MDPI

Review

Data Fusion in Agriculture: Resolving Ambiguities and Closing Data Gaps

Jayme Garcia Arnal Barbedo

Embrapa Digital Agriculture, Campinas 13083-886, SP, Brazil; jayme.barbedo@embrapa.br; Tel.: +55-19-3211-5880

Abstract: Acquiring useful data from agricultural areas has always been somewhat of a challenge, as these are often expansive, remote, and vulnerable to weather events. Despite these challenges, as technologies evolve and prices drop, a surge of new data are being collected. Although a wealth of data are being collected at different scales (i.e., proximal, aerial, satellite, ancillary data), this has been geographically unequal, causing certain areas to be virtually devoid of useful data to help face their specific challenges. However, even in areas with available resources and good infrastructure, data and knowledge gaps are still prevalent, because agricultural environments are mostly uncontrolled and there are vast numbers of factors that need to be taken into account and properly measured for a full characterization of a given area. As a result, data from a single sensor type are frequently unable to provide unambiguous answers, even with very effective algorithms, and even if the problem at hand is well defined and limited in scope. Fusing the information contained in different sensors and in data from different types is one possible solution that has been explored for some decades. The idea behind data fusion involves exploring complementarities and synergies of different kinds of data in order to extract more reliable and useful information about the areas being analyzed. While some success has been achieved, there are still many challenges that prevent a more widespread adoption of this type of approach. This is particularly true for the highly complex environments found in agricultural areas. In this article, we provide a comprehensive overview on the data fusion applied to agricultural problems; we present the main successes, highlight the main challenges that remain, and suggest possible directions for future research.

Keywords: data fusion; sensors; variability; precision agriculture; artificial intelligence

1. Introduction

The number (and quality) of sensors used to collect data in different contexts have been steadily growing. Even complex environments, such as agricultural areas, are now being "sensed" via a wide variety of equipment, generating vast amounts of data that can be explored to provide useful information about the area being observed. As a result, the number of studies attempting to explore the wealth of information contained in the sensed data have increased [1–3]. However, it is often challenging to translate the advancements achieved in experiments to the conditions found in practice. There are two main reasons for this. First, the studies described in scientific texts are usually limited in scope, because the data used in these experiments usually do not cover all of the variabilities associated with the problem at hand. As a result, while the results reported in those articles may seem encouraging, they often reveal nothing about the performance of the proposed technique under real, unconstrained conditions. Second, even if the data adequately cover the variable conditions found in practice, the adopted sensing technology may not be capable of acquiring enough information to unambiguously resolve the data and provide enough information. For example, even powerful artificial intelligence models fed with RGB digital images are often unsuccessful in recognizing plant diseases from their symptoms, because different disorders can produce similar visual signs [4].

One way to reduce the gaps caused by data limitations is to apply data fusion techniques. The term "data fusion" can be defined as "the process of combining data from multiple sources to produce more accurate, consistent, and concise information than that provided by any individual data source" [5]. Other stricter definitions do exist to better fit narrower contexts. This type of approach has been applied to agricultural problems since the first half of the 1990s [6], and there has been an increase in the use of this approach. Arguably, the main challenge involved in the use of data fusion techniques involves finding the best approach to fully explore the synergy and complementarities that potentially exist between different types of data and data sources. This is particularly true with data having significantly disparate characteristics (for example, digital images and meteorological data).

It is difficult to find a formalization for the data fusion process that fits all agricultural applications, given the variety of data sources and approaches. The formalization presented by Bleiholder and Naumann [7], although derived in a slightly different context, adopts a three-step view of the data fusion process that is applicable in most cases. In the first step, the corresponding attributes that are used to describe the information in different sources need to be identified. Such a correspondence can be easily identified if the data sources are similar, but it can be challenging as the different types of data are being used. This is one of the main reasons for the existence of the three types of data fusion described in the following paragraph. In the second step, the different objects that are described in the data sources need to be identified and aligned. This step is particularly important when data sources are images, because misalignments can lead to inconsistent representations and, as a result, to unreliable answers. Once the data are properly identified and consistent, the actual data fusion can be applied in the third step. In practice, coping with existing data inconsistencies is often ignored [7]. This situation can be (at least partially) remedied by auxiliary tools, such as data profile techniques, which can reduce inconsistencies by extracting and exploring the metadata associated to the data being fused [8].

The most common categorization divides data fusion techniques into three groups [9]: (a) raw data level (also denoted "low-level" and "early integration"), in which different types of data (raw or preprocessed) are simply concatenated into a single matrix, being used in cases in which pieces of data are of the same nature and were properly normalized. (b) Feature level (also denoted "mid-level" and "intermediate integration"), in which features are first extracted from different types of data and then concatenated into a matrix, being mostly used when pieces of data can be treated in such a way they generate features that are compatible and complementary. (c) Decision level (also denoted "high level" and "late integration"), in which classification and regression algorithms are applied separately to each type of datum and then the outputs generated by each model are combined, being more appropriate when data sources are too distinct to be combined at an earlier stage. An alternative classification of data fusion methods was proposed by Ouhami et al. [10]: probability-based, evidence-based, and knowledge-based. Although both classifications are useful, the first one is more appropriate in the context of this work (Figure 1).

In the specific case of agriculture, data can be collected at three different scales—proximal, aerial, and orbital (satellites) (Figure 1). Applications that use proximal data include navigation systems for autonomous vehicles [11–17], fruit detection [18–21], plant disease detection [22–24], delineation of homogeneous management zones [25–29], soil analysis [30–36], plant phenotyping [37], among others. Aerial data (collected using UAVs) is used mostly for detection of certain objects (e.g., certain plant species and fruits) [38] and for estimation of agricultural variables (e.g., soil moisture and nitrogen content) [39–41]. Satellite data are used for mapping variables as diverse as soil moisture [42–44], crop type [45–50], crop phenological states [51,52], evapotranspiration [40,53–58], nitrogen status [59–62], biomass [63,64], among others. While most data fusion approaches only use data in the same scale, a few studies have applied data originating from different scales [10,26,28,31,38,40,51,52,64–71].
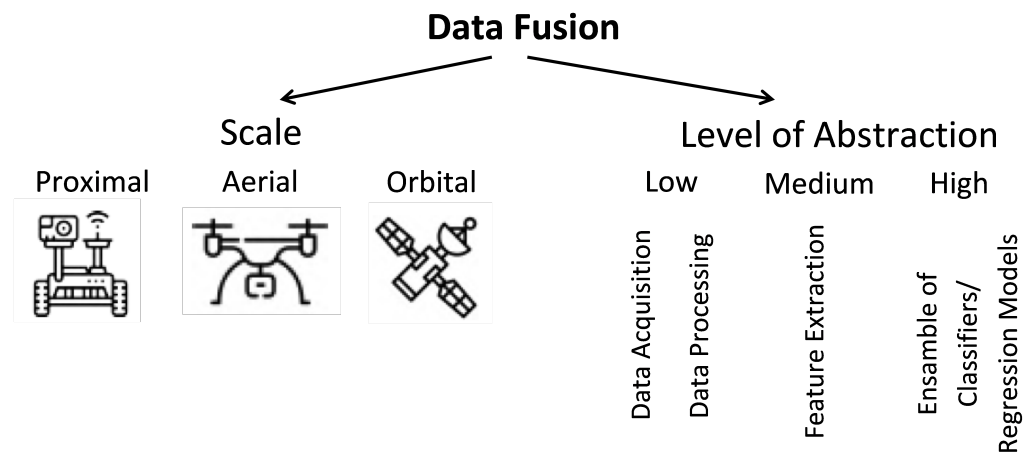
Figure 1. Categorization of data fusion approaches adopted in this work.

The objective of this article was to characterize the current state-of-the-art regarding the process of applying data fusion for agricultural applications. First, a comprehensive overview of the literature is provided, with emphasis on articles published after 2010. Then, the aspects involved in the use of data fusion to different types of data are explored in detail, together with some possible solutions for the weaknesses that prevent technologies based on data fusion from being more widely used in practice. Although there have been a few reviews dedicated toward data fusion in agriculture, they focused on specific applications and themes [9,10,68,70,72], while this one adopts a more general and systemic view of the subject.

The remainder of the article is organized as follows. Section 2 describes the literature related to the three different scales—proximal, aerial, and orbital. A discussion on several aspects relevant to data fusion in agriculture is presented in Section 3. Finally, Section 4 offers some final remarks and possible directions for future research.

2. Literature Review

A search was carried out via Google Scholar and Scopus using the keywords "data fusion" and "agriculture". Almost 400 articles were originally selected, but this number was reduced to 119 after we removed low quality documents or documents outside of the scope of this work. From this total, 50 explore proximally collected data, 45 explore satellite data, 8 explore aerial (UAV) data, and 16 use data collected at multiple scales (Figure 2). Each scale is treated separately in this section, as they have some peculiarities that require more focused descriptions and analyses. To improve the legibility of the tables containing the full list of references, many acronyms are used, all of which are defined in Table 1.
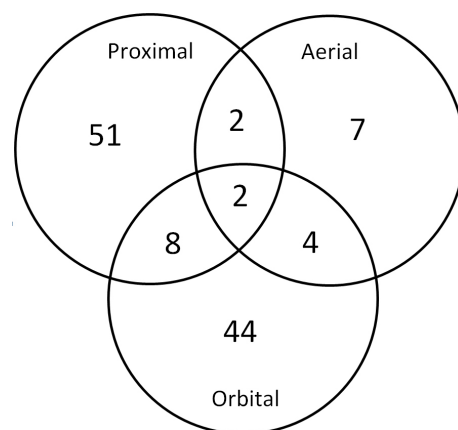
Figure 2. Number of references selected in this review.

Table 1. List of acronyms.

