of smallholder wheat fields in Bihar, India using the SCYM approach achieving $R^2 = 0.27 - 0.33$

and RMSE of 557 – 606 kg/ha for the two growing seasons studied. Shen and Evans (2021) used SCYM for estimating wheat yields over 10 growing seasons between 2003 – 2017 for two nearby fields in Western Australia. Their results showed $R^2 = 0.49$ and RMSE = 620 kg/ha against yield maps created from harvester yield monitoring data. The attempt to use SCYM to estimate village-level wheat yields in Nepal at 10-meter resolution using S2 images by Campolo et al. (2021) resulted with $R^2 = 0.24$ in comparison to crop cuts (in-situ yield samples).

Azzari et al. (2017) compared the performance of SCYM and the PEAKVI method (Becker-Reshef et al. 2010; Franch et al. 2015) in estimating the yield of maize in the United States, wheat in India, and maize in Zambia using Landsat and MODIS observations. Their comparison of these two approaches, which can theoretically be applied anywhere in the world, showed that overall both methods had similar performance in monitoring spatial variability at the county and district scales. In their attempt to estimate wheat yield in India, Azzari et al. (2017) found the performance of SCYM and MODIS-based PEAKVI resulted with $R^2 > 0.45$ in most years, while the best combination was the MODIS-based SCYM, which had an average RMSE of 560 kg/ha. These results support the conclusions of Waldner et al. (2019), which illustrated the importance of the LAI's temporal resolution for accurate yield estimation.

Despite the several attempts to use SCYM for wheat yield estimation with a range of different satellites having different spatial and temporal resolutions (e.g. Jain et al. 2016; Azzari et al. 2017; Jain et al. 2017; Campolo et al. 2021; Shen and Evans 2021), none of these attempts resulted in better performance than the proposed VeRCYe approach as demonstrated in this study. Despite the ability to generate yield estimations at the pixel level with Landsat (30 m and 16-day revisit time) or S2 (10 m and 5-day revisit time) using SCYM, the temporal resolution of these satellites is a significant drawback. While using the MODIS dataset (250 m and daily revisit time) may improve SCYM's performance (Azzari et al. 2017; Waldner et al. 2019), it will not be able to reveal the spatial variability within the field or be used for yield estimation over smallholder farms. The approach presented in this study, which used a daily 3 m LAI dataset, has the advantage of having both high spatial and temporal resolutions, which is one of the reasons it was able to accurately estimate yield at the field and sub field scales.

Around the world millions of people are highly dependent on agriculture for their livelihoods, with low yield harvests directly correlated to high levels of food insecurity (Becker-Reshef et al. 2020). Therefore, when crop conditions are extremely poor, yield prediction methods require being able to flag these failures early. Methods that use ground-based data for calibration and training of their models (as those using machine learning techniques) typically

do not have the required ground reference data to represent the yield heterogeneity of the region of interest over space and time (Benami et al. 2021). Not having training data which reflects extremely low yield scenarios may prevent these methods from producing accurate and reliable yield estimations. Dado et al. (2020) found that SCYM tended to under-predict high soybean yields and overpredict low yields. They concluded that SCYM was unable to differentiate fields which achieved high LAIs and average yields from fields with high LAIs and outstanding yields. Deines et al. (2021) reported that in their study on maize yield estimation using SCYM, 47% of outlier under-estimations occurred in a year with a severe drought (Deines et al 2020). VeRCYe on the other hand managed to estimate such extreme low-yield fields (Figure 28), despite being tested over wheat fields heavily impacted by one of the worst droughts in Australia in the last decade (Tian et al. 2020). This was achieved by identifying a field-scale failure, which was determined as a worst-case-scenario when the field's maximum RS LAI of the season was lower than 0.9. In such a worst-case-scenario the estimated yield was set to be the average of the three simulations with the lowest yield estimation.

Despite its popularity, use of the peak LAI to estimate yield is likely to achieve poor estimations (Waldner et al. 2019). LAI by itself is limited as a linear indicator for the crop's yield as this may be due to failure of plant development, biotic or abiotic stresses (Huang et al. 2019; Beyene et al. 2021). That also applies for the limited linear relationship between the VIs peak and the final yields (Kamir et al. 2020). However, it has been indicated by Dado et al. (2020), that using the peak GCVI and a window of 30 days after the peak allowed a slightly better yield estimation to be achieved than by using the GCVI peak alone.

While most VIs works well when crop is still green, they are less affective during reproductive growth when crop colour starts to change from green to yellow-brown. This limitation also applies to remotely sensed LAI estimated using optical remote sensing (See Chapter 3). Accordingly, the current study highlights the need to also match the Senescence LAI owing to its important role in grain development. Figure 24 also shows that while the simulated LAI peaks resulting from ~2,000 different combinations of possible scenarios may be similar in their timing and magnitude, only during the Senescence LAI could the pattern that better represented the final yield be identified. As optical remote sensing mainly represents the Green LAI (Haboudane et al. 2004), identifying a simulated LAI with identical pattern to the entire remotely-sensed LAI time-series very likely results in an underestimation of the final yield. For that reason, VeRCYe includes a step Step 6) which divides the remotely-sensed LAI series into two, Green LAI and Senescence LAI, and analyses each of them separately.

4.5.1. The potential of generating 3 m yield maps

A final goal of this study was to generate yield maps at the pixel level months before the harvest. Yield maps can help with estimating profitability, assessing the impacts of treatments used, establishing management zones, estimating the amount of nutrients removed by the harvested crop, improving farmers skills, reducing yield gaps and identifying areas which have predominantly large continuous gaps (Lobell et al. 2015; Fulton et al. 2018; Zhao et al. 2020). However, the first step towards reducing yield gaps is to attain accurate estimates of their magnitude, which represents their spatial and temporal variability (Hochman et al. 2012). The 3 m yield maps produced by VeRCYe can help to address these challenges, especially in regions where reliable geolocated yield data obtained from harvesters is not available (such as in many developing countries).

The accuracy of the yield maps generated by VeRCYe resulted in $R^2 = 0.32$ (RMSE of 1,213 kg/ha) using the best performing VI (RDVI). These results are equivalent to the accuracy of other yield mapping methods reported in the literature. For example, Manivasagam et al. (2021) evaluated the assimilation of LAI derived from S2 and LAI derived from fused S2-PS images into Simple Algorithm For Yield estimate (SAFY) to assess the within-field crop yield on spring wheat grown in Israel. Their results showed that the LAI derived from PS-S2 fused images had higher accuracy for yield estimation (RMSE = 690 kg/ha (69 g/m²), $R^2 = 0.45$) than did S2 images alone (RMSE = 880 kg/ha, $R^2 = 0.35$). However, the method used by Manivasagam et al. (2021) requires calibration through field-measured LAI in a few points in time, which is rarely available. Sagan et al. (2021) utilized raw satellite imagery for field-scale soybean ($R^2 = 0.87$) and corn ($R^2 = 0.57$) yield prediction using deep learning with hand-crafted features and WorldView-3 and PS imagery. Yet this method required in-situ yield data for training the model. Kamir et al. (2020) used a large training data set of yield maps obtained from harvesters to predict wheat yields in Australia using machine learning based on climate records and NDVI time series data obtained from MODIS. Their method produced pixel-level yield estimates at 250 m resolution with an R^2 of 0.77 and an RMSE of 550 kg/ha and an R^2 of 0.66 at the level of statistical units.

Dado et al. (2020), for example, used machine learning (random forest) for training their model using a unique ground-truth dataset of soybean yield maps generated from combine harvester yield monitor data across the Midwestern United States. Their harvester-trained model resulted in $R^2 = 0.43$ to 0.32. Dado et al. (2020) also evaluated SCYM performance using the same ground-truth soybean yield maps achieving an $R^2 = 0.27$. Deines et al. (2021) implemented SCYM for estimating maize yield across the United States Corn Belt at 30 m pixels and evaluated their results against harvester-based yield monitor data. First, they used

a baseline SCYM model (Jin et al. 2017b) applied to Landsat data, which resulted in $R^2 = 0.31$ (RMSE = 2,630 kg/ha). Second, Deines et al. (2021) tested an alternative SCYM model that increased the accuracy to $R^2 = 0.4$ (RMSE = 2,450 kg/ha), while Jeffries et al. (2019) tested a variation of SCYM for mapping sub-field maize yields in Nebraska, USA, which resulted in R^2 average value of 0.12 (R^2 ranged 0.003 to 0.37). However, in contrast to these studies which implemented the SCYM approach, which commonly provides yield estimations at 30 m pixel size, the yield maps produced in this current study provided a ten times higher spatial resolution with the same overall accuracy.

4.5.2. Sowing dates as model inputs

Sowing dates are major inputs for crop models, which are commonly used to explore the expected yield effects of different management practices (Zheng et al. 2012; Holzworth et al. 2014; Chenu et al. 2017; Flohr et al. 2017). However, sowing dates are a source of considerable uncertainty for regional studies (Mathison et al. 2017). Accurate regional information about sowing dates at farm scale can be used to reduce the uncertainty of crop simulations (Mathison et al. 2017), however, obtaining accurate sowing dates from ground reports is very difficult and time consuming (Sacks et al. 2010; Marinho et al. 2014). Satellite-based studies on yield estimation traditionally used officially reported sowing dates (e.g. Sakamoto et al. 2005; Marinho et al. 2014; Jin et al. 2016) or sowing dates based on farmers' reporting and surveys (e.g. Ortiz-Monasterio and Lobell 2007; Jain et al. 2016; Manivasagam et al. 2021), while other methods use a sowing date window (e.g. Lobell et al. 2015; Azzari et al. 2017). Deines et al. (2021) found that the three sowing dates used in their study, which represent the 10th, 50th, and 90th percentiles in their study area, failed to capture the full variation present within that region. Therefore, they tested a SCYM implementation with three additional sowing dates in their APSIM simulations, increasing the number of simulations per site-year from 50 to 100.

