



**HAL**  
open science

# Investigation of Common Big Data Analytics and Decision-Making Requirements Across Diverse Precision Agriculture and Livestock Farming Use Cases

Spiros Mouzakis, Giannis Tsapelas, Sotiris Pelekis, Simos Ntanopoulos,  
Dimitris Askounis, Sjoukje Osinga, Ioannis N. Athanasiadis

## ► To cite this version:

Spiros Mouzakis, Giannis Tsapelas, Sotiris Pelekis, Simos Ntanopoulos, Dimitris Askounis, et al.. Investigation of Common Big Data Analytics and Decision-Making Requirements Across Diverse Precision Agriculture and Livestock Farming Use Cases. 13th International Symposium on Environmental Software Systems (ISESS), Feb 2020, Wageningen, Netherlands. pp.139-150, 10.1007/978-3-030-39815-6\_14 . hal-03361883

**HAL Id: hal-03361883**

**<https://inria.hal.science/hal-03361883v1>**

Submitted on 1 Oct 2021

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

# Investigation of Common Big Data Analytics and Decision-Making Requirements across Diverse Precision Agriculture and Livestock Farming Use Cases

Spiros Mouzakis<sup>1</sup>, Giannis Tsapelas<sup>1</sup>, Sotiris Pelekis<sup>1</sup>, Simos Ntanopoulos<sup>1</sup>, Dimitris Askounis<sup>1</sup>, Sjoukje Osinga<sup>2</sup>, Ioannis N. Athanasiadis<sup>2</sup> [0000-0003-2764-0078]

<sup>1</sup> Decision Support Systems Laboratory, National Technical University of Athens, 15780, Greece

<sup>2</sup> Wageningen University, 6700 EW, Wageningen, The Netherlands

**Abstract.** The purpose of this paper is to present the investigation of common requirements and needs of users across a diverse set of precision agriculture and livestock farming use cases that was based on a series of interviews with experts and farmers. The requirements were based on nine interviews that were conducted in order to identify common requirements and challenges in terms of data collection and management, Big Data technologies, High Performance Computing infrastructure and decision making. The common requirements that derived from the interviews and user requirement analysis per use case can serve as basis for identifying functional and non-functional requirements of a technological solution of high re-usability, interoperability, adaptability and overall efficiency in terms of addressing common needs for precision agriculture and livestock farming.

**Keywords:** precision agriculture, livestock farming, big data analytics, decision-making, user requirements

## 1 Introduction

The recent technological advancements in Big Data, Artificial Intelligence (AI), High Performance Computing (HPC), Cloud Services and Internet of Things (IoT) have the potential to enable farmers to overcome long-standing challenges in exploiting the vast amount of data that can be collected, in order to increase efficiency and productivity while reducing the initial farm input costs [1, 2, 3]. Big data and AI enable novel precision agriculture opportunities that allow the performing of queries and analytics on a distributed and diverse set of collected data (from IoT devices, images, video, satellite data, etc.) that may lead to better and faster predictions and vital insights for farming decisions. IoT sensors on fields and crops [4, 5, 6] can provide a significant amount of

information, critical for the decision-making process such as soil conditions, high-fidelity weather conditions [7, 8], fertilising requirements, water availability and pest infestations. In addition, aerial images captured by unmanned aerial vehicles [9], or drones, which can patrol fields, can provide early warnings related to potential problems, such as diseases or deviations from expected growth rates; or offer indicators of crop ripeness and quality. Individual plants may also be monitored for nutrients and growth rates. Satellites [10] can facilitate detection of relevant changes in field with captured satellite imagery, identification of crop threats such as nutrients deficiency or insect damage, while GPS units on tractors can help determine optimal usage of heavy equipment and precise management of field operations. At the same time, livestock production management exploits technology to quantitatively measure the behaviour, health and performance of animals, including real-time monitoring of reproduction, health and welfare of livestock and the corresponding environmental impact [11, 12]. The data sources utilised in livestock management include amongst others on-line sound and video observations, feeding intake, drinking behaviour data, data from sensors on the animals and data from milking robots. There is an increasing literature of individual use-cases on precision agriculture and livestock farming applications, nevertheless the identification of common requirements and challenges across different use cases is currently missing.