Acronym	Meaning	Acronym	Meaning
AMSR-E	Advanced Microwave Scanning Radiometer on the Earth Observing System	MLP	Multilayer Perceptron
		MLR	Multiple Linear Regression
ANN	Artificial Neural Network	MOA	Model Output Averaging
ASTER	Advanced Spaceborne Thermal Emission and Reflection	MODIS	Moderate-Resolution Imaging Spectroradiometer
		MSDF-ET	Multi-Sensor Data Fusion Model for Actual Evapotranspiration Estimation
BK	Block Kriging		
BPNN	Backpropagation Neural Network	MSPI	Maximum Sum of Probabilities Intersections
CACAO	Consistent Adjustment of the Climatology to Actual Observations	NB	Naïve Bayes
		NDSI	Normalized Difference Spectral Index
CHRIS	Compact High Resolution Imaging Spectrometer	NDVI	Normalized Difference Vegetation Index
CNN	Convolutional Neural Network	NIR	Near-infrared Spectroscopy
CP-ANN	Counter-Propagation Artificial Neural Networks	NMDI	Normalized Multiband Drought Index
CV	Computer Vision	OLI	Operational Land Imager
DEM	Digital Elevation Model	PCA	Principal Component Analysis
DNN	Deep Neural Network	PDI	Perpendicular Drought Index
DRF	Distributed Random Forest	PLSR	Partial Least Square Regression
ECa	Apparent Soil Electrical Conductivity	RF	Random Forest
EDXRF	Energy dispersive X-Ray Fluorescence	RFR	Random Forest Regression
EKF	Extended Kalman Filter	RGB	Red–Green–Blue
ELM	Extreme Learning Machine	RGB-D	Red–Green–Blue-Depth
EMI	Electromagnetic Induction	RK	Regression Kriging
ESTARFM	Enhanced Spatial and Temporal Adaptive Reflective Fusion Model	RTK	Real Time Kinematic
		SADFAET	Spatiotemporal Adaptive Data Fusion Algorithm for Evapotranspiration Mapping
ET	Evapotranspiration		
FARMA	Fusion Approach for Remotely-Sensed Mapping of Agriculture	SAR	Synthetic Aperture Radar
		SF	Sensor Fusion
GBM	Gradient Boosting Machine	SfM	Structure from Motion
GKSFM	Gaussian Kernel-Based Spatiotemporal Fusion Model	SKN	Supervised Kohonen Networks
GLM	Generalized Linear Model	SMLR	Stepwise Multiple Linear Regression
GNSS	Global navigation satellite system	SPA	Successive Projections Algorithm
HUTS	High-resolution Urban Thermal Sharpener	SPOT	Satellite Pour l'Observation de la Terre
INS	Inertial Navigation System	SRTM	Shuttle Radar Topographic Mission
IoT	Internet of Things	STARFM	Spatial and Temporal Adaptive Reflective Fusion Model
ISTDFA	Improved Spatial and Temporal Data Fusion Approach		
		SVR	Support Vector Regression
kNN	k-Nearest Neighbors	TLS	Terrestrial Laser Scanning
LAI	Leaf Area Index	TRMM	Tropical Rainfall Measuring Mission
LPT	Laplacian Pyramid Transform	TVDI	Temperature Vegetation Dryness Index
LR	Linear Regression	UAV	Unmanned Aerial Vehicle
LSTM-NN	Long Short-Term Memory Neural Network	XGBoost	Extreme Gradient Boosting

To better organize the literature presented in this section and the subsequent discussion, both the data fusion techniques and the data being fused were categorized in a few more general classes (Table 2).

Table 2. Categories adopted for the data fusion techniques and the data being fused.

No.	Classes of Data Fusion Technique	No.	Classes of Data Being Fused
1	Regression methods	1	RGB images
2	STARFM-like statistical methods	2	Multispectral images
3	Geostatistical tools	3	Hyperspectral images
4	PCA and derivatives	4	Thermal images
5	Kalman filter	5	Laser scanning
6	Machine learning	6	SAR images
7	Deep learning	7	Spectroscopy
8	Decision rules	8	Fluorescence images
9	Majority rules	9	Soil measurements
10	Model output averaging	10	Environmental/weather measurements
11	Others	11	Inertial measurements
		12	Position measurements
		13	Topographic records and elevation models
		14	Historical data
		15	Others

2.1. Proximal Scale

The majority of studies dedicated to the proximal scale are concentrated in three main areas: prediction of soil properties, delineation of homogeneous zones, and robotic navigation and control. Applications, such as disease and fruit detection, prediction of water content and water stress, estimation of phonological state and yield prediction, are also present. Ten of the references also explored satellite data, and five studies combined proximal and aerial data. Data sources included cameras (RGB, multispectral, thermal, hyperspectral) spectrometers, conductance and resistivity sensors, GPS, inertial sensors, weather data, among many others. With such a variety of sensors available for field applications, efforts to explore their complementarities have been steadily growing (Table 3), but most problems still lack reliable solutions [73].

Table 3. References considered in this study–proximal scale. L, M, and H mean low-, mid-, and high-level data fusion, respectively. The numbers in the fourth column are those adopted in Table 2 for each "fused data" class.

Reference	Application	Fusion Technique	Fused Data	Mean Accuracy
[30]	Estimation of soil indices	SF (L), MOA (H)	7	0.80–0.90
[74]	Sustainable greenhouse management	Decision rules (L)	10	N/A
[73]	Human—robot interaction	LSTM-NN (L)	11	0.71–0.97
[25]	Delineation of homogeneous zones in viticulture	GAN (L), geostatistical tools (L)	2, 9	N/A
[26] [a]	Delineation of homogeneous zones	Kriging and other geostatistical tools (L)	2, 9	N/A
[51] [a]	Estimation of crop phenological states	Particle filter scheme (L)	2, 6, 10	0.93–0.96
[18]	Fruit detection	LPT (L) and fuzzy logic (L)	1, 4	0.80–0.95
[31] [a]	In-field estimation of soil properties	RK (L), PLSR (L)	3, 9	>0.5
[75]	Delineation of homogeneous management zones	Kriging (L), Gaussian anamorphosis (L)	9, 15	0.66
[76]	Delineation of homogeneous management zones	Kriging (L), Gaussian anamorphosis (L)	9, 15	N/A
[27]	Delineation of homogeneous management zones	Kriging (L), Gaussian anamorphosis (L)	9, 15	N/A
[77]	Crop nutritional status determination	PCA (L)	7, 8	0.7–0.9
[22]	Detection of olive quick decline syndrome	CNN (M)	1	0.986
[65] [b]	Monitoring Agricultural Terraces	Coregistering and information extraction (L/M)	5	N/A
[78]	Prediction of canopy water content of rice	BPNN (M), RF (M), PLSR (M)	2	0.98–1.00
[11]	Localization of a wheeled mobile robot	Dempster–Shafer (L) and Kalman filter (L)	11, 12	0.97
[19]	Immature green citrus fruit detection	Color-thermal probability algorithm (H)	1, 4	0.90–0.95
[28] [a]	Delineation of management zones	K-means clustering (L)	2, 9, 14	N/A
[79]	Segmentation for targeted application of products	Discrete wavelets transform (M)	1	0.92
[12]	System for agricultural vehicle positioning	Kalman filter (L)	11, 12	N/A
[13]	System for agricultural vehicle positioning	Kalman filter (L)	11, 12	N/A
[67] [a]	Yield gap attribution in maize	Empirical equations (L)	15	0.37–0.74
[32]	Soil environmental quality assessment	Analytic hierarchy process, weighted average (L)	15	N/A
[33]	Predict soil properties	PLSR (L)	7, 9, 13	0.80–0.96
[14]	System for agricultural vehicle positioning	Discrete Kalman filter (L)	11, 13	N/A
[34]	Estimating soil macronutrients	PLSR (L)	7, 9	0.70–0.95
[20]	Citrus fruit detection and localization	Daubechies wavelet transform (L)	1, 2	0.91
[15]	Estimation of agricultural equipment roll angle	Kalman filtering (L)	11	N/A
[80]	Predicting toxic elements in the soil	PLSR, PCA, and SPA (L/M)	7, 8	0.93–0.98
[68] [a]	Review: image fusion technology in agriculture	N/A	N/A	N/A
[81]	Heterogeneous sensor data fusion	Deep multimodal encoder (L)	10	N/A
[82]	Agricultural vulnerability assessments	Binary relevance (L), RF (L), and XGBoost (L)	10, 14	0.67–0.98
[35]	Prediction of multiple soil properties	SMLR (L), PLSR (L), PCA/SMLR combination (L)	7, 9	0.60–0.95
[83]	Prediction of environment variables	Sparse model (L), LR (L), SVM (L), ELM (L)	10	0.96

Table 3. *Cont.*

Reference	Application	Fusion Technique	Fused Data	Mean Accuracy
[64] [a]	Estimation of biomass in grasslands	Simple quadratic combination (L)	2, 15	0.66–0.88
[23]	Plant disease detection	Kohonen self-organizing maps (M)	3, 8	0.95
[84]	Water stress detection	Least squares support vectors machine (M)	3, 8	0.99
[85]	Delineation of water holding capacity zones	ANN (L), MLR (L)	7, 9	0.94–0.97
[86]	Potential of site-specific seeding (potato)	PLSR (L)	2, 9	0.64–0.90
[87]	3D characterization of fruit trees	Pixel level mapping between the images (L)	4, 5	N/A
[88]	Measurements of sprayer boom movements	Summations of normalized measurements (L)	11	N/A
[10] [a,b]	Review: IoT and data fusion for crop disease	N/A	N/A	N/A
[89]	Prediction of wheat yield and protein	Canonical powered partial least-squares (L)	7, 10	0.76–0.94
[69] [a]	Wheat yield prediction	CP-ANN (L), XY-fused networks (L), SKN (L)	2, 7	0.82
[90]	Topsoil clay mapping	PLSR (L) and kNN (L)	7, 9, 13	0.94–0.96
[21]	Fruit detection	CNN (L); scoring system (H)	1, 2	0.84
[37]	3D reconstruction for agriculture phenotyping	Linear interpolation (L)	1, 10	N/A
[29]	Delineation of site-specific management zones	CoKriging (L)	2	0.55–0.77
[91]	Orchard mapping and mobile robot localization	Laser data projection onto the RGB images (L)	1, 5	0.97
[24]	Modelling crop disease severity	2 ANN architectures (L)	10, 15	0.90–0.98
[92]	Tropical soil fertility analysis	SVM (L), PLS (L), least squares modeling (L)	2, 8	0.30–0.95
[93]	Internet of things applied to agriculture	Hydra system (L/M/H)	9, 10, 15	0.93–0.99
[70] [a,b]	Review: data fusion in agricultural systems	N/A	N/A	N/A
[36]	Soil health assessment	PLSR (L)	7, 9	0.78
[94]	Prediction of Soil Texture	SMLR (L), PLSR (L) and PCA (L)	7, 8	0.61–0.88
[95]	Rapid determination of soil class	Outer product analysis (L)	7	0.65
[16]	Navigation of autonomous vehicle	MSPI algorithm with Bayesian estimator (L)	11, 12	N/A
[38] [b]	Detection of cotton plants	Discriminant analysis (M)	2, 7	0.97
[96]	Map-based variable-rate manure application	K-means clustering (L)	2, 9	0.60–0.93
[17]	Navigation of autonomous vehicles	Kalman filter (L)	11, 12	N/A
[97]	Robust tomato recognition for robotic harvesting	Wavelet transform (L)	1	0.93
[98]	Navigation of autonomous vehicle	Self-adaptive PCA, dynamic time warping (L)	1, 11	N/A
[99]	Recognition of wheat spikes	Gram–Schmidt fusion algorithm (L)	1, 2	0.60–0.79

[a] Also explores satellite data. [b] Also explores aerial data.

2.2. Aerial Scale

Studies employing UAVs to solve agricultural problems are growing in number, but they are still outnumbered by proximal and orbital approaches. Most studies are dedicated to crop monitoring and object detection (weed, crops, etc.), although applications, such as phenotyping and water management, are also present. Almost all techniques are based on some kind of digital image (RGB, multispectral, thermal, hyperspectral). Many approaches explore the complementarity of aerial images with proximal (four articles) and orbital (six articles) data. Only eight studies employed the aerial data alone (Table 4).

Table 4. References considered in this study–aerial scale. L, M, and H mean low-, mid-, and high-level data fusion, respectively. The numbers in the fourth column are those adopted in Table 2 for each "fused data" class.