The approach used in this study was to minimize the uncertainty associated with an unknown important model input, i.e. the field's sowing date, by detecting the actual date the farmer sowed the field and using it as an input to the crop model simulation. Here the sowing dates for the analysed fields were detected using the approach presented in Chapter 2, which resulted in RMSE = 2.7 days and 0.5-day gap for the median between the reported and detected sowing dates.

To evaluate the contribution of using the detected sowing dates as model inputs, this study used a sowing window approach adopted from Waldner et al. (2019). The sowing criteria initiate the crop sowing if rain ≥ 12 mm over 3 days regardless of soil moisture from 26 April to

15 July (this study used ≥ 12 mm instead of ≥ 15 mm originally used by Waldner et al. (2019)). In case the sowing criteria were not met during the sowing window, the crop was automatically sown on the 15th of July. The result of the analysis shows a significant improvement in the accuracy of the yield estimation when using the detected sowing dates as inputs to the model instead of a rule-based sowing window. As shown in Figure 31, using the sowing window with the adjusted fused-based LAI, the R^2 and RMSE between the yield estimates and the reported yield was 0.71 and 1,271 kg/ha respectively. While using the detected sowing dates resulted in $R^2 = 0.88$ and RMSE of 757 kg/ha.

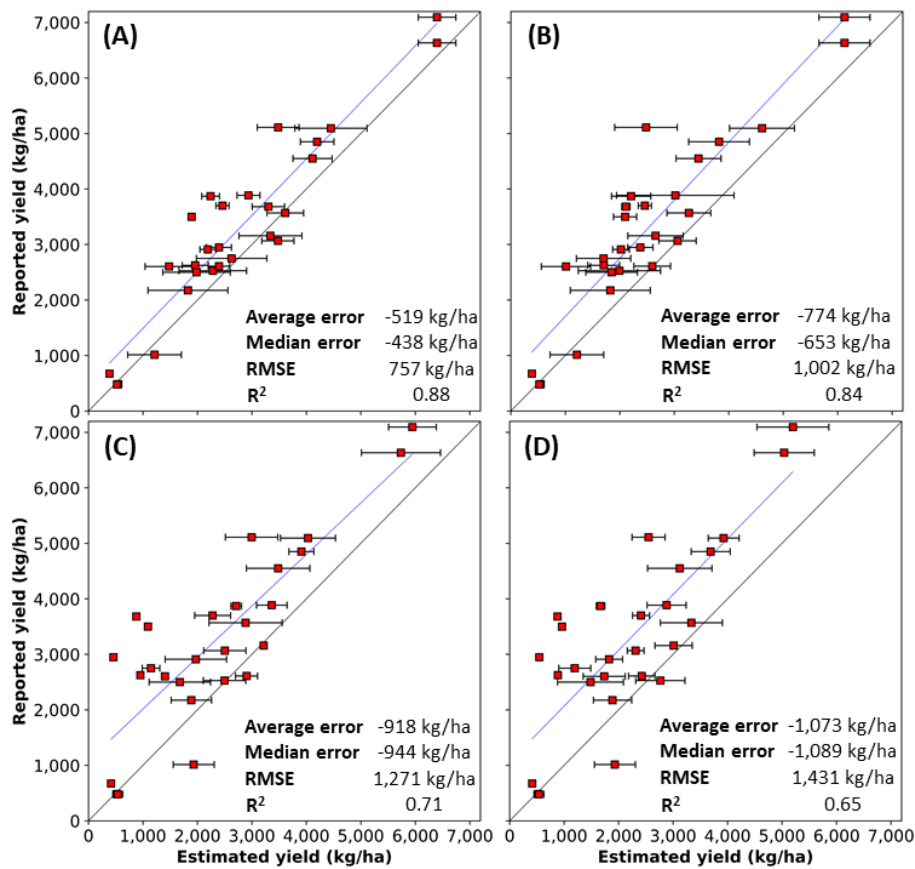


Figure 31. A comparison between accuracy of VeRCYe's yield estimations when using the detected sowing dates for each field and its accuracy when a rule-based sowing window used to determine the fields' sowing dates for APSIM simulations. (A) shows the outcome of VeRCYe using the adjusted fused-based LAI with the detected sowing dates, (B) is VeRCYe's results using the fused-based LAI dataset equivalent to the generic S2-LAI (original) with the detected sowing dates, (C) shows VeRCYe's results using the adjusted fused-based LAI with a sowing window and (D) is VeRCYe's results using the fused-based generic S2-LAI with a sowing window. In the figure each red square represents of the in 27 fields for which yield was reported by farmers; the whiskers represent the standard deviation of the estimated yield; the black line represents the 1:1 line and the blue line represent the trendline.

4.5.3. Limitations and prospects

Even with the promising results presented here, there are some limitations that should be noted.

The assumption that yields and LAI can be accurately simulated by crop growth models, such as APSIM, is the basis of VeRCYe. Despite its international reputation, the APSIM wheat model is not perfect. Brown et al. (2018), who evaluated the model's performance, found the APSIM wheat model to estimate wheat yield with $R^2 = 0.84$ and $RMSE = 100.5$ kg/ha. In addition, there is a less than perfect agreement between APSIM simulated LAI values and the remotely sensed LAI (Waldner et al. 2019), even without the imbedded noise in the satellites data (Sadeh et al. 2021). For example, Ahmed et al. (2016) reported that APSIM tends to slightly overestimate LAI. Despite the decision to use APSIM in this study, VeRCYe could theoretically be implemented using other crop growth models. Future studies should evaluate its performance with other crop models, such as WOFOST (Van Diepen et al. 1989; Ma et al. 2013). It is possible that some models will be more relevant for the crop and environment targeted.

Skakun et al. (2021) assessed within-field corn and soybean yield variability while comparing imagery from several sensors with a range of spatial resolutions. They showed that imagery with spatial resolution of 3 m, such as PS, is critical to explaining the within-field yield variability, while moving to coarser resolution data of 10 m, 20 m, and 30 m reduced the explained variability. As VeRCYe managed to achieve higher accuracy than SCYM in predicting wheat yields (compared to other studies) at the field-level, one could expect that using its pixel-level the accuracy will be better. However, VeRCYe's performance in pixel-level yield estimation was overall equivalent to the pixel-level yield estimations reported in the literature for SCYM (Dado et al. 2020; Deines et al. 2021). For the question what is causing the difference between the accuracy at the field-level V.S. the pixel-level. It is known in the literature that as the yield estimation is aggregated, for example from the field to the farm and county, the error of the estimation declines. This is probably owing to the compensation of the error by the large sample size. The same mechanism works when comparing pixel-level to field-level yield data. However, one thing that can affect this difference here is the use of the Conversion Factor (CF) in its current form. Future studies should test different approaches to improve the accuracy of VeRCYe's yield maps. One way to approach this is instead of using one Conversion Factor (CF) (Equation 7) applied over the whole field, zones with low, average and high LAI values should be treated with modified CF.

Currently VeRCYe uses the Remotely Sensed LAI_{max} (the day when RS LAI was detected as being the maximum during the growing season at the field scale) to convert the LAI values into yield at the pixel level. However, use of LAI by itself is limited as a linear indicator for yield as it is not possible to tell if low LAI was caused by plant development, biotic or abiotic stress (Huang et al. 2019).

A yield estimation at the field-scale or its associated yield map (accurate as they can be) only provides information about yield, and the map itself cannot identify the yield impacting factors. By contrast to most other yield estimation methods, VeRCYe identifies the best representing model simulations out of a couple of thousands of simulations that span a realistic range of possible environments and on-farm management practices. This can theoretically enable the on-farm management practices used in the selected best-fit simulations to be extracted for investigation. For example, when analysing field-scale yields over a specific region, the practices resulting in the highest or lowest yields can be identified and management practices that may help farmers to improve their productivity recommended. Having said that, it is possible that the method produces accurate yield estimations but for the wrong reasons. Having multiple changing parameters may end up with more knobs that can be turned, which can increase the chance of getting right-looking answers from an incorrect set of parameters. Therefore, further research is needed to verify if the optimal APSIM parameters actually reflect the conditions on the ground, which requires a very detailed record of farm management practices used by the farmers. However, such an analysis was beyond the scope of this study.

This study used weather data covering the whole growing season, while in real-time forecasting such data will not be available. In such operational mode, VeRCYe can use the season forecasted weather or the average daily weather from the past (e.g. last 30 years). In addition, the method can provide yield estimations as soon as the duration 40% of the top LAI is finished. Theoretically, this can be reduced even to 5% that will enable earlier estimation, however this is likely to reduce the accuracy of the yield estimations. Future study will explore what will be the best timepoint to estimate yield using the proposed method.

Future studies should also explore different methods (such as machine learning) to select the best fit model simulated LAI to the remotely sensed LAI, and to test the ability of the method to provide accurate field and pixel scale yield forecasting prior to the harvest using projected weather data. Although the proposed method was tested to estimate wheat yield, it is very likely that it will also be useful to estimate the yield of other crop types as well. However, the results of this study indicate that adjustment of generic S2-LAI data to better estimate the wheat Green LAI as proposed in Chapter 3 (Sadeh et al. 2021), is needed in order to achieve

better yield estimation. Moreover, it is likely that such an adjustment will also be needed when implementing VeRCYe with other crop types.

While this method relies on the availability of LAI data from optical sensors, these are highly sensitive to the presence of clouds and shadows in the imagery, which is likely to limit the ability to perform at its best over certain regions. In contrast to optical sensing, SAR sensors have the advantages of all-weather capabilities, and therefore a potential improvement to VeRCYe should include the use of SAR-based LAI or SAR-optical fused LAI.

The advantage of being able to estimate field and farm productivity remotely without the need of having “boots on the ground” has been magnified by the outbreak of COVID-19. While lockdowns and prolonged COVID-19 quarantine measures delay/limit supply of essential products such as fertilizers, herbicides, machinery or even the availability of seasonal workers, affecting the farmers performance, it has also decreased feed wheat and wheat-based product demand (FAO 2021). VeRCYe can potentially help to monitor these influences remotely across different regions without the need to risk surveyors in collecting ground data.