The purpose of this paper is to present the investigation of common requirements and needs of users across a diverse set of precision agriculture and livestock farming use cases that was based on a series of interviews with experts and farmers.

## **2 Methods and Approach**

### **2.1 Overall Approach**

For the identification of the requirements a close collaboration with twelve agri-food industry stakeholders took place between January and June 2019 and consisted of three phases:

- a) Structured interviews with the stakeholders describing the current situation, the expectations for an ideal, future situation and the challenges to be addressed in between
- b) Using the initial current and future situations from the first phase, nine Usage Scenarios were created in a collaborative manner with data engineers, analysts and HPC experts
- c) A list of common user stories was distilled from all the usage scenarios and were ranked based on their foreseen business impact and technological complexity

The requirement investigation process took into consideration the user-varying dynamics in the precision agriculture and livestock farming scenarios, that is the different sectors, expectations, backgrounds and interests of users involved in different tasks such as data collection, data cleaning, modelling, decision making, and application of the outcome of the tasks for particular purposes.

Each user task is distinguished by several aspects: the data needed for this task, the skills required to perform the task, the timing of this task, and so on. People who assume the role of performing the task can be end users (e.g. farmers, farmer cooperatives), intermediaries (e.g. extension officers, veterinarians, seed providers), companies (e.g. planning institutions), data professionals (e.g. technical assistants, data analysts / scientists), and so on. Nonetheless, the user is not necessarily a person, the term could also describe a device (like a sensor or a satellite), an application program, an HPC cloud function, or any agent performing a task.

In order to avoid confusion around the term ‘user’, the usage of ‘roles’ was preferred in the analysis. A role is the particular function a user fulfils when carrying out a particular task, such as the role of data collection or the role of data cleaning. Each role has a body fulfilling that role. Several roles may be combined by the same body of fulfilment. The final user of the use case is not necessarily the same as the end user or stakeholder who benefits from the solution that the use case provides. A use case may end at the point where an intermediary is required to bring the outcome to a farmer in order to help the farmer solve a problem. The farmer himself/ herself may not be within the scope of the technological solution, but he is a stakeholder nevertheless. To avoid confusion on that, we use the term ‘stakeholder’ exclusively for the one who ultimately benefits from use case solutions developed in the context of the ICT solution but who is not necessarily part of the scope of the use case, and hence not necessarily himself/ herself a user of the technological solution.

## **2.2 Interviews**

Initially, a set of individual online and physical one-to-one interviews were conducted between March and April 2019. During the interviews, workflow diagrams were identified and drawn together with the interviewees, based on their descriptions. When more than one situation could apply for “current situation” or “future situation”, they were all included in the interview. Hence, each interview resulted to a set of workflow diagrams that consisted of at least two, but often more diagrams.

Following, each diagram was discussed with respect to the roles involved in every step of the workflow. These roles were placed in the diagrams, labelled and described as accurately as possible at that time.

Finally, the stakeholders for each use case were identified as well, with elaborated descriptions of how they can be distinguished from one another, and how they inform themselves and communicate within and outside their peer networks. This in-depth description of the stakeholders and the ways they operate guided the process of deriving the requirements that best match their needs and expectations.

## **2.3 User Scenarios Co-Design Process**

Following the interviews, nine usage scenarios were co-designed through the collaboration of agrifood-industry stakeholders and technical staff (data engineers, analysts, HPC experts). The usage scenarios are descriptions of hypothetical real-world examples. They are essential in order to highlight the interactions of people and organisations

with a system. They contain detailed references to the steps, events, and/or actions which occur during these interactions. More specifically, the scenarios in this document attempt to describe, exhaustively and on a high-level, all the possible ways the users will interact with the proposed technical solutions and accurately depict their critical business actions. Hence, the scenarios show how the various users wish to improve their workflow with available state-of-the-art solutions. Critical properties were specified by the information provided in the scenarios regarding e.g. the type of device the users will access the developed application from, locations where data can be stored, privacy and security constraints, existing solutions. Agrifood stakeholders outlined their business needs and challenges while the technical partners reviewed the scenarios and updated them with proposed state-of-the-art ICT solutions (infrastructure, data workflows, software tools and algorithms). In order for the reader to develop a better understanding of the concept, a summary of an indicative user scenario, as it was designed by the stakeholders of a fish farm, is provided in section 2.4.