Reference	Application	Fusion Technique	Fused Data	Mean Accuracy
[100]	Root zone soil moisture estimation	NN (M), DRF (M), GBM (M), GLM (M)	2, 11	0.90–0.95
[101]	Gramineae weed detection in rice fields	Haar wavelet transformation (L)	1, 2	0.70–0.85
[65] [a]	Monitoring agricultural terraces	Coregistering and information extraction (L)	5	N/A
[66] [b]	Spectral–temporal response surfaces	Bayesian data imputation (L)	2, 3	0.77–0.83
[102]	Phenotyping of soybean	PLSR (L), SVR (L), ELR (L)	1, 2, 4	0.83–0.90
[39]	Soybean yield prediction	PLSR (M), RF (M), SVR (M), 2 types of DNN (M)	1, 2, 4	0.72
[52] [b]	Crop monitoring	PLSR (M), RF (M), SVR (M), ELR (M)	1, 2	0.60–0.93
[40] [b]	Evapotranspiration estimation	MSDF-ET (L)	1, 2, 4	0.68–0.77
[10] [a,b]	Review: IoT and data fusion for crop disease	N/A	N/A	N/A
[103]	Arid and semi-arid land vegetation monitoring	Decision tree (L/M)	3, 5	0.84–0.89
[41]	Biomass and leaf nitrogen content in sugarcane	PCA and linear regression (L)	2, 5	0.57
[70] [a,b]	Review: data fusion in agricultural systems	N/A	N/A	N/A
[104]	Navigation system for UAV	EKF (L)	11, 12	0.98
[38] [a]	Detection of cotton plants	Discriminant analysis (M)	2	0.97
[71] [b]	Vineyard monitoring	PLSR (M), SVR (M), RFR (M), ELR (M)	2	0.98

[a] Also explores proximal data. [b] Also explores satellite data.

2.3. Orbital Scale

A large portion of the articles employing satellite images aimed to either compensate for data gaps present in a primary data source by fusing it with another source of data (for example, combining optical and SAR images) [6,45,47–49,51,105,106], or increase the spatial resolution of the relatively coarse images collected by satellites with high revisit frequencies [42–44,55,57,58,107–110]. In the latter, the fused results usually inherit the details of the high spatial resolution images and the temporal revisit the frequencies of their counterparts, although the quality of the fused data usually do not match that obtained through actual missions, especially when surface changes are rapid and subtle [72]. As argued by Tao et al. [111], different sensors and image processing algorithms lead inevitably to data with some level of inconsistency, which can make rapid changes difficult to detect.

Landsat and MODIS images and products still dominate, but other satellite constellations, such as Sentinel, Worldview, GeoEye, and others, are being increasingly adopted. Data fusion has been applied to satellite images for quite some time, and well established techniques, such as STARFM and its variants, are still often used, but the interest for machine learning techniques, especially in the form of deep learning models, has been growing consistently. Water management in its several forms (evapotranspiration estimation, mapping of irrigated areas, drought detection, etc.) is by far the most common application. Yield estimation, crop monitoring, land cover classification, and prediction of soil properties are also common applications.

A major challenge associated with the orbital scale is the existence of highly heterogeneous regions with a high degree of fragmentation [109,112]. Solutions to this problem are not trivial and, as stated by Masiza et al. [113], "...successful mapping of a fragmented agricultural landscape is a function of objectively derived datasets, adapted to geographic context, and an informed optimization of mapping algorithms". However, there are cases in which target areas can have sizes smaller than the pixel resolution of the satellite images [53]. In theses cases, pairing the images with images or other types of data obtained at higher resolutions (aerial or proximal) may be the only viable solution. Satellite data were fused together with proximal and aerial data in ten and six studies, respectively (Table 5).

Table 5. References considered in this study–orbital scale. L, M, and H mean low-, mid-, and high-level data fusion, respectively. The numbers in the fourth column are those adopted in Table 2 for each "fused data" class.

Reference	Application	Fusion Technique	Fused Data	Mean Accuracy
[42]	Soil moisture mapping	ESTARFM (L)	2	0.70–0.84
[45]	Crop type mapping	2D and 3D U-Net (L), SegNet (L), RF (L)	2, 6	0.91–0.99
[43]	Estimation of surface soil moisture	ESTARFM (L)	2	0.55–0.92
[26] [a]	Delineation of homogeneous zones	Kriging and other geostatistical tools	2, 9	N/A
[51] [a]	Estimation of crop phenological states	Particle filter scheme (L/M)	2, 6, 10	0.93–0.96
[53]	Evapotranspiration mapping at field scales	STARFM (L)	2	0.92–0.95
[31] [a]	In-field estimation of soil properties	RK (L), PLSR (L)	3, 9	>0.5
[59]	Estimation of wheat grain nitrogen uptake	BK (L)	2, 3	N/A
[44]	Surface soil moisture monitoring	Linear regression analysis and Kriging (L/M)	2, 15	0.51–0.84
[46]	Crop discrimination and classification	Voting system (H)	2, 6	0.96
[9]	Review on multimodality and data fusion in RS	N/A	N/A	N/A
[47]	Crop Mapping	Pixelwise matching (H)	2, 6	0.94
[72]	Review on fusion between MODIS and Landsat	N/A	N/A	N/A
[107]	Mapping crop progress	STARFM (L)	2	0.54–0.86
[66] [b]	Generation of spectral—temporal response	Bayesian data imputation (L)	2, 3	0.77–0.83
[28] [a]	Delineation of management zones	K-means clustering (L)	2, 9, 14	N/A
[114]	Mapping irrigated areas	Decision tree (L)	2	0.67–0.93
[54]	Evapotranspiration mapping	Empirical exploration of band relationships (L)	2, 4	0.20–0.97
[67] [a]	Yield gap attribution in maize	Empirical equations (L)	15	0.37–0.74
[63]	Change detection and biomass estimation in rice	Graph-based data fusion (L)	2	0.17–0.90
[108]	Leaf area index estimation	STARFM (L)	2	0.69–0.76
[55]	Evapotranspiration estimates	STARFM (M)	2	N/A
[115]	Classification of agriculture drought	Optimal weighting of individual indices (M)	2	0.80–0.92
[56]	Mapping daily evapotranspiration	STARFM (L)	2	N/A
[20]	Mapping of cropping cycles	STARFM (L)	2	0.88–0.91
[116]	Evapotranspiration partitioning at field scales	STARFM (L)	2	N/A
[68] [a]	Review: image fusion technology in agriculture	N/A	N/A	N/A
[52] [b]	Crop monitoring	PLSR (M), RF (M), SVR (M), ELR (M)	1, 2, 4	0.60–0.93
[113]	Mapping of smallholder crop farming	XGBoost (L/M and H), RF (H), SVM (H), ANN (H), NB (H)	2, 6	0.96–0.98
[64] [a]	Estimation of biomass in grasslands	Simple quadratic combination (L/M)	2, 15	0.66–0.88
[40] [b]	Evapotranspiration estimation	MSDF-ET (L)	1, 2, 4	0.68–0.77
[117]	Semantic segmentation of land types	Majority rule (H)	2	0.99
[118]	Eucalyptus trees identification	Fuzzy information fusion (L)	2	0.98
[10] [a,b]	Review: IoT and data fusion for crop disease	N/A	N/A	N/A
[69] [a]	Wheat yield prediction	CP-ANN (M), XY-fused networks (M), SKN (M)	2, 7	0.82
[112]	Drought monitoring	RF (M)	2, 15	0.29–0.77
[48]	Crop type classification and mapping	RF (L)	2, 6, 13	0.37–0.94
[119]	Time series data fusion	Environmental data acquisition module	10	N/A
[57]	Evapotranspiration prediction in vineyard	STARFM (L)	2	0.77–0.81
[109]	Daily NDVI product at a 30-m spatial resolution	GKSFM (M)	2	0.88
[49]	Crop classification	Committee of MLPs (L)	2, 6	0.65–0.99
[6]	Multisource classification of remotely sensed data	Bayesian formulation (L)	2, 6	0.74
[111]	Fractional vegetation cover estimation	Data fusion and vegetation growth models (L)	2	0.83–0.95
[120]	Land cover monitoring	FARMA (L)	2, 6	N/A
[121]	Crop ensemble classification	mosaicking (L), classifier majority voting (H)	2	0.82–0.85
[70] [a,b]	Review: data fusion in agricultural systems	N/A	N/A	N/A
[50]	In-season mapping of crop type	Classification tree (M)	2	0.93–0.99

Table 5. *Cont.*

Reference	Application	Fusion Technique	Fused Data	Mean Accuracy
[122]	Building frequent landsat-like imagery	STARFM (L)	2	0.63–0.99
[58]	Evapotranspiration mapping	SADFAET (M)	2	N/A
[123]	Temporal land use mapping	Dynamic decision tree (M)	2	0.86–0.96
[124]	High-resolution leaf area index estimation	STDFA (L)	2	0.98
[125]	Monitoring cotton root rot	ISTDFA (M)	2	0.79–0.97
[110]	Monitoring crop water content	Modified STARFM (L)	2	0.44–0.85
[105]	Soil moisture content estimation	Vector concatenation, followed by ANN (M)	2, 6	0.39–0.93
[126]	Impact of tile drainage on evapotranspiration	STARFM (L)	2	0.23–0.91
[127]	Estimation of leaf area index	CACAO method (L)	2	0.88
[106]	Mapping winter wheat in urban region	SVM (M), RF (M)	2, 6	0.98
[128]	Leaf area index estimation	ESTARFM (L), linear regression model (M)	2	0.37–0.95
[71] [b]	Vineyard monitoring	PLSR (M), SVR (M), RFR (M), ELR (M)	2	0.98

[a] Also explores proximal data. [b] Also explores aerial data.

Another important challenge is the difficulty of obtaining/collecting reference data for validation of the techniques applied. This problem can be particularly difficult if the reference data need to be gathered in-loco. It is also important to consider that, even if reference data can be collected, differences in granularity and the positions of the sample points can make the comparison with the fused data difficult or even unfeasible [112].

These and other challenges related to data fusion applied to satellite data were discussed in depth in [9], so they are not explored in detail here, although some of them are revisited in a more general context in the next section.

3. Discussion

3.1. Comparison of the Results Yielded by Fused and Individual Sources of Data

The last columns of Tables 3–5 show the accuracies reported in each study considered in this article. There are a few important observations to be made before analysing those results. First, the value "N/A" is used in three situations: when accuracy values are not applicable (for example in review articles), when accuracy values are not available, and when the performance of the proposed models is evaluated using either qualitative criteria or metrics that cannot be summarized in a few numbers. Second, some studies report a wide range of accuracy values. This happens when different experimental setups are adopted and also when different variables are considered. This is particularly prevalent in studies dealing with soil variables, in which case, the effectiveness of models can vary significantly from variable to variable. Third, the application of data fusion in agriculture vary significantly, so does the way accuracies are assessed, which partially explains the wide differences observed in the tables. More importantly, even the results reported in studies dealing with similar applications cannot, in general, be directly compared, unless the exact same datasets were used in the experiments. Because of these limitations, the analysis in this section focuses on the effects of data fusion, and not on differences between studies. However, it is worth noting that not all studies explicitly compare results produced with and without data fusion.

The impact of data fusion varies with the type of application. Studies that focused on the fusion of digital images invariably showed improvements with respect to the results obtained using single data sources [18,19,39,84,99,106], although in some cases the improvement may not be substantial enough to justify the capture of additional data [21,71]. The success of data fusion applied to digital images can be linked not only to the complementarities shown by different types of images, but also to the fact that the images can be easily made compatible using simple normalization operations when needed. As a result, proven techniques, such as deep learning, can be applied. In the case of agriculture, image data fusion has been particularly effective at the orbital level, both for artificially increasing the spatial resolution of sources with revisit frequencies [43,53,57,107,108,114] and for com-

pensating cloud cover using the information present in SAR images [45,47,51,113]. In these cases, the improvement can exceed 50% [66].