4.6. Conclusions

The VeRCYe method proposed in this study overcame the limitation with previous studies relying on ground data to estimate wheat yield by combining the advantages of both high spatio-temporal remote sensing and crop model simulations. This not only enables model inputs in the form of sowing dates to be detected from space, but also overcame the historical trade-off between high spatial and temporal resolutions for remotely sensed estimation of crop yield at the field and pixel scales. As any requirement for ground calibration data typically limits model applicability to the regions from which the in-situ data were collected, this new approach which does not rely on such data is needed. Accordingly, VeRCYe does not require any crop statistics or in-situ measurements, making it broadly applicable across regions, including where ground calibration data are not available. This method was found to be effective for not only producing field-scale yield estimations, but also to generate yield maps at 3 m resolution up to four months before crop harvest. Despite being tested for wheat, VeRCYe can potentially be used to estimate the yield of other crop types as well, with minimum adaptation only. This study outlines an innovative approach to monitor the farmer’s management practices at 3 m, from sowing through monitoring the crops performance throughout the season until the farmer decides to harvest and the yield becomes available to be trayed as a food product. Furthermore, the information generated using this method can be used to understand yield

variability from a regional scale to the pixel scale and may provide insights on the causes and their spatial distribution.

Chapter 5 – Conclusions and future directions

While most of the common satellite-based yield estimation methods rely extensively on in-situ data for model training and calibration, they typically provide a local solution for the area where they have been calibrated. Accordingly, this thesis has presented an innovative, scalable and flexible yield estimation method named VeRsatile Crop Yield Estimator (VeRCYe) to overcome these issues. VeRCYe combines the power of both high resolution spaceborne remote sensing and crop modelling, to predict wheat yield at the field and pixel scales using remote sensing without using ground-based data.

This chapter briefly summarizes the main conclusions of the thesis. It also includes some of the future directions which forthcoming studies should explore. In summary, this thesis consists of three components (i) sowing date detection, (ii) creation of daily 3 m LAI data and (iii) yield estimation at field and pixel scales. These components are combined to achieve the main goal of creating a robust method for crop yield estimation with global applicability from the pixel level to regional scale for different types of broadacre crops. Each component of this study has a range of potential applications, as well as limitations.

5.1. Sowing date detection

The method to detect sowing dates, as outlined in Chapter 2, may be used to produce and update near-daily low-cost field level statistics with a large spatial coverage (with an RMSE of 1.9 days). The method was also proven effective for detecting harvested areas at the field and sub-field scales with very high accuracy (RMSE of 2.6 days). Moreover, it was shown that sowing dates detected using this method could be used as a model input to estimate yield.

Currently the sowing detection method is still semi-automated, as users need to identify the pixel values that correspond to a change in the image histogram, in order to determine the threshold to classify pixels as 'changed' or 'not changed' (Figure 5). In order to fully automate the process this threshold needs to be identified automatically. This potentially can be achieved by implementing machine-learning techniques to identify the pixels that represent change in the histogram. Furthermore, while the no-tillage sowing practice is now common worldwide for rainfed crops, this method is currently limited to regions where no-tillage sowing is being implemented.

Perhaps the most important question of this contribution in the context of the main objective of this thesis is, how much if any improvement can be achieved in yield estimation using the

detected sowing dates as a model input? The answer for that was supposed to come through sensitivity analysis, which although planned, were not able to be completed within the constraints of completing this thesis. Therefore, future studies should compare the effect of using VerCYe with either the detected dates (as performed in this study) or a sowing window. Moreover, additional sensitivity tests should be performed to determine the maximum error (expressed as RMSE) in days, from which using detected sowing dates will no longer minimise the uncertainty of the estimated yield.

5.2. Data fusion into high-resolution LAI

The fusion methodology outlined in Chapter 3 helped to resolve the challenges posed by inconsistency in PlanetScope (PS) signals and poor spatio-temporal resolution of Sentinel-2 (S2), producing a time-series of LAI images that preserved both the high spatial and temporal resolution of PS and the spectral quality of S2. In practical terms, S2 consistent surface reflectance RGB-NIR images and crop LAI were generated at 3 m resolution on a daily basis with an RMSE = 0.35-0.63 and R^2 of 0.92. These high spatio-temporal resolution time-series are valuable for monitoring crop growth and health, and can improve the effectiveness of farming practices. While using the fused high-resolution LAI enhanced the yield estimation using VerCYe at the field and pixel scales, the S2-based Green LAI estimates tended to underestimate high Green LAI values ($LAI > \sim 3$), similarly to other remotely sensed LAI estimations. These underestimations are likely due to asymptotic saturation of the surface reflectance data caused by the high biomass density, which cannot be assessed fully without 3D information on canopy structure. As such data cannot be obtained from space, this thesis introduced an adjustment that successfully overcame the underestimation of high Green LAI values.

It is proposed that generic LAI products, such as the S2-LAI product, should be adjusted to the crop of interest in order to achieve better yield estimation. As a prospect for future improvements and research directions, it is suggested that future studies should test the proposed fusion method over other crop types and evaluate the LAI adjustment in other crops. Furthermore, future studies should explore the possibility of adding more sensors in the fusion process (e.g. Landsat, SPOT or SkySat) and examine the suitability of this method to fuse sensor data other than S2 and PS. An interesting research direction would be to explore how feasible it will be to fuse satellite imagery with imagery acquired from drones. Succeeding in this would remove the need for costly and time-consuming aircraft flights, which would be very beneficial for many applications. In addition, only four spectral bands were fused here in the RGB-NIR range. Suitability of the method for improving the spatial and temporal resolution of

data obtained from sensors operated in the shortwave infrared (SWIR) and the thermal infrared (TIR) wavelengths should also be explored.

5.3. VeRCYe

VeRCYe is an innovative approach that leverages the power of high-resolution remote sensing and crop models to estimate and map crop yield at the field ($R^2 = 0.88$, RMSE of 757 kg/ha) and pixel ($R^2 = 0.32$, RMSE of 1,213 kg/ha) scales. Compared with the results from other studies, it is found that VeRCYe performed as well or better, with the added advantage that it does not need the *in-situ* calibration or training data required by other methods.

The promising results of VeRCYe in this study indicate that it can potentially be implemented globally, so long as the crop type, the field's location and its boundaries are provided, as well as basic knowledge of the common local farm management practices. Using the fused S2-PS LAI dataset as input to VeRCYe enabled yield estimation over very small fields (<0.1 ha), which could be ideal for estimating crop yield of smallholders in developing countries. However, if the focus is on yield forecasting over developed countries such as Australia and the USA, where the size of the fields is typically very large, future studies should explore using adjusted S2-LAI data alone as the remotely-sensed LAI input to VeRCYe. However, owing to its temporal resolution, such dataset is likely to achieve limited accuracy. According to previous studies, it is likely that the high temporal resolution LAI data used in this study had the most substantial influence on VeRCYe's accuracy and not its spatial resolution.

As the chances for having cloud-free images every five days using the S2-LAI data alone is unlikely, a future study should explore the possibility of using Synthetic-Aperture Radar (SAR) as a means of estimating crop LAI, either solely using SAR imagery or by fusing it with optical sensors. For example, fusing SAR data from Sentinel-1 (S1) with optical data from S2, or by training S1 to estimate S2-LAI using machine learning techniques. Furthermore, it is reasonable to assume that owing to the sensitivity of SAR to the geometry and roughness of the surface will be more suitable for estimating the LAI of the senescent leaves and its accuracy would not be limited only for actively photosynthesising leaves. In addition, using SAR-based LAI will also solve the problem of having clouds in the optical image, which highly affects the estimated LAI values.

As VeRCYe was designed to identify the simulations in which their simulated LAI best matched the remotely sensed LAI, this enables farm management practices used as inputs in these selected simulations to be identified. Future work should examine whether the modelled practices actually correspond to what was done on the ground. An interesting future direction of study would also be looking into these practices and analysing them to identify which

practices ended up with the highest yield in certain regions. This can potentially be used to close the yield gap in areas where farm management practices and crop statistics are unavailable, such as in developing countries. However, it is possible that the method produces accurate yield estimations for the wrong reasons. Another aspect is validating the representation of the model parameters against the actual practices used by the farmer, which will be very challenging as many farmers do not keep records of such data (e.g. dates and amount of fertilizer applied, sowing and harvest dates, sowing depth, cultivar used etc.).

Another potential feature of VeRCYe which should be further explored is the possibility of using this method to map important agronomic properties other than crop yield. At the base of VeRCYe's yield map production lays the assumption that LAI is highly correlated with yield. However, APSIM can also simulate and plot other crop parameters such as the biomass, nutrient conditions and plant available water as well as the soil's nutrient status (e.g. carbon, NH_4 and NO_3). The method is currently designed to use field-scale yield to calculate a Conversion Factor, which is used to convert the LAI map to a yield map at the pixel level. Accordingly, the spatial distribution of these other parameters of interest can also be estimated so long as they are highly correlated with the crop LAI and the model is capable of accurately simulating them. Having the ability to map these parameters at such a high resolution without the need for expensive and time-consuming ground samples, may open the door for a wide range of applications.

To conclude, this study illustrated that the method proposed in this thesis is capable of (1) identifying when a field was sown, (2) monitoring the crops performance and health daily at a 3 m resolution in form of VIs and LAI, (3) estimating its field-scale yield, (4) producing a yield map of the field at the pixel level months before the harvest and finally, (5) detecting when the field is harvested. Therefore, the methods developed during this study have many different applications as well as challenges that need to be addressed towards an operational use in the near future.

