Backman, Juha; Linkolehto, Raimo; Koistinen, Markku; Nikander, Jussi; Ronkainen, Ari; Kaivosoja, Jere; Suomi, Pasi; Pesonen, Liisa

**Cropinfra research data collection platform for ISO 11783 compatible and retrofit farm equipment**

*Published in:* Computers and Electronics in Agriculture

**DOI:**
10.1016/j.compag.2019.105008

Published: 01/11/2019

**Document Version**
Publisher's PDF, also known as Version of record

*Please cite the original version:*
Cropinfra research data collection platform for ISO 11783 compatible and retrofit farm equipment

Juha Backman⁎, Raimo Linkolehto, Markku Koistinen, Jussi Nikander, Ari Ronkainen, Jere Kaivosoja, Pasi Suomi, Liisa Pesonen

a Natural Resources Institute Finland (Luke), Production Systems, Tietotie 2, 02150 Espoo, Finland
b Aalto University, School of Engineering, Department of Built Environment, Otakaari 4, Espoo, Finland

ARTICLE INFO

Keywords:
ISOBUS
Document database
Development platform
Data capture
Data transfer

ABSTRACT

The agricultural machinery produces an increasing number of measurements during operations. The primary use of these measurements is to control agricultural operations on the farm. Data that describes the in-field variation in plant growth potential and growing conditions is the basis for precision farming. The secondary use for the gathered information is documentation of work and work performance for business purposes. Researcher also benefits from the increasing measurement capabilities. Biologists and agronomists can model the crops and agronomic phenomena. Work scientists can analyse the agricultural work processes. And finally, machines with additional accurate sensors can be used for agricultural machine product development and technological research purposes.

This paper concentrates on an independent research data collection platform (Cropinfra) which can be used to collect data for all above mentioned purposes. Data can be collected both from ISOBUS (ISO 11783) compliant machines as well as older and proprietary systems and stored to database for further analysis. The farm machines in Cropinfra are supplemented with extra sensors that are more accurate than existing in commercial machines. Therefore, the Cropinfra can be used as a reference measurement system to verify the correct operation of the machines as well as to produce data for biological research purposes. This paper will also present how the cloud connection of the data collection system can be realized. The solution was designed to be compatible with the existing ISO 11783-10 standard. The examples presented in this paper verify that the solution works in real farming environment. The data has been used in numerous research projects already, and in the future the data will be an important asset when machine learning and other artificial intelligence methods will be studied and utilized.

1. Introduction

Agricultural machinery has gone through significant changes during last decades. The introduction of electronic control units (ECU) has opened up new ways for carrying out agricultural operations (Stone et al., 2008). Especially the GNSS (Global Navigation Satellite System) based solutions enable position based precision farming (Thomasson et al., 2019). At the same time the sensors and other measurement devices have also developed tremendously and their unit price has come down (SACHS, 2014). Agricultural machinery produces an increasing number of measurements during operations. This information can be used for many purposes. One primary purpose is to use data gathered to control agricultural operations. Soil conditions, crop growth and health can be monitored and used to make fertilization and plant protection decisions and prescription maps for precision farming operations (Söderström et al., 2016). Another use for the information is documentation of work and work performance for business purposes. For example, an agricultural contractor can verify the work they have carried out by showing the measurement data to the customer (Suonenieto, 2019; Dataväxt, 2019). In addition to the land owners and farmers, researchers can also benefit from the increasing measurement capabilities. Biologists and agronomists can model crops and agronomic phenomena, verify hypotheses and produce new knowledge (Kaivosoja et al., 2013). Machinery data is an important source of so called big data in agriculture (Wolpert et al., 2017). Also work scientists can analyse the agricultural work processes in more detail. And finally, machines with additional and accurate sensors can be used for Product development of agricultural machinery and technological research.

⁎ Corresponding author.
E-mail address: juha.backman@luke.fi (J. Backman).

https://doi.org/10.1016/j.compag.2019.105008
Received 21 November 2018; Received in revised form 30 July 2019; Accepted 12 September 2019

© 2019 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY license (http://creativecommons.org/licenses/BY/4.0/).
purposes (Öhman et al., 2004; Suomi and Oksanen, 2015) Data can be collected from the machinery using a number of different ways. The existing electronic control units (ECUs) can create their own proprietary logs, which can be downloaded to USB stick or, in some cases, to cloud servers. Another way is to use an independent logging unit which does the same, but is not directly bundled with any manufacturer’s control units. In both cases, the data collection includes four phases: measurement devices or sensors, data capture, data transfer, and data storage. However, the last two phases of the process can be fundamentally different depending on the technology used.