The use of data fusion for the estimation of soil variables in agricultural areas is more complex, due to the significant differences between the variables that are normally considered. In these cases, data fusion can be effective at improving the estimation of some variables [30,33,35,36], but may fail to produce any improvement for others [31,33,35,95], mostly because those variables do not correlate well with any of the current soil sensors. Some studies also remark that no single data fusion approach works for all soil variables of interest, so a comprehensive variable estimation may require the use of multiple techniques [35,92].

Conversely, the usefulness of data fusion for determining homogeneous zones and producing soil maps, which usually employ soil property measurements, can also be highly dependent on the sensed data, as the inclusion of certain variables can actually decrease accuracy due to the weak link with the properties being used to determine homogeneous areas [90]. This is not a trivial task, and although some studies report encouraging results [28,95], in many cases, further research is needed to better understand how to properly explore the data collected by the sensors [25,27,76]. This has led some authors to remark that the choice of data is more important than the data fusion method employed [90].

Another application in which the use of data fusion has some particular characteristics is the positioning and navigation of autonomous vehicles to be deployed in agricultural areas. In all studies considered in this article, fusing different sensors led to decreased errors, as long as the data fusion methods were calibrated correctly [11–14,16,17,98]. It is worth noting that a variety of different vehicles were considered in those studies, and the error requirements varied from case-to-case. As a result, the level of success of the data fusion process is relative; that is, the same error levels could be considered a success or a failure depending on the vehicle and environment where it will be used.

Some of the studies considered in this article use types of data that are more particular and do not fit a general category, normally in combination with some of the more commonly used variables [24,44,64,67,93]. In general, the performance of the data fusion reported on in those cases compares favorably with single data sources, although some difficulties related with data compatibility have also been reported [82,129].

3.2. Data Fusion Techniques

The variety of data fusion approaches found in the literature indicates that there is no technique that fits all (or even the majority) of the possible applications. Tavares et al. [92] argued that the best data fusion approach depends on the application and attributes considered, and that the selection of an appropriate method should be conducted using independent sample subsets (independent validations) to avoid bias. It is worth noting that, depending on the problem being addressed, experiments sometimes reveal that data fusion is simply not effective [41]. It is important to keep in mind that data fusion is not always the best approach to prevent the practice of attempting to force the results to fit the original hypothesis, which often leads to biased and unrealistic claims.

The use of conventional regression models (proximal and aerial scales) and well-established spatiotemporal methods (orbital scale) is still prevalent. These types of methods have been exhaustively tested and have consistently yielded good results. Because different studies usually employ different datasets, it is not possible to make a direct comparison, but the general trend seems to indicate a lack of progress and a certain level of redundancy between studies using this type of approach. Machine learning techniques have also been applied for some time, but their full potential has yet to be realized. There are a few reasons for this. First, practicing data fusion using machine learning and artificial intelligence models is far from straightforward. Different types of data need to be handled properly in order to make them compatible within the confines of the model. Since many machine learning techniques have poor interpretability [83], it is often difficult to determine how to

achieve an acceptable degree of compatibility. Second, there are a lack of databases that properly cover the entire variability associated to a given application. Although this is a problem for any type of technique [48], the effects of data gaps become more evident when machine learning techniques are applied. Third, machine learning techniques (deep learning in particular) require large amounts of computational power for model training. This problem has become much less damaging in recent years as the computational power availability increases [1]. This, combined with the fact that deep leaning has shown remarkable potential for combining and extracting information from different types of images [21,45,81], will probably lead to this type of technique being increasingly used in the future.

3.3. Data Fusion Level

Even the level of fusion to be adopted is not a straightforward choice. The majority of studies employ low-level fusion, arguably because this is a more straightforward and computationally lighter approach [121]. Some studies seem to indicate that higher fusion levels tend to produce better results [30,77,80], arguing that the poorer results observed when lower level fusion is applied is likely the result of data redundancies arising from complementary information from distinct sensors [30]. Other studies have come to the opposite conclusion, with results favoring the low-level approach [113]. Because few studies have compared the different fusion levels, no definitive conclusions on the matter can be drawn. More comparative studies tackling different fusion levels are needed for a better understanding on how to treat data under different contexts and conditions.

Low-level fusion attempts to directly combine different types of data directly. Since these have different value ranges and formats, some kind of normalization is almost always needed in order to make those different databases compatible [77]. If the only difference is the range of values, normalization tends to be relatively straightforward. On the other hand, if the types of data employed are of different natures (for example, images and meteorological data), the normalization process can be complex or even unfeasible, in which case, the use of higher data fusion levels may be inevitable.

Mid-level fusion can be viewed as a two-step information extraction procedure, as data are first processed in order to generate meaningful features, which are then combined into the final answer. Although in most cases the features are extracted from different types of data (the features themselves are often also different), in some cases, different features are extracted from the same dataset [79]. Calling the combination of features originating from the same dataset "data fusion" is not appropriate in most cases, but sometimes those features represent such different aspects of the data, where the techniques used to combine them are akin those employed in actual data fusion.

As mentioned before, high-level fusion tends to yield solid results, but this approach is the least used among the three levels of fusion. The explanation for this seems to be related to two main factors. First, high-level fusion can be computationally expensive, especially during training [121]. This is particularly true if one or more of the classifiers are based on deep neural networks. Second, choosing the right variables to feed a single model usually is challenging enough, so the process of selecting variables for multiple models may become somewhat impractical. Nevertheless, the use of high-level fusion tends to grow as the availability of computational resources becomes less of a problem, as the research on data fusion matures, and as the characteristics of each type of model/classifier become better understood.

3.4. Differences between Fusion Techniques

Directly comparing the results yielded by different data fusion techniques is difficult for several reasons: the datasets used in the studies are different in terms of data distribution and characteristics of the experimental areas, assessment criteria are diverse, and constraints on the experimental setup usually make the results not generalizable. In addition, results reported in the literature are usually quite similar, but such similitude is often related to

limitations in the representativeness of the datasets, and not to the methods themselves. Thus, instead of comparing the results yielded by different approaches, a qualitative comparison is carried out, with the classes of techniques and applications shown in Table 2 as basis.

Almost all regression methods used in the context of data fusion are of the linear type. They attempt to fit a line that fits the data observed, thus providing a means for predicting values that are not present in the training datasets. In general, this type of technique can be used only when the different sources of data are of the same type or can yield features that are highly compatible. This kind of approach is still commonly used to derive soil properties, as the sensors used in this kind of application generate strings of numbers that can be easily made compatible, but it is seldom adopted in other applications.

STARFM is a data fusion method created specifically to explore the temporal and spatial characteristics of satellite images in order to generate synthetic images that combine the high temporal and spatial resolutions provided by different sensors [130]. This technique and its derivatives [122,131,132] have been successfully applied for more than 15 years, being particular prevalent for fusing Landsat and MODIS images [20,42,53,55,57,58,107,108, 110,116,122,126]. Despite the growth of machine learning and deep learning techniques, the fusion of satellite imagery is still dominated by this kind of approach. There are two main reasons for this. First, these techniques were specifically designed for this type of application and have been continuously perfected, achieving high levels of accuracy and robustness. Second, although machine learning and deep learning techniques can deliver good results, they are susceptible to regularization and overfitting issues that cannot be easily avoided with satellite images [133].

Geostatistical tools, such as kriging, cokriging, and Gaussian anamorphosis, are frequently used for the delineation of homogeneous zones in agricultural areas. This is accomplished by combining different types of variables into a map according to predefined criteria. This type of approach has some competition from machine learning-based models (e.g., K-means clustering), but the latter is usually applied only when digital images are employed in combination with soil measurements [28].

PCA aims at removing redundancies that are often found in data, retaining only the components that provide new information. Although PCA has been used in a wide variety of situations and types of data, in the case of data fusion, its use has been most applied when images are combined with spectroscopy data [77,80], or as a preliminary step to prepare the data to be processed by the actual fusion models [35,41].

In the context of data fusion, Kalman filters are almost exclusively applied to combine inertial and position measurements to aid in the navigation and positioning of autonomous vehicles. This is a well-established approach that is dominating this kind of application. Although some alternative techniques have been proposed [16,98], there is not enough evidence in the literature to favor any given method over Kalman filters.

Machine learning techniques have been applied to a wide variety of problems since (at least) the 1990s. In the case of data fusion in agriculture, techniques, such as fuzzy logic [18], random forest [78,82], support vector machines [83,92,113], k-nearest neighbors [90], and shallow neural networks [24,69,85] have been employed, often showing advantages when digital images were involved, but without ever dominating other strategies. This began to change with the inception of deep neural networks. Deep learning is a branch of machine learning in which the models have deep architectures; that is, the neural networks have many layers with well-defined purposes [134]. They are particularly well suited to deal with digital images and, with the exception of satellite images, this type of technique is quickly becoming the standard for image data fusion [39]. The downside of this success is that deep learning is frequently to unsuitable datasets. Deep learning models are known to require a large number of samples to properly capture the data distributions of the classes to be considered [1]. The minimum number of samples depends on several factors, but in general, the larger the variabilities associated with the problem, the more samples are needed. This is bad news for agricultural applications because the number of factors

that introduce variabilities in image datasets captured in the field is very large [1]. When the dataset lacks variability and training and test samples come from the same dataset (which is almost always the case), accuracies tend to be unrealistically high, and the trained model will almost certainly suffer from overfitting and will tend to fail when fed with new samples. Image augmentation is often used to mitigate this problem, but in many cases this process is not applied correctly, thus aggravating the overfitting problem [4]. The literature has strong evidence that deep learning is indeed the best approach when dealing with all types of digital images [21,22,39], as long some of the pitfalls associated with this type of approach are carefully avoided.

Decision rules [74], majority rules [46,117,121], and model output averaging [30] are all relatively simple techniques applied in high-level fusion with the purpose of combining the information extracted by different models. Because model selection tends to have a greater impact on the fusion effectiveness, the combination of the model outputs is usually carried out using one of these standard approaches.

3.5. Limitations of Current Studies

One problem that plagues most studies based on field data (not only those based on data fusion) is that the results are not realistically weighed against the limitations of the dataset used in the experiments. In particular, in the case of agriculture, it is very unlikely that a single dataset will cover all the variability that can be found in practice [1]. As emphatically put by Øvergaard et al. [89], "how well a model fits a calibration data set does not reveal any information on how good the model performance will be for other data, i.e., for real predictions". In their work, the model was tested using data collected in a different year, under the justification that validating a model with data captured under very similar conditions to those used for training will almost invariably lead to unrealistic results, which unfortunately is usually the case. Indeed, when this approach was applied by Zhou et al. [71], the results obtained for different years were strongly disparate, thus revealing that the data used for training was not representative enough. Veum et al. [36] added that it is difficult to find the ideal sample distribution due to too much homogeneity and presence of extraneous factors. Additionally, even if the process of training, testing and validating is carried out properly, virtually all datasets used in research will have certain limitations, and these always need to be properly considered when analyzing the results. Unfortunately, with some exceptions [33,63,94,123], this is seldom the case. As a result, unrealistic claims are often found in scientific texts, hampering the progress on the subject.