## 2.4 An Example of a User Scenario

Datafish is an SME that develops precision livestock farming ICT solutions for aquaculture by deploying IoT sensors and AI. Datafish wants to develop a platform solution that will aid Cretafish, a fish-farming company in Crete, in monitoring fish behavior in fish tanks and optimizing the decision-making in fish diseases, fish feeding and environmental impacts of the fish farm. Cretafish is currently adjusting the amount of food to be thrown in the cages by observing the fish at the surface level during the feedings. Datafish will incorporate near real-time processing of aerial drone and satellite images as well as sensor collected data such as water temperature, salinity, maximum chlorophyll index (MCI) etc. Datafish will also include several components that support multiple data operations and visualisations, machine learning modeling capabilities and reporting services, in order to enable the generation of valuable information to be leveraged by Cretafish.

### **Data Operations and Simple Visualisations Component.**

- *Data Management.* This module automatically receives relevant data available data to the platform (e.g. weather data). Additionally, it allows for an authorised Datafish employee to gain access to the datasets that are generated by sensors and cameras. She/ He can regularly clean and upload them to the solution's platform where they are securely stored in a private cloud.
- *Data Exploration.* This module permits the exploration of each dataset e.g. aerial images for observing fish behaviour, temperature and O2 levels to monitor fish health.
- *Simple Visualisations.* This module provides interactive graphs such as bar-charts that combine generated feeding patterns with weather data in order to assess their correlation.

**Experiment/Analytics Environment.** An authorised Datafish data analyst, uses this component to gain access to a development environment built to support her/ his favorite programming language -e.g. Python- as well as her/ his most often-used machine learning algorithms and datasets. She/ He writes the code in order to achieve:

- Fish feeding optimisation
- Algae bloom prediction
- Fish disease prediction

**Dashboards/Advanced Visualisations Component.** The component includes a real-time dynamic heatmap of the fish tank and advanced graphs showing the amount of food, the recommended optimal feeding plan and an overview of adjustable fish parameters. It also provides historical versions of the heatmap for on-the-go comparisons, as well as a camera feed with zoom capabilities that will enable the inspection of individual and group fish behavior and will facilitate disease detection.

**Reporting Component.** The component enables Datafish to automate the reporting process so that Cretafish receives daily updates that include parts of the Dashboards and a summary report on each cage. Cretafish can also view the dashboards in real-time when feeding is taking place so to optimise future decision-making.

**Alerts Component.** Cretafish uses this component in order to receive disease warnings based on anomaly detection techniques and alerts relevant to the duration and amount of feeding.

**Model Template Component.** After approximately six months of application, this system becomes a model template ready to be used by other fish farms through this component.

## 2.5 Requirements Elicitation and Ranking

A list of user stories was elicited from the case-specific usage scenarios in the form of “As a <Role>”, “I want to <user requirement>”, “So that <benefit>” and grouped together based on the use cases. Thereafter the stakeholders were asked to rank the derived requirements based on:

- a) Business value: The business value was asked to be provided by the agrifood stakeholders, in a qualitative manner using a scale between 1-5 points, by estimating the importance and foreseen business impact of this requirement for the overall process.
- b) Technical Complexity: The Technical complexity was calculated by the technical partners through voting, on a scale for 1 to 5, with 1 meaning “least complex” and 5 “most complex”

This methodology poses as an extension to an approach proposed by Lant [13] which calculates a score for every user story, against two different attributes, that of business value and urgency. Through the methodology presented above and by qualitatively examining the requirements and the dependencies among some of them, we concluded on a final set of common requirements for Big Data analysis and decision-making across the nine different use cases of precision agriculture and livestock farming practice.