References

- Ababaei, B., & Chenu, K. (2020). Heat shocks increasingly impede grain filling but have little effect on grain setting across the Australian wheatbelt. *Agricultural and Forest Meteorology*, 284, 107889.
- Abdi, H., & Williams, L.J. (2010). Principal component analysis. *Wiley interdisciplinary reviews: computational statistics*, 2(4), 433-459.
- Ahmed, M., Akram, M.N., Asim, M., Aslam, M., Hassan, F.-u., Higgins, S., et al. (2016). Calibration and validation of APSIM-Wheat and CERES-Wheat for spring wheat under rainfed conditions: Models evaluation and application. *Computers and Electronics in Agriculture*, 123, 384-401.
- Al-Amri, S.S., Kalyankar, N.V., & Khamitkar, S.D. (2010). A comparative study of removal noise from remote sensing image. *arXiv preprint arXiv:1002.1148*
- Australian Bureau of Meteorology. 2020. "Australian Bureau of Meteorology." Bureau of Meteorology, Accessed 12/8/2020. <http://www.bom.gov.au/climate/data/?ref=fr>.
- Australian Bureau of Statistics. 2021. "Agricultural Commodities, Australia, 2018-19." Australian Bureau of Statistics (ABS), Accessed 7.2.2021. <https://www.abs.gov.au/statistics/industry/agriculture/agricultural-commodities-australia/latest-release#broadacre-crops>.
- Azzari, G., Jain, M., & Lobell, D.B. (2017). Towards fine resolution global maps of crop yields: Testing multiple methods and satellites in three countries. *Remote Sensing of Environment*, 202, 129-141.
- Azzari, G., & Lobell, D.B. (2017). Landsat-based classification in the cloud: An opportunity for a paradigm shift in land cover monitoring. *Remote Sensing of Environment*, 202, 64-74.
- Baker, C.J., & Saxton, K.E. (2006). The 'what' and 'why' of no-tillage farming. *No-Tillage Seeding in Conservation Agriculture: Second Edition* (pp. 1-10)
- Bartier, P.M., & Keller, C.P. (1996). Multivariate interpolation to incorporate thematic surface data using inverse distance weighting (IDW). *Computers & Geosciences*, 22(7), 795-799.
- Battude, M., Al Bitar, A., Morin, D., Cros, J., Huc, M., Sicre, C.M., et al. (2016). Estimating maize biomass and yield over large areas using high spatial and temporal resolution Sentinel-2 like remote sensing data. *Remote Sensing of Environment*, 184, 668-681.
- Becker-Reshef, I., Justice, C., Barker, B., Humber, M., Rembold, F., Bonifacio, R., et al. (2020). Strengthening agricultural decisions in countries at risk of food insecurity: The GEOGLAM Crop Monitor for Early Warning. *Remote Sensing of Environment*, 237, 111553.
- Becker-Reshef, I., Vermote, E., Lindeman, M., & Justice, C. (2010). A generalized regression-based model for forecasting winter wheat yields in Kansas and Ukraine using MODIS data. *Remote Sensing of Environment*, 114(6), 1312-1323.
- Bégué, A., Arvor, D., Bellon, B., Betbeder, J., de Aballeyra, D., Ferraz, R.P.D., et al. (2018). Remote sensing and cropping practices: A review. *Remote Sensing*, 10(1)
- Benami, E., Jin, Z., Carter, M.R., Ghosh, A., Hijmans, R.J., Hobbs, A., et al. (2021). Uniting remote sensing, crop modelling and economics for agricultural risk management. *Nature Reviews Earth & Environment*, 2(2), 140-159.
- Beyene, A.N., Zeng, H., Wu, B., Zhu, L., Gebremicael, T.G., Zhang, M., et al. (2021). Coupling remote sensing and crop growth model to estimate national wheat yield in Ethiopia. *Big Earth Data*, 1-18.

- Bøgh, E., Thorsen, M., Butts, M., Hansen, S., Christiansen, J., Abrahamsen, P., et al. (2004). Incorporating remote sensing data in physically based distributed agro-hydrological modelling. *Journal of Hydrology*, 287(1-4), 279-299.
- Bognár, P., Kern, A., Pásztor, S., Lichtenberger, J., Koronczay, D., & Ferencz, C. (2017). Yield estimation and forecasting for winter wheat in Hungary using time series of MODIS data. *International Journal of Remote Sensing*, 38(11), 3394-3414.
- Brown, H., Huth, N., & Holzworth, D. (2018). Crop model improvement in APSIM: using wheat as a case study. *European Journal of Agronomy*, 100, 141-150.
- Bruzzone, L., & Prieto, D.F. (2000). Automatic analysis of the difference image for unsupervised change detection. *IEEE Transactions on Geoscience and Remote Sensing*, 38(3), 1171-1182.
- Bsaibes, A., Courault, D., Baret, F., Weiss, M., Olioso, A., Jacob, F., et al. (2009). Albedo and LAI estimates from FORMOSAT-2 data for crop monitoring. *Remote Sensing of Environment*, 113(4), 716-729.
- Burke, M., & Lobell, D.B. (2017). Satellite-based assessment of yield variation and its determinants in smallholder African systems. *Proceedings of the National Academy of Sciences*, 114(9), 2189-2194.
- Byrne, G.F., Crapper, P.F., & Mayo, K.K. (1980). Monitoring land-cover change by principal component analysis of multitemporal landsat data. *Remote Sensing of Environment*, 10(3), 175-184.
- Cai, Y., Guan, K., Lobell, D., Potgieter, A.B., Wang, S., Peng, J., et al. (2019). Integrating satellite and climate data to predict wheat yield in Australia using machine learning approaches. *Agricultural and Forest Meteorology*, 274, 144-159.
- Campolo, J., Güereña, D., Maharjan, S., & Lobell, D.B. (2021). Evaluation of soil-dependent crop yield outcomes in Nepal using ground and satellite-based approaches. *Field Crops Research*, 260, 107987.
- Celik, T. (2009). Unsupervised change detection in satellite images using principal component analysis and k-means clustering. *IEEE Geoscience and Remote Sensing Letters*, 6(4), 772-776.
- Chen, J.M., Pavlic, G., Brown, L., Cihlar, J., Leblanc, S.G., White, H.P., et al. (2002). Derivation and validation of Canada-wide coarse-resolution leaf area index maps using high-resolution satellite imagery and ground measurements. *Remote Sensing of Environment*, 80(1), 165-184.
- Chen, Y., Donohue, R.J., McVicar, T.R., Waldner, F., Mata, G., Ota, N., et al. (2020). Nationwide crop yield estimation based on photosynthesis and meteorological stress indices. *Agricultural and Forest Meteorology*, 284, 107872.
- Chenu, K., Cooper, M., Hammer, G., Mathews, K.L., Dreccer, M., & Chapman, S.C. (2011). Environment characterization as an aid to wheat improvement: interpreting genotype–environment interactions by modelling water-deficit patterns in North-Eastern Australia. *Journal of experimental botany*, 62(6), 1743-1755.
- Chenu, K., Dehifard, R., & Chapman, S.C. (2013). Large-scale characterization of drought pattern: a continent-wide modelling approach applied to the Australian wheatbelt—spatial and temporal trends. *New Phytologist*, 198(3), 801-820.
- Chenu, K., Porter, J.R., Martre, P., Basso, B., Chapman, S.C., Ewert, F., et al. (2017). Contribution of Crop Models to Adaptation in Wheat. *Trends in Plant Science*, 22(6), 472-490.
- Clevers, J.G.P.W. (1991). Application of the WDVl in estimating LAI at the generative stage of barley. *ISPRS Journal of Photogrammetry and Remote Sensing*, 46(1), 37-47.
- Clevers, J.G.P.W., Kooistra, L., & van den Brande, M.M.M. (2017). Using Sentinel-2 data for retrieving LAI and leaf and canopy chlorophyll content of a potato crop. *Remote Sensing*, 9(5)

- Coventry, D., Reeves, T., Brooke, H., & Cann, D. (1993). Influence of genotype, sowing date, and seeding rate on wheat development and yield. *Australian Journal of Experimental Agriculture*, 33(6), 751-757.
- Dado, W.T., Deines, J.M., Patel, R., Liang, S.-Z., & Lobell, D.B. (2020). High-Resolution Soybean Yield Mapping Across the US Midwest Using Subfield Harvester Data. *Remote Sensing*, 12(21), 3471.
- Dalglish, N., Hochman, Z., Huth, N., & Holzworth, D. (2016). Field Protocol to APSoil characterisations. In. Australia: CSIRO.
- Dash, J., & Ogutu, B.O. (2016). Recent advances in space-borne optical remote sensing systems for monitoring global terrestrial ecosystems. *Progress in Physical Geography*, 40(2), 322-351.
- Daughtry, C., Gallo, K., Goward, S., Prince, S., & Kustas, W. (1992). Spectral estimates of absorbed radiation and phytomass production in corn and soybean canopies. *Remote Sensing of Environment*, 39(2), 141-152.
- Daughtry, C.S.T. (2001). Discriminating crop residues from soil by shortwave infrared reflectance. *Agronomy Journal*, 93(1), 125-131.
- Deines, J.M., Patel, R., Liang, S.-Z., Dado, W., & Lobell, D.B. (2021). A million kernels of truth: insights into scalable satellite maize yield mapping and yield gap analysis from an extensive ground dataset in the US Corn Belt. *Remote Sensing of Environment*, 253, 112174.
- Delegido, J., Verrelst, J., Rivera, J.P., Ruiz-Verdú, A., & Moreno, J. (2015). Brown and green LAI mapping through spectral indices. *International Journal of Applied Earth Observation and Geoinformation*, 35, 350-358.
- Deng, J.S., Wang, K., Deng, Y.H., & Qi, G.J. (2008). PCA-based land-use change detection and analysis using multitemporal and multisensor satellite data. *International Journal of Remote Sensing*, 29(16), 4823-4838.
- Derpsch, R., Friedrich, T., Kassam, A., & Li, H. (2010). Current status of adoption of no-till farming in the world and some of its main benefits. *International Journal of Agricultural and Biological Engineering*, 3(1), 1-25.
- Dhakar, R., Sehgal, V.K., Chakraborty, D., Sahoo, R.N., & Mukherjee, J. (2019). Field scale wheat LAI retrieval from multispectral Sentinel 2A-MSI and Landsat 8-OLI imagery: effect of atmospheric correction, image resolutions and inversion techniques. *Geocarto International*, 1-21.
- Djamai, N., & Fernandes, R. (2018). Comparison of SNAP-derived Sentinel-2A L2A product to ESA product over Europe. *Remote Sensing*, 10(6), 926.
- Djamai, N., Fernandes, R., Weiss, M., McNairn, H., & Goïta, K. (2019). Validation of the Sentinel Simplified Level 2 Product Prototype Processor (SL2P) for mapping cropland biophysical variables using Sentinel-2/MSI and Landsat-8/OLI data. *Remote Sensing of Environment*, 225, 416-430.
- Dong, T., Liu, J., Shang, J., Qian, B., Ma, B., Kovacs, J.M., et al. (2019). Assessment of red-edge vegetation indices for crop leaf area index estimation. *Remote Sensing of Environment*, 222, 133-143.
- Donohue, R.J., Lawes, R.A., Mata, G., Gobbett, D., & Ouzman, J. (2018). Towards a national, remote-sensing-based model for predicting field-scale crop yield. *Field Crops Research*, 227, 79-90.
- Dronova, I., Gong, P., Wang, L., & Zhong, L. (2015). Mapping dynamic cover types in a large seasonally flooded wetland using extended principal component analysis and object-based classification. *Remote Sensing of Environment*, 158, 193-206.