3.6. Types of Data

In many cases, the level of useful information that can be extracted from different types of data are asymmetrical. However, it is important to note that the usefulness of certain types of data in isolation is not an indicator of how it will perform in combination with other data sources. In fact, the most important indicator for including or not a variable is its degree of complementarity with respect to the other ones [113]. In this context, variables that perform poorly in isolation can greatly improve the effectiveness of other variables if they have a high degree of complementarity [19]. In some cases, even if all variables perform poorly individually, they can yield good results when fused together [64]. Conversely, variables that perform well in isolation can match poorly with other variables if they share a high degree of redundancy. Poor synergies [86,90] and high levels of redundancy [102] between different types of data can even lead to the deterioration of the results. However, it is also important to consider that, in some contexts, even if results do not improve with the inclusion of new data types, the robustness to conditions not considered during the development of the models may increase [30]. Moreover, the complementarity of different types of data can be increased by the application of techniques capable of decreasing redundancy (e.g., PCA). In any case, determining if it is worth to include or not a certain type of data are not trivial and usually requires thorough experiments exploring different data combinations [33,36].

Different types of data can have peculiarities that need to be taken into account in order to maximize the performance of the data fusion process. For example, the amount and quality of the information contained in thermal images is highly dependent on the time of day and weather conditions at the time of collection [18,19]. Moreover, some types of data tend to be noisy and prone to outliers, especially if generated from low cost sensors [16], which can potentially introduce error instead of improving the results. In cases like these, applying noise reduction and outlier removal techniques is recommended [33,73]. A comprehensive analysis of several outlier removal techniques in the context of water management is presented by Torres et al. [93]. Despite the problems associated with spurious data, many studies simply ignore this issue, thus limiting the level of detail and quality of the resulting fused information [68].

3.7. Other Issues

In its most basic definition, the term "support" refers to the size or volume associated with each data value [25]. Depending on the context, however, this term may encompass other factors. For example, in the field of geostatistical research, spatial support is related not only to the area and volume associated with the data, but also includes the shape and orientation of the spatial units involved in the measurements [27]. Differences in data representativeness can be very pronounced [95], and if not properly addressed, can cause the data fusion process to fail [25,90]. As stated by Castrignanò et al. [27], "at present, the advantages of using multi-sensor data cannot yet conceal the complexity of the problem encountered in combining disparate spatial data". These authors added that while spatial data fusion has advanced greatly in the last years, the apparent progress can be deceiving as it is often built upon unrealistic assumptions regarding support differences. Given the complexity of the problem, more rigorous validations with more realistic data and assumptions are needed to enable the development of methods with real potential for practical applications.

Overfitting is the phenomenon in which even small noisy variations in the data distribution are captured by the model, making predictions unreliable when applied to independent data [1]. This problem is particularly pervasive with small datasets with limited variability [105]. There are many techniques that can help preventing overfitting, such as data augmentation and cross-validation. Unfortunately, with a few exceptions [34,35,60,77,105], the problems of overfitting is often ignored, thus rendering the results unreliable.

Data gap filling is usually applied to satellite data to compensate for cloud cover and other factors that cause data to be lost or unusable. This can be done either by using data from other satellites or employing data collected at other levels, especially proximal, but the latter does not always improves results [51]. Data gap filling can also refer to the improvement of spatial completeness for a more detailed representation of a given area [42]. Having multiple sensors, even if they are of different types, can also help to deal with situations in which one of the data sources become unavailable for a certain amount of time due to hardware or communication problems [14]. In these cases, even if the effectiveness of the system is diminished, useful information can still be produced until the problem is resolved.

Owing to different time of acquisition, platform modes, and spectral bands of the source images, direct fusion leads to a loss of information and even the failure of fusion [68]. Currently, the most effective way to minimize this problem is by applying image registration, which has, as a main goal, to spatially match different images [19]. This step, which is often pointed out as the most important step in data fusion processes involving images [65], is essential to guarantee that the data present in different types of images are correctly fused into more complete and informative data [18]. This can be challenging, especially if the images being matched are obtained at different scales, at different bands of the spectrum, with different spatial resolutions, or are captured in areas with complex objects and topography. Additionally, the agricultural environment is highly dynamic, with plant canopies changing positions due to the wind or animals moving across the field,

among other elements that may change between image captures [37]. To make matters even more complicated, significant differences between images may exist, especially if these are obtained by different acquisition methods [68]. In many cases, some kind of transformation capable of aligning the points in different images according to a given reference system is required [65].

4. Conclusions

Considerable effort is being made toward finding effective ways at dealing with the wealth of data that are currently being generated in the agricultural environment. Considerable progress has been made, but there are many factors that still prevent techniques that are based on data fusion from being more widely employed in practice. Three of these factors appear to be particularly relevant. First, although large amounts of data are being generated, the complexity of the agricultural environment is such that the databases being generated are not enough to cover all variabilities found in practice. As a result, models and methods proposed in the literature tend to fail under real practical conditions. Second, as the datasets used in research almost invariably have some gaps, the results reported on in scientific articles need to be weighed against those limitations. Unfortunately, this is seldom the case, and often those same results serve as the basis for the development of technologies that will likely fail. Third, even if the technology is robust enough for practical use, there are many technological, economic, political, social, and environmental barriers that prevent its adoption [135].

Despite these hurdles, the potential for growth is still there. In addition to technical advancements, there is some progress towards improving the quantity and quality of the data being collected. The practice of data sharing is steadily growing, with many research groups making datasets available under the findable, accessible, interoperable, and reusable (FAIR) principles [1,136]. When data come from a variety of sources, the representativeness of datasets tend to be much better. Additionally, "citizen science principles" [137,138]—calling for the involvement of individuals outside of the research community, in effort to build datasets—are being applied across different disciplines, with encouraging results. Once a technology truly brings benefits to potential users, adoption barriers tend to weaken. As a result of all of these efforts, the gap between academic research and practical adoption will likely continue to decrease.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Barbedo, J. Deep learning applied to plant pathology: The problem of data representativeness. *Trop. Plant Pathol.* **2022**, *47*, 85–94 . [CrossRef]
2. Kamilaris, A.; Kartakoullis, A.; Prenafeta-Boldú, F.X. A review on the practice of big data analysis in agriculture. *Comput. Electron. Agric.* **2017**, *143*, 23–37. [CrossRef]
3. Coble, K.H.; Mishra, A.K.; Ferrell, S.; Griffin, T. Big Data in Agriculture: A Challenge for the Future. *Appl. Econ. Perspect. Policy* **2018**, *40*, 79–96. [CrossRef]
4. Barbedo, J.G. Factors influencing the use of deep learning for plant disease recognition. *Biosyst. Eng.* **2018**, *172*, 84–91. [CrossRef]
5. Munir, A.; Blasch, E.; Kwon, J.; Kong, J.; Aved, A. Artificial Intelligence and Data Fusion at the Edge. *IEEE Aerosp. Electron. Syst. Mag.* **2021**, *36*, 62–78. [CrossRef]
6. Solberg, A.; Jain, A.; Taxt, T. Multisource classification of remotely sensed data: Fusion of Landsat TM and SAR images. *IEEE Trans. Geosci. Remote Sens.* **1994**, *32*, 768–778. [CrossRef]
7. Bleiholder, J.; Naumann, F. Data Fusion. *ACM Comput. Surv.* **2009**, *41*, 1–40. [CrossRef]
8. Caruccio, L.; Cirillo, S. Incremental Discovery of Imprecise Functional Dependencies. *J. Data Inf. Qual.* **2020**, *12*, 1–25. [CrossRef]

9.   Dalla Mura, M.; Prasad, S.; Pacifici, F.; Gamba, P.; Chanussot, J.; Benediktsson, J.A. Challenges and Opportunities of Multimodality and Data Fusion in Remote Sensing. *Proc. IEEE* **2015**, *103*, 1585–1601. [CrossRef]

10.   Ouhami, M.; Hafiane, A.; Es-Saady, Y.; El Hajji, M.; Canals, R. Computer Vision, IoT and Data Fusion for Crop Disease Detection Using Machine Learning: A Survey and Ongoing Research. *Remote Sens.* **2021**, *13*, 2486. [CrossRef]

11.   Erfani, S.; Jafari, A.; Hajiahmad, A. Comparison of two data fusion methods for localization of wheeled mobile robot in farm conditions. *Artif. Intell. Agric.* **2019**, *1*, 48–55. [CrossRef]

12.   Guo, L.; Zhang, Q. Wireless Data Fusion System for Agricultural Vehicle Positioning. *Biosyst. Eng.* **2005**, *91*, 261–269. [CrossRef]

13.   Han, J.H.; Park, C.H.; Kwon, J.H.; Lee, J.; Kim, T.S.; Jang, Y.Y. Performance Evaluation of Autonomous Driving Control Algorithm for a Crawler-Type Agricultural Vehicle Based on Low-Cost Multi-Sensor Fusion Positioning. *Appl. Sci.* **2020**, *10*, 4667. [CrossRef]

14.   Khot, L.; Tang, L.; Steward, B.; Han, S. Sensor fusion for improving the estimation of roll and pitch for an agricultural sprayer. *Biosyst. Eng.* **2008**, *101*, 13–20. [CrossRef]

15.   Li, Y.; Jia, H.; Qi, J.; Sun, H.; Tian, X.; Liu, H.; Fan, X. An Acquisition Method of Agricultural Equipment Roll Angle Based on Multi-Source Information Fusion. *Sensors* **2020**, *20*, 2082. [CrossRef]

16.   Zaidner, G.; Shapiro, A. A novel data fusion algorithm for low-cost localisation and navigation of autonomous vineyard sprayer robots. *Biosyst. Eng.* **2016**, *146*, 133–148. [CrossRef]

17.   Zhang, Q.; Chen, Q.; Xu, Z.; Zhang, T.; Niu, X. Evaluating the navigation performance of multi-information integration based on low-end inertial sensors for precision agriculture. *Precis. Agric.* **2021**, *22*, 627–646. [CrossRef]

18.   Bulanon, D.; Burks, T.; Alchanatis, V. Image fusion of visible and thermal images for fruit detection. *Biosyst. Eng.* **2009**, *103*, 12–22. [CrossRef]

19.   Gan, H.; Lee, W.; Alchanatis, V.; Ehsani, R.; Schueller, J. Immature green citrus fruit detection using color and thermal images. *Comput. Electron. Agric.* **2018**, *152*, 117–125. [CrossRef]

20.   Li, P.; Lee, S.H.; Hsu, H.Y.; Park, J.S. Nonlinear Fusion of Multispectral Citrus Fruit Image Data with Information Contents. *Sensors* **2017**, *17*, 142. [CrossRef]

21.   Sa, I.; Ge, Z.; Dayoub, F.; Upcroft, B.; Perez, T.; McCool, C. DeepFruits: A Fruit Detection System Using Deep Neural Networks. *Sensors* **2016**, *16*, 1222. [CrossRef] [PubMed]

22.   Cruz, A.C.; Luvisi, A.; De Bellis, L.; Ampatzidis, Y. X-FIDO: An Effective Application for Detecting Olive Quick Decline Syndrome with Deep Learning and Data Fusion. *Front. Plant Sci.* **2017**, *8*, 1741. [CrossRef] [PubMed]