### **3 Interview Results per Use Case**

#### **3.1 Organic Soya Yield and Protein-Content Prediction**

This use case involves the prediction of yield and protein-content maps based on satellite imagery and additional information concerning electromagnetic soil scans, drone images and sensory data. Time-series of satellite images are very indicative of the relative yield and protein content on the field, i.e. they can pinpoint the areas in which soybean grows better or worse. By knowing the absolute values of yield and protein content on the whole farm, these traits can be reverse-engineered in these specific areas of the field and derive the corresponding maps.

During the interview, the need for data fusion was recognised, as data will be acquired from heterogeneous sources (EC probes, satellites, drones) at different times. Although monitoring in one farm may not be computationally intensive, monitoring the soybean production at field-scale throughout Europe requires huge processing power and smart algorithms, optimised for parallel execution. The size of the dataset and the classification algorithm that has to be trained and executed, require efficient data management and strong processing power through HPC infrastructure and Big Data solutions.

#### **3.2 Climate-Smart Predictive Models for Viticulture**

The purpose of this use case is to support complex, highly non-linear, models for vine and grape growth with respect to the extreme number of variables (data types) that have been shown to affect the quality and quantity of the produced yields. Such crop models could estimate vine and grape growth and crop yield at larger scales, with spatial sources of information on soil, water, land use, and other factors. In this way, much larger predictions of yield could be achieved across regional scales.

During the interview it was clarified that this use case involves large matrix operations, which are memory-intensive computations. The limited amount of computer memory available in a non-HPC enabled deployment does not allow for an efficient parallelisation of the data processes entailed in the solution.

#### **3.3 Climate Services for Organic Fruit Production**

Integration and comparison of fruit bud development estimation models with temperature and air humidity forecasts and other ancillary data can be used for risk probability mapping in order to establish an early warning system that can help farms to minimise damage effects through protective methods for frost and hail. This use case focuses on climate predictors that are correlated with either frost or hail occurrences and then can be used for planning risk prevention operations.

During the interview the need for examining the integration of a frost and hail early warning system as a climate service into a decision support system for horticultural and fruit-tree farmers was presented. This service is based on exploring and analyzing the

best probabilistic predictions of these extreme climate events for site-specific spatial scales. Observational and simulated climate variables need to be integrated with crop modelling approaches to estimate the probable risk. This requires exploration and comparison of different methods together with forecast quality assessment of the predictions, in which synchronous observed and predicted values are compared.

### 3.4 Autonomous Robotic Systems within Arable Arametworks

This case considers the provision of autonomous robotic systems within an arable farm. Dictated by the weather, farming tasks have often to be carried out within a short time window. Consequently, equipment has increased in size to complete the work rapidly. One alternative solution is for farmers to manage fleets of smaller, autonomous vehicles and carry out the tasks as required.

During the interview, the to-be situation focused on field robots obtaining sensory data for soil chemical analysis (regularly/monthly) and hyperspectral imaging (HSI) for determining soil and crop conditions. Such data can be used for real time object level plant identification, individual plant harvest readiness assessment (in near harvest periods) and plant-level automated harvesting (which is currently labour intensive). The challenge in this use case lies in the precise processing of sensory data not only for identifying plant, weed and arable land readiness for harvesting, but also for activating nearby actuators distributed across a number of vehicles.

### 3.5 Optimizing Computations for Crop Yield Forecasting

Crop yield monitoring can be used as a tool for agricultural monitoring (e.g. early warning & anomaly detection), index-based insurance (index estimates) and farmer advisory services to facilitate precision agriculture and timely identify in-field phenomics by helping to provide greater yields and contributing to better food security.

The interview process highlighted that current computation loads over a single server have been reduced to meet hardware limitations. Additionally, more detailed weather forecasts from ECWMF together with parcel-specific data, and data processed for Sentinel Satellite Imagery may allow to predict crop yields at parcel level. Hence, one of the main challenges is distributing the computational load over several computational resources and exploring the potential of machine learning algorithms to forecast yields at parcel level.