-
- Drusch, M., Del Bello, U., Carlier, S., Colin, O., Fernandez, V., Gascon, F., et al. (2012). Sentinel-2: ESA's optical high-resolution mission for GMES operational services. *Remote Sensing of Environment*, 120, 25-36.
- Du, Q., & Fowler, J.E. (2007). Hyperspectral image compression using JPEG2000 and principal component analysis. *IEEE Geoscience and Remote Sensing Letters*, 4(2), 201-205.
- Duchemin, B., Fieuzal, R., Rivera, M.A., Ezzahar, J., Jarlan, L., Rodriguez, J.C., et al. (2015). Impact of sowing date on yield and water use efficiency of wheat analyzed through spatial modeling and FORMOSAT-2 images. *Remote Sensing*, 7(5), 5951-5979.
- El Hajj, M., Baghdadi, N., Cheviron, B., Belaud, G., & Zribi, M. (2016). Integration of remote sensing derived parameters in crop models: Application to the PILOTE model for hay production. *Agricultural Water Management*, 176, 67-79.
- Fang, H., Baret, F., Plummer, S., & Schaepman-Strub, G. (2019). An overview of global leaf area index (LAI): Methods, products, validation, and applications. *Reviews of Geophysics*, 57(3), 739-799.
- FAO (2021). Monthly news report on grains - June 2021. In, *FAO Markets and Trade Division: FAO Markets and Trade Division*.
- Feng, P., Wang, B., Li Liu, D., Waters, C., Xiao, D., Shi, L., et al. (2020). Dynamic wheat yield forecasts are improved by a hybrid approach using a biophysical model and machine learning technique. *Agricultural and Forest Meteorology*, 285, 107922.
- Ferencz, C., Bogнар, P., Lichtenberger, J., Hamar, D., Tarcsai, G., Timár, G., et al. (2004). Crop yield estimation by satellite remote sensing. *International Journal of Remote Sensing*, 25(20), 4113-4149.
- Filippi, P., Jones, E.J., Wimalathunge, N.S., Somarathna, P.D., Pozza, L.E., Ugbaje, S.U., et al. (2019). An approach to forecast grain crop yield using multi-layered, multi-farm data sets and machine learning. *Precision Agriculture*, 20(5), 1015-1029.
- Flohr, B.M., Hunt, J.R., Kirkegaard, J.A., & Evans, J.R. (2017). Water and temperature stress define the optimal flowering period for wheat in south-eastern Australia. *Field Crops Research*, 209, 108-119.
- Flohr, B.M., Hunt, J.R., Kirkegaard, J.A., Evans, J.R., Trevaskis, B., Zwart, A., et al. (2018). Fast winter wheat phenology can stabilise flowering date and maximise grain yield in semi-arid Mediterranean and temperate environments. *Field Crops Research*, 223, 12-25.
- Franch, B., Vermote, E., Roger, J.-C., Murphy, E., Becker-Reshef, I., Justice, C., et al. (2017). A 30+ Year AVHRR Land Surface Reflectance Climate Data Record and Its Application to Wheat Yield Monitoring. *Remote Sensing*, 9(3)
- Franch, B., Vermote, E.F., Becker-Reshef, I., Claverie, M., Huang, J., Zhang, J., et al. (2015). Improving the timeliness of winter wheat production forecast in the United States of America, Ukraine and China using MODIS data and NCAR Growing Degree Day information. *Remote Sensing of Environment*, 161, 131-148.
- Frantz, D., Röder, A., Udelhoven, T., & Schmidt, M. (2015). Enhancing the Detectability of Clouds and Their Shadows in Multitemporal Dryland Landsat Imagery: Extending Fmask. *IEEE Geoscience and Remote Sensing Letters*, 12(6), 1242-1246.
- Fulton, J., Hawkins, E., Taylor, R., & Franzen, A. (2018). Yield monitoring and mapping. *Precision agriculture basics*, 63-77.
- Gallego, F.J., Kussul, N., Skakun, S., Kravchenko, O., Shelestov, A., & Kussul, O. (2014). Efficiency assessment of using satellite data for crop area estimation in Ukraine. *International Journal of Applied Earth Observation and Geoinformation*, 29, 22-30.
-

- Gao, F., Anderson, M.C., Kustas, W.P., & Wang, Y. (2012). Simple method for retrieving leaf area index from Landsat using MODIS leaf area index products as reference. *Journal of Applied Remote Sensing*, 6(1), 063554.
- Gao, F., Anderson, M.C., Zhang, X., Yang, Z., Alfieri, J.G., Kustas, W.P., et al. (2017). Toward mapping crop progress at field scales through fusion of Landsat and MODIS imagery. *Remote Sensing of Environment*, 188, 9-25.
- Gao, F., Masek, J., Schwaller, M., & Hall, F. (2006). On the blending of the Landsat and MODIS surface reflectance: Predicting daily Landsat surface reflectance. *IEEE Transactions on Geoscience and Remote Sensing*, 44(8), 2207-2218.
- Gašparović, M., & Jogun, T. (2018). The effect of fusing Sentinel-2 bands on land-cover classification. *International Journal of Remote Sensing*, 39(3), 822-841.
- Gašparović, M., Medak, D., Pilaš, I., Jurjević, L., & Balenović, I. (2018). Fusion of Sentinel-2 and PlanetScope Imagery for Vegetation Detection and Monitorin. In, *Volumes ISPRS TC I Mid-term Symposium Innovative Sensing-From Sensors to Methods and Applications*
- Gil-Yepes, J.L., Ruiz, L.A., Recio, J.A., Balaguer-Beser, Á., & Hermosilla, T. (2016). Description and validation of a new set of object-based temporal geostatistical features for land-use/land-cover change detection. *ISPRS Journal of Photogrammetry and Remote Sensing*, 121, 77-91.
- Gitelson, A., & Merzlyak, M.N. (1994). Spectral reflectance changes associated with autumn senescence of *Aesculus hippocastanum* L. and *Acer platanoides* L. leaves. Spectral features and relation to chlorophyll estimation. *Journal of plant physiology*, 143(3), 286-292.
- Gitelson, A.A. (2004). Wide dynamic range vegetation index for remote quantification of biophysical characteristics of vegetation. *Journal of plant physiology*, 161(2), 165-173.
- Gitelson, A.A., Viña, A., Arkebauer, T.J., Rundquist, D.C., Keydan, G., & Leavitt, B. (2003). Remote estimation of leaf area index and green leaf biomass in maize canopies. *Geophysical Research Letters*, 30(5), n/a-n/a.
- Gitelson, A.A., Vina, A., Ciganda, V., Rundquist, D.C., & Arkebauer, T.J. (2005). Remote estimation of canopy chlorophyll content in crops. *Geophysical Research Letters*, 32(8)
- Gonçalves, M.L., Netto, M.L.A., Costa, J.A.F., & Zullo Junior, J. (2008). An unsupervised method of classifying remotely sensed images using Kohonen self-organizing maps and agglomerative hierarchical clustering methods. *International Journal of Remote Sensing*, 29(11), 3171-3207.
- Guo, Z. (2013). Mapping the planting dates: An effort to retrieve crop phenology information from MODIS NDVI time series in Africa. In, *International Geoscience and Remote Sensing Symposium (IGARSS)* (pp. 3281-3284)
- Haas, R., Deering, D., Rouse Jr, J., & Schell, J. (1975). Monitoring vegetation conditions from LANDSAT for use in range management. In, *NASA Earth Resources Survey Symposium* (pp. 43-52). NASA. Lyndon B. Johnson Space Center: United States.
- Haboudane, D., Miller, J.R., Pattey, E., Zarco-Tejada, P.J., & Strachan, I.B. (2004). Hyperspectral vegetation indices and novel algorithms for predicting green LAI of crop canopies: Modeling and validation in the context of precision agriculture. *Remote Sensing of Environment*, 90(3), 337-352.
- Hadria, R., Duchemin, B., Baup, F., Le Toan, T., Bouvet, A., Dedieu, G., et al. (2009). Combined use of optical and radar satellite data for the detection of tillage and irrigation operations: Case study in Central Morocco. *Agricultural Water Management*, 96(7), 1120-1127.
- Hammer, G., Hansen, J., Phillips, J., Mjelde, J., Hill, H., Love, A., et al. (2001). Advances in application of climate prediction in agriculture. *Agricultural Systems*, 70(2), 515-553.