23.   Moshou, D.; Bravo, C.; Oberti, R.; West, J.; Bodria, L.; McCartney, A.; Ramon, H. Plant disease detection based on data fusion of hyper-spectral and multi-spectral fluorescence imaging using Kohonen maps. *Real-Time Imaging* **2005**, *11*, 75–83. [CrossRef]

24.   Shankar, P.; Johnen, A.; Liwicki, M. Data Fusion and Artificial Neural Networks for Modelling Crop Disease Severity. In Proceedings of the 2020 IEEE 23rd International Conference on Information Fusion (FUSION), Rustenburg, South Africa, 6–9 July 2020; pp. 1–8. [CrossRef]

25.   Anastasiou, E.; Castrignanò, A.; Arvanitis, K.; Fountas, S. A multi-source data fusion approach to assess spatial-temporal variability and delineate homogeneous zones: A use case in a table grape vineyard in Greece. *Sci. Total Environ.* **2019**, *684*, 155–163. [CrossRef]

26.   De Benedetto, D.; Castrignano, A.; Diacono, M.; Rinaldi, M.; Ruggieri, S.; Tamborrino, R. Field partition by proximal and remote sensing data fusion. *Biosyst. Eng.* **2013**, *114*, 372–383. [CrossRef]

27.   Castrignanò, A.; Buttafuoco, G.; Quarto, R.; Parisi, D.; Viscarra Rossel, R.; Terribile, F.; Langella, G.; Venezia, A. A geostatistical sensor data fusion approach for delineating homogeneous management zones in Precision Agriculture. *CATENA* **2018**, *167*, 293–304. [CrossRef]

28.   Guerrero, A.; De Neve, S.; Mouazen, A.M. Data fusion approach for map-based variable-rate nitrogen fertilization in barley and wheat. *Soil Tillage Res.* **2021**, *205*, 104789. [CrossRef]

29.   Shaddad, S.; Madrau, S.; Castrignanò, A.; Mouazen, A. Data fusion techniques for delineation of site-specific management zones in a field in UK. *Precis. Agric.* **2016**, *17*, 200–217. [CrossRef]

30.   Afriyie, E.; Verdoodt, A.; Mouazen, A.M. Data fusion of visible near-infrared and mid-infrared spectroscopy for rapid estimation of soil aggregate stability indices. *Comput. Electron. Agric.* **2021**, *187*, 106229. [CrossRef]

31.   Casa, R.; Castaldi, F.; Pascucci, S.; Basso, B.; Pignatti, S. Geophysical and Hyperspectral Data Fusion Techniques for In-Field Estimation of Soil Properties. *Vadose Zone J.* **2013**, *12*, vzj2012.0201. [CrossRef]

32.   Huo, Z.; Tian, J.; Wu, Y.; Ma, F. A Soil Environmental Quality Assessment Model Based on Data Fusion and Its Application in Hebei Province. *Sustainability* **2020**, *12*, 6804. [CrossRef]

33.   Ji, W.; Adamchuk, V.I.; Chen, S.; Mat Su, A.S.; Ismail, A.; Gan, Q.; Shi, Z.; Biswas, A. Simultaneous measurement of multiple soil properties through proximal sensor data fusion: A case study. *Geoderma* **2019**, *341*, 111–128. 2019.01.006. [CrossRef]

34.   La, W.; Sudduth, K.; Kim, H.; Chung, S. Fusion of spectral and electrochemical sensor data for estimating soil macronutrients. *Trans. ASABE* **2016**, *59*, 787–794.

35.   Mahmood, H.S.; Hoogmoed, W.B.; van Henten, E.J. Sensor data fusion to predict multiple soil properties. *Precis. Agric.* **2012**, *13*, 628–645. [CrossRef]

36.   Veum, K.S.; Sudduth, K.A.; Kremer, R.J.; Kitchen, N.R. Sensor data fusion for soil health assessment. *Geoderma* **2017**, *305*, 53–61. [CrossRef]

37. Sampaio, G.S.; Silva, L.A.; Marengoni, M. 3D Reconstruction of Non-Rigid Plants and Sensor Data Fusion for Agriculture Phenotyping. *Sensors* **2021**, *21*, 4115. [CrossRef]
38. Zhang, H.; Lan, Y.; Suh, C.P.C.; Westbrook, J.; Clint Hoffmann, W.; Yang, C.; Huang, Y. Fusion of remotely sensed data from airborne and ground-based sensors to enhance detection of cotton plants. *Comput. Electron. Agric.* **2013**, *93*, 55–59. [CrossRef]
39. Maimaitijiang, M.; Sagan, V.; Sidike, P.; Hartling, S.; Esposito, F.; Fritschi, F.B. Soybean yield prediction from UAV using multimodal data fusion and deep learning. *Remote Sens. Environ.* **2020**, *237*, 111599. [CrossRef]
40. Mokhtari, A.; Ahmadi, A.; Daccache, A.; Drechsler, K. Actual Evapotranspiration from UAV Images: A Multi-Sensor Data Fusion Approach. *Remote Sens.* **2021**, *13*, 2315. [CrossRef]
41. Shendryk, Y.; Sofonia, J.; Garrard, R.; Rist, Y.; Skocaj, D.; Thorburn, P. Fine-scale prediction of biomass and leaf nitrogen content in sugarcane using UAV LiDAR and multispectral imaging. *Int. J. Appl. Earth Obs. Geoinf.* **2020**, *92*, 102177. [CrossRef]
42. Abowarda, A.S.; Bai, L.; Zhang, C.; Long, D.; Li, X.; Huang, Q.; Sun, Z. Generating surface soil moisture at 30 m spatial resolution using both data fusion and machine learning toward better water resources management at the field scale. *Remote Sens. Environ.* **2021**, *255*, 112301. [CrossRef]
43. Bai, L.; Long, D.; Yan, L. Estimation of Surface Soil Moisture With Downscaled Land Surface Temperatures Using a Data Fusion Approach for Heterogeneous Agricultural Land. *Water Resour. Res.* **2019**, *55*, 1105–1128. [CrossRef]
44. Chen, C.F.; Valdez, M.C.; Chang, N.B.; Chang, L.Y.; Yuan, P.Y. Monitoring Spatiotemporal Surface Soil Moisture Variations During Dry Seasons in Central America With Multisensor Cascade Data Fusion. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2014**, *7*, 4340–4355. [CrossRef]
45. Adrian, J.; Sagan, V.; Maimaitijiang, M. Sentinel SAR-optical fusion for crop type mapping using deep learning and Google Earth Engine. *ISPRS J. Photogramm. Remote Sens.* **2021**, *175*, 215–235. [CrossRef]
46. Chen, S.; Useya, J.; Mugiyo, H. Decision-level fusion of Sentinel-1 SAR and Landsat 8 OLI texture features for crop discrimination and classification: Case of Masvingo, Zimbabwe. *Heliyon* **2020**, *6*, e05358. [CrossRef]
47. Forkuor, G.; Conrad, C.; Thiel, M.; Ullmann, T.; Zoungrana, E. Integration of Optical and Synthetic Aperture Radar Imagery for Improving Crop Mapping in Northwestern Benin, West Africa. *Remote Sens.* **2014**, *6*, 6472–6499. [CrossRef]
48. Pott, L.P.; Amado, T.J.C.; Schwalbert, R.A.; Corassa, G.M.; Ciampitti, I.A. Satellite-based data fusion crop type classification and mapping in Rio Grande do Sul, Brazil. *ISPRS J. Photogramm. Remote Sens.* **2021**, *176*, 196–210. 2021.04.015. [CrossRef]
49. Skakun, S.; Kussul, N.; Shelestov, A.Y.; Lavreniuk, M.; Kussul, O. Efficiency Assessment of Multitemporal C-Band Radarsat-2 Intensity and Landsat-8 Surface Reflectance Satellite Imagery for Crop Classification in Ukraine. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2016**, *9*, 3712–3719. [CrossRef]
50. Villa, P.; Stroppiana, D.; Fontanelli, G.; Azar, R.; Brivio, P.A. In-Season Mapping of Crop Type with Optical and X-Band SAR Data: A Classification Tree Approach Using Synoptic Seasonal Features. *Remote Sens.* **2015**, *7*, 12859–12886. [CrossRef]
51. De Bernardis, C.; Vicente-Guijalba, F.; Martinez-Marin, T.; Lopez-Sanchez, J.M. Contribution to Real-Time Estimation of Crop Phenological States in a Dynamical Framework Based on NDVI Time Series: Data Fusion With SAR and Temperature. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2016**, *9*, 3512–3523. [CrossRef]
52. Maimaitijiang, M.; Sagan, V.; Sidike, P.; Daloye, A.M.; Erkbol, H.; Fritschi, F.B. Crop Monitoring Using Satellite/UAV Data Fusion and Machine Learning. *Remote Sens.* **2020**, *12*, 1357. [CrossRef]
53. Cammalleri, C.; Anderson, M.; Gao, F.; Hain, C.; Kustas, W. Mapping daily evapotranspiration at field scales over rainfed and irrigated agricultural areas using remote sensing data fusion. *Agric. For. Meteorol.* **2014**, *186*, 1–11. [CrossRef]
54. Guzinski, R.; Nieto, H.; Sandholt, I.; Karamitilios, G. Modelling High-Resolution Actual Evapotranspiration through Sentinel-2 and Sentinel-3 Data Fusion. *Remote Sens.* **2020**, *12*, 1433. [CrossRef]
55. Knipper, K.; Kustas, W.; Anderson, M.; Alfieri, J.; Prueger, J.; Hain, C.; Gao, F.; Yang, Y.; McKee, L.; Nieto, H.; et al. Evapotranspiration estimates derived using thermal-based satellite remote sensing and data fusion for irrigation management in California vineyards. *Irrig. Sci.* **2019**, *37*, 431–449. [CrossRef]
56. Li, Y.; Huang, C.; Gu, J. Mapping daily evapotranspiration using ASTER and MODIS images based on data fusion over irrigated agricultural areas. In Proceedings of the 2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Fort Worth, TX, USA, 23–28 July 2017; pp. 4394–4397. [CrossRef]
57. Semmens, K.A.; Anderson, M.C.; Kustas, W.P.; Gao, F.; Alfieri, J.G.; McKee, L.; Prueger, J.H.; Hain, C.R.; Cammalleri, C.; Yang, Y.; et al. Monitoring daily evapotranspiration over two California vineyards using Landsat 8 in a multi-sensor data fusion approach. *Remote Sens. Environ.* **2016**, *185*, 155–170. [CrossRef]
58. Wang, T.; Tang, R.; Li, Z.L.; Jiang, Y.; Liu, M.; Niu, L. An Improved Spatio-Temporal Adaptive Data Fusion Algorithm for Evapotranspiration Mapping. *Remote Sens.* **2019**, *11*, 761. [CrossRef]
59. Castaldi, F.; Castrignanò, A.; Casa, R. A data fusion and spatial data analysis approach for the estimation of wheat grain nitrogen uptake from satellite data. *Int. J. Remote Sens.* **2016**, *37*, 4317–4336. [CrossRef]
60. Chlingaryan, A.; Sukkarieh, S.; Whelan, B. Machine learning approaches for crop yield prediction and nitrogen status estimation in precision agriculture: A review. *Comput. Electron. Agric.* **2018**, *151*, 61–69. [CrossRef]
61. Brinkhoff, J.; Dunn, B.W.; Robson, A.J.; Dunn, T.S.; Dehaan, R.L. Modeling Mid-Season Rice Nitrogen Uptake Using Multispectral Satellite Data. *Remote Sens.* **2019**, *11*, 1837. [CrossRef]
62. Nutini, F.; Confalonieri, R.; Crema, A.; Movedi, E.; Paleari, L.; Stavrakoudis, D.; Boschetti, M. An operational workflow to assess rice nutritional status based on satellite imagery and smartphone apps. *Comput. Electron. Agric.* **2018**, *154*, 80–92. [CrossRef]