### 3.6 Pig Weighing Optimisation

This use case has three main goals: (1) To estimate the mean and standard deviation of the live weight of grower/finisher pigs in a pen based on video images; (2) To track the weight of individual pigs in a pen based on video images; (3) To incorporate the growth curve estimated by the Convolutional Neural Networks (CNNs) in previously developed models for early warning of diarrhea. Currently there is no video-based weight estimation available and an *eye-balling* estimate is used on common farms due to manual weighing on big farms being a laborious task.



The interview process revealed that the training of the above described CNNs with the already available big data is an inherently immensely computationally demanding task which cannot be handled by limited computational power and memory infrastructures. Moreover, image and video processing and analysis is a task that can highly benefit from HPCs. The sensors that are installed are generating large amounts of video and signal data and this will be an even bigger issue as additional sensors might be installed in the future.

### **3.7 Sustainable Pig Production**

This use case focuses on improving pig health and welfare, works on fulfilling the potential of each pig through its life and on increasing the quality of the end-product for the market and the consumers. This requires the usage and data fusion of various data sources coming from multiple on-farm sensors and software systems, image analysis, management data and slaughterhouse records.

During the interview, the need for exploring techniques for data fusion was identified, particularly for data of different size and sampling frequencies. Additional needs include high-throughput processing of big data with multivariate algorithms and advanced machine learning, as deep learning, to automatically detect data anomalies. The outcomes can include warnings (alerts) of problems with health, welfare or productivity, development of longitudinal trends, and data on individual pigs and prediction of end-product quality.

### **3.8 Open Sea Fishing**

The goal of this use case is to achieve higher sustainability of a fishing fleet, rebuild overfished stocks and prevent overfishing. First, the integration of data from the entire fleet's electronic logbooks is necessary and requires a series of improvements of the on-board database systems of commercial fishing vessels. Additionally, the collection, storage and processing of on-board sensor data is required, together with a visual-based processing of catch imagery deriving from RGB cameras for fish-selection purposes. During the interview it was revealed that the preprocessing involved is computationally intensive and requiring a scalable infrastructure that offers parallel processing. Moreover, a multivariate analysis model that integrates all available sensor and price data as well as an artificial neural network to find optimal operational parameters for minimal costs is required and is highly computationally intensive.

### **3.9 Aquaculture Monitoring and Feeding Optimisation**

The purpose of this use case is to investigate fish behaviour on a deeper level. To do this, methods like segmentation and region proposal, object tracking, video analysis and machine learning are used to analyse water movements from colour, to detect problems in nets and cages, and to determine fish positions. This information will be combined with other data such as weather information and sensor measurements (mainly related to oxygen and current speed) in order to develop an efficient feed management

system that can help companies to make optimal use of the food, reduce costs and also reduce the impact on the environment.

The interview revealed that data processing needs to be performed within short time frame. Being able to process data fast, and extract insights are the big challenges of this use case, demanding high throughput, computational intensity and short turnaround times.

#### 4 Derived Common Requirements from All Use Cases

For each use case presented above, a set of usage scenarios was created along with the corresponding set of user stories. Thereafter all user stories were analysed and classified in eight categories (data management, data storage, data exploration, data analysis, data process, visualisations, support, alerts). Common requirements across all use cases and user stories were identified and are presented in the following Table 1.

**Table 1.** Common requirements from all use cases investigated

Category	Requirement
Data Management	Guarantee data, model & outputs integrity through the entire data value chain in order to develop a trustful relationship between stakeholders (e.g. farmer & buyer)
Data Management	Be able to upload new datasets, with different schemas, in a common data repository, from different sources
Data Management	Support automated check-in of sensor-collected data
Data Management	Daily transfer & access to daily updated data (e.g. parcel specific weather forecasts, satellite imagery data, logbooks, satellite VMS data etc.)
Data Management	HPC infrastructure for near real-time image processing in order to allow near real-time notifications
Data Management	Ability to share results and visualisations with other users to increase collaboration opportunities and efficiency
Data Management	Ability to store as template a whole project, or smaller workflows of it, ready to be trained with new input datasets
Data Management	Access to parcel-specific soil data, crop parameters, crop calendars, historical weather archives, regional statistics & historic yield data (parcel-specific N/A but desired) in order to correlate them to conduct analysis and feed them to crop models
Data Management	Expose data and results through APIs for increased interoperability with other systems
Data Storage	A private cloud storage is required with fine-grained access control
Data Storage	Encryption of data on the private cloud storage