- Herrmann, I., Pimstein, A., Karnieli, A., Cohen, Y., Alchanatis, V., & Bonfil, D.J. (2011). LAI assessment of wheat and potato crops by VEN μ S and Sentinel-2 bands. *Remote Sensing of Environment*, 115(8), 2141-2151.
- Hobbs, P.R., Sayre, K., & Gupta, R. (2008). The role of conservation agriculture in sustainable agriculture. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 363(1491), 543-555.
- Hochman, Z., Gobbett, D., Holzworth, D., McClelland, T., van Rees, H., Marinoni, O., et al. (2012). Quantifying yield gaps in rainfed cropping systems: A case study of wheat in Australia. *Field Crops Research*, 136, 85-96.
- Holzworth, D., Huth, N.I., Fainges, J., Brown, H., Zurcher, E., Cichota, R., et al. (2018). APSIM Next Generation: Overcoming challenges in modernising a farming systems model. *Environmental Modelling & Software*, 103, 43-51.
- Holzworth, D.P., Huth, N.I., deVoil, P.G., Zurcher, E.J., Herrmann, N.I., McLean, G., et al. (2014). APSIM – Evolution towards a new generation of agricultural systems simulation. *Environmental Modelling & Software*, 62, 327-350.
- Houborg, R., & McCabe, M. (2016). High-Resolution NDVI from Planet's Constellation of Earth Observing Nano-Satellites: A New Data Source for Precision Agriculture. *Remote Sensing*, 8(9)
- Houborg, R., & McCabe, M. (2018a). Daily Retrieval of NDVI and LAI at 3 m Resolution via the Fusion of CubeSat, Landsat, and MODIS Data. *Remote Sensing*, 10(6), 890.
- Houborg, R., & McCabe, M.F. (2018b). A Cubesat Enabled Spatio-Temporal Enhancement Method (CESTEM) Utilizing Planet, Landsat and MODIS Data. *Remote Sensing of Environment*, 209, 211-226.
- Houborg, R., & McCabe, M.F. (2018c). A hybrid training approach for leaf area index estimation via Cubist and random forests machine-learning. *ISPRS Journal of Photogrammetry and Remote Sensing*, 135, 173-188.
- Houborg, R., McCabe, M.F., & Gao, F. (2016). A Spatio-Temporal Enhancement Method for medium resolution LAI (STEM-LAI). *International Journal of Applied Earth Observation and Geoinformation*, 47, 15-29.
- Huang, J., Gómez-Dans, J.L., Huang, H., Ma, H., Wu, Q., Lewis, P.E., et al. (2019). Assimilation of remote sensing into crop growth models: Current status and perspectives. *Agricultural and Forest Meteorology*, 276, 107609.
- Huang, J., Ma, H., Su, W., Zhang, X., Huang, Y., Fan, J., et al. (2015). Jointly Assimilating MODIS LAI and et Products into the SWAP Model for Winter Wheat Yield Estimation. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 8(8), 4060-4071.
- Idso, S.B., Jackson, R.D., & Reginato, R.J. (1977). Remote-sensing of crop yields. *Science*, 196(4285), 19-25.
- Ines, A.V., Das, N.N., Hansen, J.W., & Njoku, E.G. (2013). Assimilation of remotely sensed soil moisture and vegetation with a crop simulation model for maize yield prediction. *Remote Sensing of Environment*, 138, 149-164.
- Israel Meteorological Service. 2020. Accessed 18.08.2020. <https://ims.gov.il/en/climateAtlas>.
- Jain, M., Singh, B., Srivastava, A.A.K., Malik, R.K., McDonald, A.J., & Lobell, D.B. (2017). Using satellite data to identify the causes of and potential solutions for yield gaps in India's Wheat Belt. *Environmental Research Letters*, 12(9)

- Jain, M., Srivastava, A., Balwinder, S., Joon, R., McDonald, A., Royal, K., et al. (2016). Mapping Smallholder Wheat Yields and Sowing Dates Using Micro-Satellite Data. *Remote Sensing*, 8(10), 860-878.
- Jeffrey, S.J., Carter, J.O., Moodie, K.B., & Beswick, A.R. (2001). Using spatial interpolation to construct a comprehensive archive of Australian climate data. *Environmental Modelling & Software*, 16(4), 309-330.
- Jeffries, G.R., Griffin, T.S., Fleisher, D.H., Naumova, E.N., Koch, M., & Wardlow, B.D. (2019). Mapping sub-field maize yields in Nebraska, USA by combining remote sensing imagery, crop simulation models, and machine learning. *Precision Agriculture*, 1-17.
- Jiang, Z., Huete, A.R., Didan, K., & Miura, T. (2008). Development of a two-band enhanced vegetation index without a blue band. *Remote Sensing of Environment*, 112(10), 3833-3845.
- Jin, N., Tao, B., Ren, W., Feng, M., Sun, R., He, L., et al. (2016). Mapping Irrigated and Rainfed Wheat Areas Using Multi-Temporal Satellite Data. *Remote Sensing*, 8(3), 207.
- Jin, S., Yang, L., Danielson, P., Homer, C., Fry, J., & Xian, G. (2013). A comprehensive change detection method for updating the National Land Cover Database to circa 2011. *Remote Sensing of Environment*, 132, 159-175.
- Jin, Z., Azzari, G., Burke, M., Aston, S., & Lobell, D.B. (2017a). Mapping Smallholder Yield Heterogeneity at Multiple Scales in Eastern Africa. *Remote Sensing*, 9(9), 931.
- Jin, Z., Azzari, G., & Lobell, D.B. (2017b). Improving the accuracy of satellite-based high-resolution yield estimation: A test of multiple scalable approaches. *Agricultural and Forest Meteorology*, 247, 207-220.
- Jin, Z., Azzari, G., You, C., Di Tommaso, S., Aston, S., Burke, M., et al. (2019). Smallholder maize area and yield mapping at national scales with Google Earth Engine. *Remote Sensing of Environment*, 228, 115-128.
- Jordan, C.F. (1969). Derivation of leaf-area index from quality of light on the forest floor. *Ecology*, 50(4), 663-666.
- Kamir, E., Waldner, F., & Hochman, Z. (2020). Estimating wheat yields in Australia using climate records, satellite image time series and machine learning methods. *ISPRS Journal of Photogrammetry and Remote Sensing*, 160, 124-135.
- Kanungo, T., Mount, D.M., Netanyahu, N.S., Piatko, C.D., Silverman, R., & Wu, A.Y. (2002). An efficient k-means clustering algorithm: Analysis and implementation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(7), 881-892.
- Kassam, A., Friedrich, T., Derpsch, R., & Kienzle, J. (2015). Overview of the worldwide spread of conservation agriculture. *Field Actions Science Report*, 8
- Khan, A., Hansen, M.C., Potapov, P.V., Adusei, B., Pickens, A., Krylov, A., et al. (2018). Evaluating landsat and rapideye data for winter wheat mapping and area estimation in punjab, pakistan. *Remote Sensing*, 10(4), 489.
- Kimm, H., Guan, K., Jiang, C., Peng, B., Gentry, L.F., Wilkin, S.C., et al. (2020). Deriving high-spatiotemporal-resolution leaf area index for agroecosystems in the US Corn Belt using Planet Labs CubeSat and STAIR fusion data. *Remote Sensing of Environment*, 239, 111615.
- Knyazikhin, Y., & Myneni, R. (2018). VIIRS Leaf Area Index (LAI) and Fraction of Photosynthetically Active Radiation Absorbed by Vegetation (FPAR) User Guide

- Kolecka, N., Ginzler, C., Pazur, R., Price, B., & Verburg, P.H. (2018). Regional scale mapping of grassland mowing frequency with Sentinel-2 time series. *Remote Sensing*, 10(8)
- Kussul, N., Lavreniuk, M., Skakun, S., & Shelestov, A. (2017). Deep Learning Classification of Land Cover and Crop Types Using Remote Sensing Data. *IEEE Geoscience and Remote Sensing Letters*, 14(5), 778-782.
- Labus, M., Nielsen, G., Lawrence, R., Engel, R., & Long, D. (2002). Wheat yield estimates using multi-temporal NDVI satellite imagery. *International Journal of Remote Sensing*, 23(20), 4169-4180.
- Lai, Y., Pringle, M., Kopittke, P.M., Menzies, N.W., Orton, T.G., & Dang, Y.P. (2018). An empirical model for prediction of wheat yield, using time-integrated Landsat NDVI. *International Journal of Applied Earth Observation and Geoinformation*, 72, 99-108.
- Lawes, R., Chen, C., & van Rees, H. (2018). The National Paddock Survey—What causes the yield gap across Australian paddocks? In, *GRDC Update, Perth, Western Australia, February 2018*: GRDC.
- Leach, N., Coops, N.C., & Obrknezev, N. (2019). Normalization method for multi-sensor high spatial and temporal resolution satellite imagery with radiometric inconsistencies. *Computers and Electronics in Agriculture*, 164, 104893.
- LI-COR, I. (1992). LAI-2000 Plant Canopy Analyzer. In, *Operating Manual*
- Li, W., Jiang, J., Guo, T., Zhou, M., Tang, Y., Wang, Y., et al. (2019). Generating Red-Edge Images at 3 M Spatial Resolution by Fusing Sentinel-2 and Planet Satellite Products. *Remote Sensing*, 11(12), 1422.
- Liu, W., Huang, J., Wei, C., Wang, X., Mansaray, L.R., Han, J., et al. (2018). Mapping water-logging damage on winter wheat at parcel level using high spatial resolution satellite data. *ISPRS Journal of Photogrammetry and Remote Sensing*, 142, 243-256.
- Lobell, D.B., Asner, G.P., Ortiz-Monasterio, J.I., & Benning, T.L. (2003). Remote sensing of regional crop production in the Yaqui Valley, Mexico: Estimates and uncertainties. *Agriculture, Ecosystems and Environment*, 94(2), 205-220.
- Lobell, D.B., Ortiz-Monasterio, J.I., Sibley, A.M., & Sohu, V.S. (2013). Satellite detection of earlier wheat sowing in India and implications for yield trends. *Agricultural Systems*, 115, 137-143.
- Lobell, D.B., Thau, D., Seifert, C., Engle, E., & Little, B. (2015). A scalable satellite-based crop yield mapper. *Remote Sensing of Environment*, 164, 324-333.
- Louis, J., Debaecker, V., Pflug, B., Main-Knorn, M., Bieniarz, J., Mueller-Wilm, U., et al. (2016). Sentinel-2 SEN2COR: L2A processor for users. In, *Proceedings of the Living Planet Symposium, Prague, Czech Republic* (pp. 9-13)
- Lu, D., Mausel, P., Brondízio, E., & Moran, E. (2004). Change detection techniques. *International Journal of Remote Sensing*, 25(12), 2365-2407.
- Luo, Y., Guan, K., & Peng, J. (2018). STAIR: A generic and fully-automated method to fuse multiple sources of optical satellite data to generate a high-resolution, daily and cloud-/gap-free surface reflectance product. *Remote Sensing of Environment*, 214, 87-99.
- Ma, G., Huang, J., Wu, W., Fan, J., Zou, J., & Wu, S. (2013). Assimilation of MODIS-LAI into the WOFOST model for forecasting regional winter wheat yield. *Mathematical and Computer Modelling*, 58(3-4), 634-643.
- Manfron, G., Delmotte, S., Busetto, L., Hossard, L., Ranghetti, L., Brivio, P.A., et al. (2017). Estimating inter-annual variability in winter wheat sowing dates from satellite time series in Camargue, France. *International Journal of Applied Earth Observation and Geoinformation*, 57, 190-201.