63. Jimenez-Sierra, D.A.; Benítez-Restrepo, H.D.; Vargas-Cardona, H.D.; Chanussot, J. Graph-Based Data Fusion Applied to: Change Detection and Biomass Estimation in Rice Crops. *Remote Sens.* **2020**, *12*, 2683. [CrossRef]

64. Moeckel, T.; Safari, H.; Reddersen, B.; Fricke, T.; Wachendorf, M. Fusion of Ultrasonic and Spectral Sensor Data for Improving the Estimation of Biomass in Grasslands with Heterogeneous Sward Structure. *Remote Sens.* **2017**, *9*, 98. [CrossRef]

65. Cucchiaro, S.; Fallu, D.J.; Zhang, H.; Walsh, K.; Van Oost, K.; Brown, A.G.; Tarolli, P. Multiplatform-SfM and TLS Data Fusion for Monitoring Agricultural Terraces in Complex Topographic and Landcover Conditions. *Remote Sens.* **2020**, *12*, 1946. [CrossRef]

66. Gevaert, C.M.; Suomalainen, J.; Tang, J.; Kooistra, L. Generation of Spectral–Temporal Response Surfaces by Combining Multispectral Satellite and Hyperspectral UAV Imagery for Precision Agriculture Applications. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2015**, *8*, 3140–3146. [CrossRef]

67. Hu, S.; Mo, X.; Huang, F. Retrieval of photosynthetic capability for yield gap attribution in maize via model-data fusion. *Agric. Water Manag.* **2019**, *226*, 105783. [CrossRef]

68. Li, D.; Song, Z.; Quan, C.; Xu, X.; Liu, C. Recent advances in image fusion technology in agriculture. *Comput. Electron. Agric.* **2021**, *191*, 106491. [CrossRef]

69. Pantazi, X.; Moshou, D.; Alexandridis, T.; Whetton, R.; Mouazen, A. Wheat yield prediction using machine learning and advanced sensing techniques. *Comput. Electron. Agric.* **2016**, *121*, 57–65. [CrossRef]

70. Üstundag, B. Data Fusion in Agricultural Information Systems. In *Agro-Geoinformatics*; Di, L., Üstundag, B., Eds.; Springer: New York, NY, USA, 2021; pp. 103–141.

71. Zhou, X.; Yang, L.; Wang, W.; Chen, B. UAV Data as an Alternative to Field Sampling to Monitor Vineyards Using Machine Learning Based on UAV/Sentinel-2 Data Fusion. *Remote Sens.* **2021**, *13*, 457. [CrossRef]

72. Gao, F.; Hilker, T.; Zhu, X.; Anderson, M.; Masek, J.; Wang, P.; Yang, Y. Fusing Landsat and MODIS Data for Vegetation Monitoring. *IEEE Geosci. Remote Sens. Mag.* **2015**, *3*, 47–60. [CrossRef]

73. Anagnostis, A.; Benos, L.; Tsaopoulos, D.; Tagarakis, A.; Tsolakis, N.; Bochtis, D. Human Activity Recognition through Recurrent Neural Networks for Human–Robot Interaction in Agriculture. *Appl. Sci.* **2021**, *11*, 2188. [CrossRef]

74. Aiello, G.; Giovino, I.; Vallone, M.; Catania, P.; Argento, A. A decision support system based on multisensor data fusion for sustainable greenhouse management. *J. Clean. Prod.* **2018**, *172*, 4057–4065. [CrossRef]

75. Castrignanò, A.; Landrum, C.; Benedetto, D.D. Delineation of Management Zones in Precision Agriculture by Integration of Proximal Sensing with Multivariate Geostatistics. Examples of Sensor Data Fusion. *Agric. Conspec. Sci.* **2015**, *80*, 39–45.

76. Castrignanò, A.; Buttafuoco, G.; Quarto, R.; Vitti, C.; Langella, G.; Terribile, F.; Venezia, A. A Combined Approach of Sensor Data Fusion and Multivariate Geostatistics for Delineation of Homogeneous Zones in an Agricultural Field. *Sensors* **2017**, *17*, 2794. [CrossRef] [PubMed]

77. Comino, F.; Ayora-Cañada, M.; Aranda, V.; Díaz, A.; Domínguez-Vidal, A. Near-infrared spectroscopy and X-ray fluorescence data fusion for olive leaf analysis and crop nutritional status determination. *Talanta* **2018**, *188*, 676–684. [CrossRef] [PubMed]

78. Elsherbiny, O.; Fan, Y.; Zhou, L.; Qiu, Z. Fusion of Feature Selection Methods and Regression Algorithms for Predicting the Canopy Water Content of Rice Based on Hyperspectral Data. *Agriculture* **2021**, *11*, 51. [CrossRef]

79. Guijarro, M.; Riomoros, I.; Pajares, G.; Zitinski, P. Discrete wavelets transform for improving greenness image segmentation in agricultural images. *Comput. Electron. Agric.* **2015**, *118*, 396–407. [CrossRef]

80. Li, F.; Xu, L.; You, T.; Lu, A. Measurement of potentially toxic elements in the soil through NIR, MIR, and XRF spectral data fusion. *Comput. Electron. Agric.* **2021**, *187*, 106257. [CrossRef]

81. Liu, Z.; Zhang, W.; Lin, S.; Quek, T.Q. Heterogeneous Sensor Data Fusion By Deep Multimodal Encoding. *IEEE J. Sel. Top. Signal Process.* **2017**, *11*, 479–491. [CrossRef]

82. López, I.D.; Figueroa, A.; Corrales, J.C. Multi-Label Data Fusion to Support Agricultural Vulnerability Assessments. *IEEE Access* **2021**, *9*, 88313–88326. [CrossRef]

83. Mancipe-Castro, L.; Gutiérrez-Carvajal, R. Prediction of environment variables in precision agriculture using a sparse model as data fusion strategy. *Inf. Process. Agric.* **2021**. [CrossRef]

84. Moshou, D.; Pantazi, X.E.; Kateris, D.; Gravalos, I. Water stress detection based on optical multisensor fusion with a least squares support vector machine classifier. *Biosyst. Eng.* **2014**, *117*, 15–22. [CrossRef]

85. Mouazen, A.M.; Alhwaimel, S.A.; Kuang, B.; Waine, T. Multiple on-line soil sensors and data fusion approach for delineation of water holding capacity zones for site specific irrigation. *Soil Tillage Res.* **2014**, *143*, 95–105. [CrossRef]

86. Munnaf, M.; Haesaert, G.; Van Meirvenne, M.; Mouazen, A. Map-based site-specific seeding of consumption potato production using high-resolution soil and crop data fusion. *Comput. Electron. Agric.* **2020**, *178*, 105752. 2020.105752. [CrossRef]

87. Yandún Narváez, F.J.; Salvo del Pedregal, J.; Prieto, P.A.; Torres-Torriti, M.; Auat Cheein, F.A. LiDAR and thermal images fusion for ground-based 3D characterisation of fruit trees. *Biosyst. Eng.* **2016**, *151*, 479–494. 2016.10.012. [CrossRef]

88. Ooms, D.; Lebeau, F.; Ruter, R.; Destain, M.F. Measurements of the horizontal sprayer boom movements by sensor data fusion. *Comput. Electron. Agric.* **2002**, *33*, 139–162. [CrossRef]

89. Øvergaard, S.I.; Isaksson, T.; Korsaeth, A. Prediction of Wheat Yield and Protein Using Remote Sensors on Plots—Part II: Improving Prediction Ability Using Data Fusion. *J. Infrared Spectrosc.* **2013**, *21*, 133–140. [CrossRef]

90. Piikki, K.; Söderström, M.; Stenberg, B. Sensor data fusion for topsoil clay mapping. *Geoderma* **2013**, *199*, 106–116. [CrossRef]

91. Shalal, N.; Low, T.; McCarthy, C.; Hancock, N. Orchard mapping and mobile robot localisation using on-board camera and laser scanner data fusion—Part A: Tree detection. *Comput. Electron. Agric.* **2015**, *119*, 254–266. [CrossRef]

92. Tavares, T.R.; Molin, J.P.; Javadi, S.H.; Carvalho, H.W.P.D.; Mouazen, A.M. Combined Use of Vis-NIR and XRF Sensors for Tropical Soil Fertility Analysis: Assessing Different Data Fusion Approaches. *Sensors* **2021**, *21*, 148. [CrossRef]

93. Torres, A.B.; da Rocha, A.R.; Coelho da Silva, T.L.; de Souza, J.N.; Gondim, R.S. Multilevel data fusion for the internet of things in smart agriculture. *Comput. Electron. Agric.* **2020**, *171*, 105309. [CrossRef]

94. Wang, S.Q.; Li, W.D.; Li, J.; Liu, X.S. Prediction of soil texture using FT-NIR spectroscopy and PXRF spectrometry with data fusion. *Soil Sci.* **2013**, *178*, 626–638. [CrossRef]

95. Xu, H.; Xu, D.; Chen, S.; Ma, W.; Shi, Z. Rapid Determination of Soil Class Based on Visible-Near Infrared, Mid-Infrared Spectroscopy and Data Fusion. *Remote Sens.* **2020**, *12*, 1512. [CrossRef]

96. Zhang, J.; Guerrero, A.; Mouazen, A.M. Map-based variable-rate manure application in wheat using a data fusion approach. *Soil Tillage Res.* **2021**, *207*, 104846. [CrossRef]

97. Zhao, Y.; Gong, L.; Huang, Y.; Liu, C. Robust Tomato Recognition for Robotic Harvesting Using Feature Images Fusion. *Sensors* **2016**, *16*, 173. [CrossRef] [PubMed]

98. Zhao, W.; Li, T.; Qi, B.; Nie, Q.; Runge, T. Terrain Analytics for Precision Agriculture with Automated Vehicle Sensors and Data Fusion. *Sustainability* **2021**, *13*, 2905. [CrossRef]

99. Zhou, C.; Liang, D.; Yang, X.; Xu, B.; Yang, G. Recognition of Wheat Spike from Field Based Phenotype Platform Using Multi-Sensor Fusion and Improved Maximum Entropy Segmentation Algorithms. *Remote Sens.* **2018**, *10*, 246. [CrossRef]

100. Babaeian, E.; Paheding, S.; Siddique, N.; Devabhaktuni, V.K.; Tuller, M. Estimation of root zone soil moisture from ground and remotely sensed soil information with multisensor data fusion and automated machine learning. *Remote Sens. Environ.* **2021**, *260*, 112434. [CrossRef]

101. Barrero, O.; Perdomo, S. RGB and multispectral UAV image fusion for Gramineae weed detection in rice fields. *Precis. Agric.* **2018**, *19*, 809–822. [CrossRef]