Data Exploration	Be able to execute user-friendly queries across diverse and distributed datasets of different types and store their results. Desired subsets of whole datasets can be saved & used in latter data analytics.
Data Exploration	Perform queries on raw, cured or aggregated data, either in batch mode or in real time
Data Analysis	Create and train machine learning models using available datasets to assist in the decision-making process
Data Analysis	Reuse models stored in the personal account in different cases in order to save time and use previous experience in new cases
Data Analysis	Access to predictions about quantity to be produced (crop yield/ fish catch prediction, etc.) on a daily basis in order to improve logistics, increase profits, optimise negotiations & remain reliable to partners
Data Analysis	Provide support / integration for a variety of statistical and Big Data tools, like R and MXNet, Tensorflow or Spark frameworks
Data Analysis	Upload, train & deploy locally created models to the solution despite the different ML tools used for their development (e.g. by supporting ONNX) for easier comprehension of current situation, predictions and recommended actions
Data Analysis	Ability to create models in different programming language (R, python, Scala, Spark MLlib) depending on the expertise / preference of each user. Do not lock-in the use case in a specific programming language
Data Analysis	Be able to validate using reference historic data from previous projects to recalibrate or fine tune the system to each user's case in order to leverage existing data and assist the quick training of the models.
Data Analysis	Ability to enable semi-automatic generation of metadata, in order to save time from manual data input
Data Analysis	Support automated detection of anomalies in output from the simulation models that run (executed efficiently on HPC clusters)
Data Analysis	Apply classification algorithms on both images and other input variables
Data Analysis	Train & optimise Machine Learning and Deep Learning models and evaluate them using different parameters in a parallel way (tool examples: Horvod, Tensorflow)
Data Process	Support indications of the parts of the code that are memory or processing intensive and may require parallelisation, i.e. using HPC or Apache Spark functionality
Data Process	Process video & image data and extract features within short time windows (near real-time responses)
Data Process	Be able to manage the resources being used by the experiments [e.g. Workflow management dashboard (edit, pause, delete, view consumed sources)] in order to be aware and be able to manage resources used respectively to the workload of each task
Data Process	Make use of the best distributed frameworks, the best available software libraries for scientific computation (such as MKL or openBLAS)
Data Process	Smart distribution of data and parallelisation of tasks based on multi-objective optimisations (e.g. by origin / destination of the data,

	available resources) and best parameter setting for the productivity model (e.g. according to environmental, regional and weather conditions)
Data Process	Interpolate the available data with many (e.g. at least 20) different weather forecasts and satellite images in order to increase the robustness and practicality of predictions
Visualisations	Be able to create standard and custom visualisations over the data that users with basic ICT knowledge can easily interact with farmers/ fishers, etc.
Visualisations	Be able to combine different series of data, create custom visualisations, charts, maps and graphs of selected variables in order to gain intuition over data, models and events and other statistical observations
Visualisations	Support map-based visualisations based on HPC software that require real time data processing
Visualisations	Combine multiple time series in the same graph in order to visually compare and correlate multiple variables in time
Support	Have expert technicians install hardware (cameras and sensors) at the farm (or premises) to enrich the existing databases at minimal cost
Alerts	Ability to receive customisable real-time notifications and warnings in order to make timely decisions

## 5 Conclusions and Future Work

The present paper presented the interviews conducted in nine (9) diverse precision agriculture and livestock farming cases in order to identify common requirements and challenges in terms of data collection and management, Big Data, HPC infrastructure and decision making.