- Manivasagam, V., Sadeh, Y., Kaplan, G., Bonfil, D.J., & Rozenstein, O. (2021). Studying the Feasibility of Assimilating Sentinel-2 and PlanetScope Imagery into the SAFY Crop Model to Predict Within-Field Wheat Yield. *Remote Sensing*, 13(12), 2395.
- Marais Sicre, C., Inglada, J., Fieuzal, R., Baup, F., Valero, S., Cros, J., et al. (2016). Early Detection of Summer Crops Using High Spatial Resolution Optical Image Time Series. *Remote Sensing*, 8(7)
- Marinho, E., Vancutsem, C., Fasbender, D., Kayitakire, F., Pini, G., & Pekel, J.F. (2014). From remotely sensed vegetation onset to sowing dates: Aggregating pixel-level detections into village-level sowing probabilities. *Remote Sensing*, 6(11), 10947-10965.
- Mas, J.F. (1999). Monitoring land-cover changes: A comparison of change detection techniques. *International Journal of Remote Sensing*, 20(1), 139-152.
- Mathison, C., Deva, C., Falloon, P., & Challinor, A.J. (2017). Defining sowing and harvest dates based on the Asian Summer Monsoon. *Earth System Dynamics Discussions*, 1-39.
- Maynard, J.J., Karl, J.W., & Browning, D.M. (2016). Effect of spatial image support in detecting long-term vegetation change from satellite time-series. *Landscape ecology*, 31(9), 2045-2062.
- Mc Nairn, H., Wood, D., Gwyn, Q.H.J., Brown, R.J., & Charbonneau, F. (1998). Mapping tillage and crop residue management practices with RADARSAT. *Canadian Journal of Remote Sensing*, 24(1), 28-35.
- McCabe, M.F., Aragon, B., Houborg, R., & Mascaro, J. (2017). CubeSats in Hydrology: Ultrahigh-Resolution Insights Into Vegetation Dynamics and Terrestrial Evaporation. *Water Resources Research*, 53(12), 10017-10024.
- McCarty, J.L., Neigh, C.S.R., Carroll, M.L., & Wooten, M.R. (2017). Extracting smallholder cropped area in Tigray, Ethiopia with wall-to-wall sub-meter WorldView and moderate resolution Landsat 8 imagery. *Remote Sensing of Environment*
- Myneni, R., & Park, Y. (2015). MODIS collection 6 (C6) LAI/FPAR product user's guide. In: Feb.
- Nakalembe, C., Becker-Reshef, I., Bonifacio, R., Hu, G., Humber, M.L., Justice, C.J., et al. (2021). A review of satellite-based global agricultural monitoring systems available for Africa. *Global Food Security*, 29, 100543.
- Nguy-Robertson, A., Gitelson, A., Peng, Y., Viña, A., Arkebauer, T., & Rundquist, D. (2012). Green leaf area index estimation in maize and soybean: Combining vegetation indices to achieve maximal sensitivity. *Agronomy Journal*, 104(5), 1336-1347.
- Nguy-Robertson, A.L., Peng, Y., Gitelson, A.A., Arkebauer, T.J., Pimstein, A., Herrmann, I., et al. (2014). Estimating green LAI in four crops: Potential of determining optimal spectral bands for a universal algorithm. *Agricultural and Forest Meteorology*, 192, 140-148.
- Ortiz-Monasterio, J.I., & Lobell, D.B. (2007). Remote sensing assessment of regional yield losses due to sub-optimal planting dates and fallow period weed management. *Field Crops Research*, 101(1), 80-87.
- Ozturk, A., Caglar, O., & Bulut, S. (2006). Growth and yield response of facultative wheat to winter sowing, freezing sowing and spring sowing at different seeding rates. *Journal of Agronomy and Crop Science*, 192(1), 10-16.
- Pacheco, A.M., McNairn, H., & Merzouki, A. (2010). Evaluating TerraSAR-X for the identification of tillage occurrence over an agricultural area in Canada. In, *Proceedings of SPIE - The International Society for Optical Engineering*

- Pan, H., Chen, Z., de Wit, A., & Ren, J. (2019). Joint assimilation of leaf area index and soil moisture from Sentinel-1 and Sentinel-2 data into the WOFOST model for winter wheat yield estimation. *Sensors*, *19*(14), 3161.
- Pan, Z., Huang, J., Zhou, Q., Wang, L., Cheng, Y., Zhang, H., et al. (2015). Mapping crop phenology using NDVI time-series derived from HJ-1 A/B data. *International Journal of Applied Earth Observation and Geoinformation*, *34*, 188-197.
- Pasqualotto, N., Delegido, J., Van Wittenberghe, S., Rinaldi, M., & Moreno, J. (2019). Multi-Crop Green LAI Estimation with a New Simple Sentinel-2 LAI Index (SeLI). *Sensors*, *19*(4), 904.
- Peng, Y., & Gitelson, A.A. (2011). Application of chlorophyll-related vegetation indices for remote estimation of maize productivity. *Agricultural and Forest Meteorology*, *151*(9), 1267-1276.
- Planet Team. 2018. "Planet imagery product specifications." Planet Labs Inc, Accessed 12/04/2018. https://www.planet.com/products/satellite-imagery/files/Planet_Combined_Imagery_Product_Specs_December2017.pdf.
- Planet Team. 2020. "Planet Surface Reflectance Product v2." Planet Labs, Inc, Accessed 18.08.2020. https://assets.planet.com/marketing/PDF/Planet_Surface_Reflectance_Technical_White_Paper.pdf.
- Pohl, C., & Van Genderen, J.L. (1998). Review article multisensor image fusion in remote sensing: concepts, methods and applications. *International Journal of Remote Sensing*, *19*(5), 823-854.
- Pollock, R., & Kanemasu, E. (1979). Estimating leaf-area index of wheat with Landsat data. *Remote Sensing of Environment*, *8*(4), 307-312.
- Potgieter, A.B., Lobell, D.B., Hammer, G.L., Jordan, D.R., Davis, P., & Brider, J. (2016). Yield trends under varying environmental conditions for sorghum and wheat across Australia. *Agricultural and Forest Meteorology*, *228-229*, 276-285.
- Prasad, A.K., Chai, L., Singh, R.P., & Kafatos, M. (2006). Crop yield estimation model for Iowa using remote sensing and surface parameters. *International Journal of Applied Earth Observation and Geoinformation*, *8*(1), 26-33.
- Psomiadis, E., Dercas, N., Dalezios, N.R., & Spyropoulos, N.V. (2017). Evaluation and cross-comparison of vegetation indices for crop monitoring from sentinel-2 and worldview-2 images. In, *Remote Sensing for Agriculture, Ecosystems, and Hydrology XIX* (p. 104211B): International Society for Optics and Photonics.
- Qi, J., Chehbouni, A., Huete, A., Kerr, Y., & Sorooshian, S. (1994). A modified soil adjusted vegetation index. *Remote Sensing of Environment*, *48*(2), 119-126.
- Rajendran, S., Al-Sayigh, A.R., & Al-Awadhi, T. (2016). Vegetation analysis study in and around Sultan Qaboos University, Oman, using Geoeye-1 satellite data. *Egyptian Journal of Remote Sensing and Space Science*, *19*(2), 297-311.
- Raun, W.R., Solie, J.B., Johnson, G.V., Stone, M.L., Lukina, E.V., Thomason, W.E., et al. (2001). In-season prediction of potential grain yield in winter wheat using canopy reflectance. *Agronomy Journal*, *93*(1), 131-138.
- Raun, W.R., Solie, J.B., Johnson, G.V., Stone, M.L., Mullen, R.W., Freeman, K.W., et al. (2002). Improving nitrogen use efficiency in cereal grain production with optical sensing and variable rate application. *Agronomy Journal*, *94*(4), 815-820.
- Ray, D.K., Gerber, J.S., MacDonald, G.K., & West, P.C. (2015). Climate variation explains a third of global crop yield variability. *Nature communications*, *6*(1), 1-9.