102. Maimaitijiang, M.; Ghulam, A.; Sidike, P.; Hartling, S.; Maimaitiyiming, M.; Peterson, K.; Shavers, E.; Fishman, J.; Peterson, J.; Kadam, S.; et al. Unmanned Aerial System (UAS)-based phenotyping of soybean using multi-sensor data fusion and extreme learning machine. *ISPRS J. Photogramm. Remote Sens.* **2017**, *134*, 43–58. [CrossRef]

103. Sankey, T.; McVay, J.; Swetnam, T.; McClaran, M.; Heilman, P.; Nichols, M. UAV hyperspectral and lidar data and their fusion for arid and semi-arid land vegetation monitoring. *Remote Sens. Ecol. Conserv.* **2018**, *4*, 20–33. [CrossRef]

104. Xiang, H.; Tian, L. Development of a low-cost agricultural remote sensing system based on an autonomous unmanned aerial vehicle (UAV). *Biosyst. Eng.* **2011**, *108*, 174–190. [CrossRef]

105. Yahia, O.; Guida, R.; Iervolino, P. Novel Weight-Based Approach for Soil Moisture Content Estimation via Synthetic Aperture Radar, Multispectral and Thermal Infrared Data Fusion. *Sensors* **2021**, *21*, 3457. [CrossRef] [PubMed]

106. Zhou, T.; Pan, J.; Zhang, P.; Wei, S.; Han, T. Mapping Winter Wheat with Multi-Temporal SAR and Optical Images in an Urban Agricultural Region. *Sensors* **2017**, *17*, 1210. [CrossRef]

107. Gao, F.; Anderson, M.C.; Zhang, X.; Yang, Z.; Alfieri, J.G.; Kustas, W.P.; Mueller, R.; Johnson, D.M.; Prueger, J.H. Toward mapping crop progress at field scales through fusion of Landsat and MODIS imagery. *Remote Sens. Environ.* **2017**, *188*, 9–25. [CrossRef]

108. Kimm, H.; Guan, K.; Jiang, C.; Peng, B.; Gentry, L.F.; Wilkin, S.C.; Wang, S.; Cai, Y.; Bernacchi, C.J.; Peng, J.; et al. Deriving high-spatiotemporal-resolution leaf area index for agroecosystems in the U.S. Corn Belt using Planet Labs CubeSat and STAIR fusion data. *Remote Sens. Environ.* **2020**, *239*, 111615. [CrossRef]

109. Shen, Y.; Shen, G.; Zhai, H.; Yang, C.; Qi, K. A Gaussian Kernel-Based Spatiotemporal Fusion Model for Agricultural Remote Sensing Monitoring. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2021**, *14*, 3533–3545. [CrossRef]

110. Xu, C.; Qu, J.J.; Hao, X.; Cosh, M.H.; Zhu, Z.; Gutenberg, L. Monitoring crop water content for corn and soybean fields through data fusion of MODIS and Landsat measurements in Iowa. *Agric. Water Manag.* **2020**, *227*, 105844. [CrossRef]

111. Tao, G.; Jia, K.; Wei, X.; Xia, M.; Wang, B.; Xie, X.; Jiang, B.; Yao, Y.; Zhang, X. Improving the spatiotemporal fusion accuracy of fractional vegetation cover in agricultural regions by combining vegetation growth models. *Int. J. Appl. Earth Obs. Geoinf.* **2021**, *101*, 102362. [CrossRef]

112. Park, S.; Im, J.; Park, S.; Rhee, J. Drought monitoring using high resolution soil moisture through multi-sensor satellite data fusion over the Korean peninsula. *Agric. For. Meteorol.* **2017**, *237–238*, 257–269. [CrossRef]

113. Masiza, W.; Chirima, J.G.; Hamandawana, H.; Pillay, R. Enhanced mapping of a smallholder crop farming landscape through image fusion and model stacking. *Int. J. Remote Sens.* **2020**, *41*, 8739–8756. [CrossRef]

114. Gumma, M.K.; Thenkabail, P.S.; Hideto, F.; Nelson, A.; Dheeravath, V.; Busia, D.; Rala, A. Mapping Irrigated Areas of Ghana Using Fusion of 30 m and 250 m Resolution Remote-Sensing Data. *Remote Sens.* **2011**, *3*, 816–835. [CrossRef]

115. Kukunuri, A.N.J.; Murugan, D.; Singh, D. Variance based fusion of VCI and TCI for efficient classification of agriculture drought using MODIS data. *Geocarto Int.* **2020**. [CrossRef]

116. Li, Y.; Huang, C.; Kustas, W.P.; Nieto, H.; Sun, L.; Hou, J. Evapotranspiration Partitioning at Field Scales Using TSEB and Multi-Satellite Data Fusion in The Middle Reaches of Heihe River Basin, Northwest China. *Remote Sens.* **2020**, *12*, 3223. [CrossRef]

117. De Oliveira, J.P.; Costa, M.G.F.; Filho, C. Methodology of Data Fusion Using Deep Learning for Semantic Segmentation of Land Types in the Amazon. *IEEE Access* **2020**, *8*, 187864–187875. [CrossRef]

118. Oliveira, D.; Martins, L.; Mora, A.; Damásio, C.; Caetano, M.; Fonseca, J.; Ribeiro, R.A. Data fusion approach for eucalyptus trees identification. *Int. J. Remote Sens.* **2021**, *42*, 4087–4109. [CrossRef]

119. Samourkasidis, A.; Athanasiadis, I.N. A semantic approach for timeseries data fusion. *Comput. Electron. Agric.* **2020**, *169*, 105171. [CrossRef]

120. Thomas, N.; Neigh, C.S.R.; Carroll, M.L.; McCarty, J.L.; Bunting, P. Fusion Approach for Remotely-Sensed Mapping of Agriculture (FARMA): A Scalable Open Source Method for Land Cover Monitoring Using Data Fusion. *Remote Sens.* **2020**, *12*, 3459. [CrossRef]

121. Useya, J.; Chen, S. Comparative Performance Evaluation of Pixel-Level and Decision-Level Data Fusion of Landsat 8 OLI, Landsat 7 ETM+ and Sentinel-2 MSI for Crop Ensemble Classification. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2018**, *11*, 4441–4451. [CrossRef]

122. Wang, P.; Gao, F.; Masek, J.G. Operational Data Fusion Framework for Building Frequent Landsat-Like Imagery. *IEEE Trans. Geosci. Remote Sens.* **2014**, *52*, 7353–7365. [CrossRef]

123. Wang, L.; Wang, J.; Qin, F. Feature Fusion Approach for Temporal Land Use Mapping in Complex Agricultural Areas. *Remote Sens.* **2021**, *13*, 2517. [CrossRef]

124. Wu, M.; Wu, C.; Huang, W.; Niu, Z.; Wang, C. High-resolution Leaf Area Index estimation from synthetic Landsat data generated by a spatial and temporal data fusion model. *Comput. Electron. Agric.* **2015**, *115*, 1–11. [CrossRef]

125. Wu, M.; Yang, C.; Song, X.; Hoffmann, W.C.; Huang, W.; Niu, Z.; Wang, C.; Li, W.; Yu, B. Monitoring cotton root rot by synthetic Sentinel-2 NDVI time series using improved spatial and temporal data fusion. *Sci. Rep.* **2018**, *8*, 2016. [CrossRef] [PubMed]

126. Yang, Y.; Anderson, M.; Gao, F.; Hain, C.; Kustas, W.; Meyers, T.; Crow, W.; Finocchiaro, R.; Otkin, J.; Sun, L.; et al. Impact of Tile Drainage on Evapotranspiration in South Dakota, USA, Based on High Spatiotemporal Resolution Evapotranspiration Time Series From a Multisatellite Data Fusion System. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2017**, *10*, 2550–2564. [CrossRef]

127. Yin, G.; Verger, A.; Qu, Y.; Zhao, W.; Xu, B.; Zeng, Y.; Liu, K.; Li, J.; Liu, Q. Retrieval of High Spatiotemporal Resolution Leaf Area Index with Gaussian Processes, Wireless Sensor Network, and Satellite Data Fusion. *Remote Sens.* **2019**, *11*, 244. [CrossRef]

128. Zhou, X.; Wang, P.; Tansey, K.; Zhang, S.; Li, H.; Tian, H. Reconstruction of time series leaf area index for improving wheat yield estimates at field scales by fusion of Sentinel-2, -3 and MODIS imagery. *Comput. Electron. Agric.* **2020**, *177*, 105692. [CrossRef]

129. Da Costa Bezerra, S.F.; Filho, A.S.M.; Delicato, F.C.; da Rocha, A.R. Processing Complex Events in Fog-Based Internet of Things Systems for Smart Agriculture. *Sensors* **2021**, *21*, 7226. [CrossRef]

130. Gao, F.; Masek, J.; Schwaller, M.; Hall, F. On the blending of the Landsat and MODIS surface reflectance: predicting daily Landsat surface reflectance. *IEEE Trans. Geosci. Remote Sens.* **2006**, *44*, 2207–2218. [CrossRef]

131. Hilker, T.; Wulder, M.A.; Coops, N.C.; Seitz, N.; White, J.C.; Gao, F.; Masek, J.G.; Stenhouse, G. Generation of dense time series synthetic Landsat data through data blending with MODIS using a spatial and temporal adaptive reflectance fusion model. *Remote Sens. Environ.* **2009**, *113*, 1988–1999. [CrossRef]

132. Zhu, X.; Chen, J.; Gao, F.; Chen, X.; Masek, J.G. An enhanced spatial and temporal adaptive reflectance fusion model for complex heterogeneous regions. *Remote Sens. Environ.* **2010**, *114*, 2610–2623. [CrossRef]

133. Gu, Y.; Wylie, B.K.; Boyte, S.P.; Picotte, J.; Howard, D.M.; Smith, K.; Nelson, K.J. An Optimal Sample Data Usage Strategy to Minimize Overfitting and Underfitting Effects in Regression Tree Models Based on Remotely-Sensed Data. *Remote Sens.* **2016**, *8*, 943. [CrossRef]

134. Kamilaris, A.; Prenafeta-Boldú, F.X. Deep learning in agriculture: A survey. *Comput. Electron. Agric.* **2018**, *147*, 70–90. [CrossRef]

135. Da Silveira, F.; Lermen, F.H.; Amaral, F.G. An overview of agriculture 4.0 development: Systematic review of descriptions, technologies, barriers, advantages, and disadvantages. *Comput. Electron. Agric.* **2021**, *189*, 106405. 2021.106405. [CrossRef]

136. Wilkinson, M.D.; Dumontier, M.; Aalbersberg, I.J.; Appleton, G.; Axton, M.; Baak, A.; Blomberg, N.; Boiten, J.W.; da Silva Santos, L.B.; Bourne, P.E.; et al. The FAIR Guiding Principles for scientific data management and stewardship. *Sci. Data* **2016**. [CrossRef] [PubMed]

137. Irwin, A. *Citizen Science: A Study of People, Expertise and Sustainable Development*, 1st ed.; Routledge Press: Oxfordshire, UK, 2002.

138. Silvertown, J. A new dawn for citizen science. *Trends Ecol. Evol.* **2009**, *24*, 467–471. [CrossRef]