The interviews outlined and revealed a set of common user requirements and indicated the users' point of view concerning the necessity of a robust and scalable infrastructure that would enable better and faster predictions and data analytics taking advantage of the fusion of different and timely updated datasets. The users put emphasis on the importance of the low cost and the ease of installation and operation of such infrastructure and decision-making tools. A high development and maintenance cost could render the technological solutions inapplicable. The common requirements that derived from the interviews and user requirement analysis per use case can serve as basis for identifying functional and non-functional requirements of a technological solution of high re-usability, interoperability, adaptability and overall efficiency in terms of addressing common needs for precision agriculture and livestock farming.

Future work involves the development of a blueprint architecture and design principles for technological solutions in precision agriculture and livestock farming, as well as the development of a prototype platform that combines HPC, Big Data, Cloud Computing (services) and IoT. Their main purpose will be to provide integrated and unmediated access to a vast number of large-scale datasets of diverse types, coming from a plethora of different sources, as well as to enable the actual generation of value and extraction of insights out of these data.

## Acknowledgements

This work has been co-funded by the CYBELE project, a European Commission research program under H2020-825355. We are particularly grateful to the interviewees from the nine CYBELE case study demonstrators.

## References

1. Schimmelpfennig, D. (2016). Farm profits and adoption of precision agriculture (No. 1477-2016-121190).
2. Schimmelpfennig, D., & Ebel, R. (2016). Sequential adoption and cost savings from precision agriculture. *Journal of Agricultural and Resource Economics*, 41(1835-2016-149552), 97-115.
3. Janssen, S. J., Porter, C. H., Moore, A. D., Athanasiadis, I. N., Foster, I., Jones, J. W., & Antle, J. M. (2017). Towards a new generation of agricultural system data, models and knowledge products: Information and communication technology. *Agricultural systems*, 155, 200-212.
4. Srbinovska, M., Gavrovski, C., Dimcev, V., Krkoleva, A., & Borozan, V. (2015). Environmental parameters monitoring in precision agriculture using wireless sensor networks. *Journal of cleaner production*, 88, 297-307.
5. Mahlein, A. K. (2016). Plant disease detection by imaging sensors—parallels and specific demands for precision agriculture and plant phenotyping. *Plant disease*, 100(2), 241-251.
6. Vuran, M. C., Salam, A., Wong, R., & Irmak, S. (2018). Internet of underground things in precision agriculture: Architecture and technology aspects. *Ad Hoc Networks*, 81, 160-173.
7. Bendre, M. R., Thool, R. C., & Thool, V. R. (2015, September). Big data in precision agriculture: Weather forecasting for future farming. In *2015 1st International Conference on Next Generation Computing Technologies (NGCT)* (pp. 744-750). IEEE.
8. Keswani, B., Mohapatra, A. G., Mohanty, A., Khanna, A., Rodrigues, J. J., Gupta, D., & de Albuquerque, V. H. C. (2019). Adapting weather conditions based IoT enabled smart irrigation technique in precision agriculture mechanisms. *Neural Computing and Applications*, 31(1), 277-292.
9. Wolf, J., Bhandari, S., & Raheja, A. (2017). Unmanned Aerial Vehicles for Precision Agriculture in Orchard Crops.
10. Gevaert, C. M., Suomalainen, J., Tang, J., & Kooistra, L. (2015). Generation of spectral-temporal response surfaces by combining multispectral satellite and hyperspectral UAV imagery for precision agriculture applications. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 8(6), 3140-3146.
11. Wolfert, S., Ge, L., Verdouw, C., & Bogaardt, M. J. (2017). Big data in smart farming—a review. *Agricultural Systems*, 153, 69-80.
12. van der Heide, E. M. M., Veerkamp, R. F., van Pelt, M. L., Kamphuis, C., Athanasiadis, I., & Ducro, B. J. (2019). Comparing regression, naive Bayes, and random forest methods in the prediction of individual survival to second lactation in Holstein cattle. *Journal of Dairy Science*, 102(10), 9409-9421. <https://doi.org/10.3168/jds.2019-1629>
13. Michael Lant (2010). How to Easily Prioritize Your Agile Stories. Available at: <https://michaellant.com/2010/05/21/how-to-easily-prioritize-your-agile-stories/>