- Rondeaux, G., Steven, M., & Baret, F. (1996). Optimization of soil-adjusted vegetation indices. *Remote Sensing of Environment*, 55(2), 95-107.
- Roujean, J.-L., & Breon, F.-M. (1995). Estimating PAR absorbed by vegetation from bidirectional reflectance measurements. *Remote Sensing of Environment*, 51(3), 375-384.
- Rouse, J., Haas, R., Schell, J., & Deering, D. (1974). Monitoring vegetation systems in the Great Plains with ERTS. In, *Washington, DC* (pp. 309–317): NASA SP-351.
- Sacks, W.J., Deryng, D., Foley, J.A., & Ramankutty, N. (2010). Crop planting dates: An analysis of global patterns. *Global Ecology and Biogeography*, 19(5), 607-620.
- Sadeh, Y., Cohen, H., Maman, S., & Blumberg, D.G. (2018). Evaluation of Manning's roughness coefficient in arid environments by using SAR backscatter. *Remote Sensing*, 10(10), 1505.
- Sadeh, Y., Zhu, X., Chenu, K., & Dunkerley, D. (2019). Sowing date detection at the field scale using CubeSats remote sensing. *Computers and Electronics in Agriculture*, 157, 568-580.
- Sadeh, Y., Zhu, X., Dunkerley, D., Walker, J.P., Zhang, Y., Rozenstein, O., et al. (2021). Fusion of Sentinel-2 and PlanetScope time-series data into daily 3 m surface reflectance and wheat LAI monitoring. *International Journal of Applied Earth Observation and Geoinformation*, 96, 102260.
- Sagan, V., Maimaitijiang, M., Bhadra, S., Maimaitiyiming, M., Brown, D.R., Sidike, P., et al. (2021). Field-scale crop yield prediction using multi-temporal WorldView-3 and PlanetScope satellite data and deep learning. *ISPRS Journal of Photogrammetry and Remote Sensing*, 174, 265-281.
- Sakamoto, T., Wardlow, B.D., Gitelson, A.A., Verma, S.B., Suyker, A.E., & Arkebauer, T.J. (2010). A two-step filtering approach for detecting maize and soybean phenology with time-series MODIS data. *Remote Sensing of Environment*, 114(10), 2146-2159.
- Sakamoto, T., Yokozawa, M., Toritani, H., Shibayama, M., Ishitsuka, N., & Ohno, H. (2005). A crop phenology detection method using time-series MODIS data. *Remote Sensing of Environment*, 96(3-4), 366-374.
- Sedano, F., Kempeneers, P., Strobl, P., Kucera, J., Vogt, P., Seebach, L., et al. (2011). A cloud mask methodology for high resolution remote sensing data combining information from high and medium resolution optical sensors. *ISPRS Journal of Photogrammetry and Remote Sensing*, 66(5), 588-596.
- Serbin, G., Daughtry, C.S.T., Hunt Jr, E.R., Brown, D.J., & McCarty, G.W. (2009). Effect of soil spectral properties on remote sensing of crop residue cover. *Soil Science Society of America Journal*, 73(5), 1545-1558.
- Shen, J., & Evans, F.H. (2021). The potential of Landsat NDVI sequences to explain wheat yield variation in fields in Western Australia. *Remote Sensing*, 13(11), 2202.
- SILO (2018). SILO – Getting started guide. In Science Information Services (Ed.). Brisbane: The State of Queensland (Department of Environment and Science).
- Skakun, S., Kalecinski, N.I., Brown, M.G., Johnson, D.M., Vermote, E.F., Roger, J.-C., et al. (2021). Assessing within-Field Corn and Soybean Yield Variability from WorldView-3, Planet, Sentinel-2, and Landsat 8 Satellite Imagery. *Remote Sensing*, 13(5), 872.
- Skakun, S., Vermote, E., Roger, J.-C., & Franch, B. (2017). Combined Use of Landsat-8 and Sentinel-2A Images for Winter Crop Mapping and Winter Wheat Yield Assessment at Regional Scale. *AIMS Geosciences*, 3(2), 163-186.
- Sripada, R.P., Heiniger, R.W., White, J.G., & Meijer, A.D. (2006). Aerial color infrared photography for determining early in-season nitrogen requirements in corn. *Agronomy Journal*, 98(4), 968-977.

- SUHET (2015). Sentinel-2 User Handbook. In (p. 64): The European Space Agency (ESA).
- Sun, L., Gao, F., Anderson, M., Kustas, W., Alsina, M., Sanchez, L., et al. (2017). Daily Mapping of 30 m LAI and NDVI for Grape Yield Prediction in California Vineyards. *Remote Sensing*, 9(4)
- Sun, R., Chen, S., Su, H., Mi, C., & Jin, N. (2019). The Effect of NDVI Time Series Density Derived from Spatiotemporal Fusion of Multisource Remote Sensing Data on Crop Classification Accuracy. *ISPRS International Journal of Geo-Information*, 8(11), 502.
- Tian, F., Wu, J., Liu, L., Leng, S., Yang, J., Zhao, W., et al. (2020). Exceptional drought across southeastern Australia caused by extreme lack of precipitation and its impacts on NDVI and SIF in 2018. *Remote Sensing*, 12(1), 54.
- Toulouse, T., Rossi, L., Celik, T., & Akhloufi, M. (2016). Automatic fire pixel detection using image processing: a comparative analysis of rule-based and machine learning-based methods. *Signal, Image and Video Processing*, 10(4), 647-654.
- Urban, D., Guan, K., & Jain, M. (2018). Estimating sowing dates from satellite data over the US Midwest: A comparison of multiple sensors and metrics. *Remote Sensing of Environment*, 211, 400-412.
- Valderrama-Landeros, L.H., España-Boquera, M.L., & Baret, F. (2016). Deforestation in Michoacan, Mexico, from CYCLOPES-LAI time series (2000–2006). *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 9(12), 5398-5405.
- Van der Sande, C., De Jong, S., & De Roo, A. (2003). A segmentation and classification approach of IKONOS-2 imagery for land cover mapping to assist flood risk and flood damage assessment. *International Journal of Applied Earth Observation and Geoinformation*, 4(3), 217-229.
- Van Diepen, C.v., Wolf, J., Van Keulen, H., & Rappoldt, C. (1989). WOFOST: a simulation model of crop production. *Soil use and management*, 5(1), 16-24.
- Van Niel, T.G., & McVicar, T.R. (2004). Determining temporal windows for crop discrimination with remote sensing: A case study in south-eastern Australia. *Computers and Electronics in Agriculture*, 45(1-3), 91-108.
- Verger, A., Filella, I., Baret, F., & Peñuelas, J. (2016). Vegetation baseline phenology from kilometric global LAI satellite products. *Remote Sensing of Environment*, 178, 1-14.
- Verrelst, J., Rivera, J.P., Veroustraete, F., Muñoz-Mari, J., Clevers, J.G.P.W., Camps-Valls, G., et al. (2015). Experimental Sentinel-2 LAI estimation using parametric, non-parametric and physical retrieval methods - A comparison. *ISPRS Journal of Photogrammetry and Remote Sensing*, 108, 260-272.
- Viña, A., Gitelson, A.A., Nguy-Robertson, A.L., & Peng, Y. (2011). Comparison of different vegetation indices for the remote assessment of green leaf area index of crops. *Remote Sensing of Environment*, 115(12), 3468-3478.
- Waldner, F., Horan, H., Chen, Y., & Hochman, Z. (2019). High temporal resolution of leaf area data improves empirical estimation of grain yield. *Scientific reports*, 9(1), 1-14.
- Walker, J.P., Houser, P.R., & Willgoose, G.R. (2004). Active microwave remote sensing for soil moisture measurement: a field evaluation using ERS-2. *Hydrological processes*, 18(11), 1975-1997.
- Watson, D.J. (1947). Comparative physiological studies on the growth of field crops: I. Variation in net assimilation rate and leaf area between species and varieties, and within and between years. *Annals of botany*, 11(41), 41-76.
- Weiss, M., & Baret, F. (2016). S2ToolBox Level 2 Products: LAI, FAPAR, FCOVER, Version 1.1. *ESA Contract n° 4000110612/14/I-BG (p. 52)*: INRA Avignon, France.

- Wiegand, C., Richardson, A., & Kanemasu, E. (1979). Leaf Area Index Estimates for Wheat from LANDSAT and Their Implications for Evapotranspiration and Crop Modeling 1. *Agronomy Journal*, 71(2), 336-342.
- Wold, S., Esbensen, K., & Geladi, P. (1987). Principal component analysis. *Chemometrics and Intelligent Laboratory Systems*, 2(1-3), 37-52.
- Zadoks, J.C., Chang, T.T., & Konzak, C.F. (1974). A decimal code for the growth stages of cereals. *Weed research*, 14(6), 415-421.
- Zhang, L., Feng, H., Jin, N., & Zhang, T. (2018). Mapping irrigated and rainfed wheat areas using high spatial-temporal resolution data generated by Moderate Resolution Imaging Spectroradiometer and Landsat. *Journal of Applied Remote Sensing*, 12(4), 046023.
- Zhao, Y., Potgieter, A.B., Zhang, M., Wu, B., & Hammer, G.L. (2020). Predicting wheat yield at the field scale by combining high-resolution Sentinel-2 satellite imagery and crop modelling. *Remote Sensing*, 12(6), 1024.
- Zheng, B., Campbell, J.B., Serbin, G., & Galbraith, J.M. (2014). Remote sensing of crop residue and tillage practices: Present capabilities and future prospects. *Soil and Tillage Research*, 138, 26-34.
- Zheng, B., Chapman, S.C., Christopher, J.T., Frederiks, T.M., & Chenu, K. (2015). Frost trends and their estimated impact on yield in the Australian wheatbelt. *J Exp Bot*, 66(12), 3611-3623.
- Zheng, B., Chenu, K., Fernanda Dreccer, M., & Chapman, S.C. (2012). Breeding for the future: What are the potential impacts of future frost and heat events on sowing and flowering time requirements for Australian bread wheat (*Triticum aestivum*) varieties? *Global change biology*, 18(9), 2899-2914.
- Zhong, Y., Zhang, L., Huang, B., & Li, P. (2006). An unsupervised artificial immune classifier for multi/hyperspectral remote sensing imagery. *IEEE Transactions on Geoscience and Remote Sensing*, 44(2), 420-431.
- Zhu, W., Pan, Y., He, H., Wang, L., Mou, M., & Liu, J. (2011). A changing-weight filter method for reconstructing a high-quality NDVI time series to preserve the integrity of vegetation phenology. *IEEE Transactions on Geoscience and Remote Sensing*, 50(4), 1085-1094.
- Zhu, Z., & Woodcock, C.E. (2012). Object-based cloud and cloud shadow detection in Landsat imagery. *Remote Sensing of Environment*, 118, 83-94.