Farms can be very heterogeneous, as the history of the farm, the geography of the land used by the farm, the crops and livestock grown on the farm, as well as technological capability and business strategy of the farm all affect how the farm works. The type of production and the size of the farm define the type and size of machinery required. Furthermore, farm’s machinery is often a mixture of older and newer machines from a number of manufacturers. Rarely any manufacturer can provide all machines, devices and software that the farmer needs. Therefore there is a strong need for compatibility between different brands. In agriculture the communication between different components in tractor-implement system is defined by the commonly agreed standard ISO 11783 (ISO, 2017) – marketed under the name ISOBUS. In agriculture the communication between different components in tractor-implement system is defined by the commonly agreed standard ISO 11783 (ISO, 2017) – marketed under the name ISOBUS (AEF, 2019a). The Agricultural Electronic Industry Foundation (AEF) develops and markets the ISOBUS system.

In the previous research projects the information management of farm machinery and the data transfer from the machinery to farm management information systems (FMIS) and various on- and off-site services has been studied. Fountas et al. (2015) have proposed a Farm Machinery Management Information System based on interviews of tractor operators and farm managers. The solution was presented as a rich picture from which a conceptual model for conventional farm machinery and agricultural robots was developed. Steinberger et al. (2009) have presented a prototype implementation of ISO 11783 task controller without user interface for data logging. The data was stored in ISO XML format in a PDA and transferred to a local storage using WLAN. For telemetry purposes feasibility of OPC Unified Architecture (OPC UA) technology to transfer ISO 11783 related process data between farm machinery and the internet was studied by Oksanen et al. (2015). Oksanen et al. presented also an approach to convert ISO 11783 DDOP (Device Description Object Pool) to OPC UA information model. In the experiments the latency and bandwidth usage of communication were measured and found feasible.

Commercial solutions for data collection exist from all major farm machinery manufacturers. For example AGCO has AgCommand, John Deere has JDLink, CLAAS has TONI, etc. There exist also retrofittable fleet management and telemetry solutions from different manufacturers, for example Suonientieto AgriSmart, Wapice IoT-Ticket, etc. Currently the ISOBUS working group at AEF is working on a standard for data transfer between ISOBUS-compatible machinery and farm management systems (AEF, 2019b). The standard is expected to be ready for use during the year 2019.

This paper concentrates on an independent research data collection platform (Cropinfra) which can be used to collect data both form ISOBUS compliant machines as well as old and proprietary systems and store the data to database for further analysis. The farm machines in Cropinfra are supplemented with extra sensors that are more accurate than those existing in commercial machines. Therefore the Cropinfra can be used as a reference measurement system to verify the correct operation of the machine. The Cropinfra has developed in many national and international research projects starting from 2003 (Fig. 1). The Cropinfra platform is a concept for future farm’s data collection and management structure that is capable to serve also research purposes, and which has been verified by implementing it in a real scale experimental farm for technology development. The research farm was located in Southern Finland, in Vihti. It included 151 ha of field, all necessary buildings for farm activities, four tractors, a combine harvester and all necessary implements to operate the farm. The results of Cropinfra based research projects have recently been reported in Nikander et al. (2017), Nikander et al. (2015) and Pesonen et al. (2014).

In this paper we report the methods and means used to develop Cropinfra into a framework that can be used as a general agricultural data collection and exploitation system for research purposes. The primary objective of this paper is to depict how the existing technology on agricultural data collection can be integrated into an open research system. A secondary objective is to present the experiences and results of different ways to realise the cloud connection of data collection system that is at the moment crucial improvement to be included in the ISOBUS standard.

The rest of this paper is organized as follows. In Section 2, methods, we describe how the information management and software environment in a modern farm can be organized, focusing on the parts that are vital to this paper. In Section 3 we describe how the Cropinfra platform implements farm information management infrastructure, and how the system has been used for data capturing and management using examples. Section 4 contains the discussion and conclusions of our work.

2. Methods

This work is based on data collection from agricultural machinery, as well as the analysis, and exploitation of that data both in cloud systems and in the machinery. In this work, we concentrate on the four phases of data collection: measurement, capture, transfer, and storage. In the measurement phase, the various sensors take readings and provide values that can be stored. In the capturing phase, the sensor values are combined and given semantics by assigning each value to a specific variable. In the transfer phase the data created in the capturing phase is transferred from the field machine to a cloud service. Finally, in the storage phase the data is stored in a cloud service for further analysis and exploitation.

Data collection from tractors, implements, and other agricultural machinery is primarily based on previous development carried out for digital systems in heavy machinery and vehicles in general. These systems have then been adapted to use in agriculture, and extended in order to best serve the specific needs of the agricultural sector. The cloud systems and other software used outside the machinery mostly use standard data exchange methods. The data formats and analysis methods used are, of course, agriculture-specific in order to both capture the nature of the data being used as well as providing useful information to act as basis for managing farm, i.e. planning and executing specific agricultural operations.

2.1. Farm information management

The digital environment of a modern farm can already be extremely complex and is becoming increasingly so. Fig. 2 shows an example of a farm digital environment, in which the parts that are relevant to this study are emphasized. The Figure has been adapted from one originally published in Nikander et al. (2015). The elements of the Figure that are out of scope for this work are shown in order to provide context for the reader how the data transfer in this work would fit into the overall software environment of a future farm. In the bottom of the Fig. 2 are depicted the measurement and capturing phases of data collection that are carried out in the tractor-implement. The transfer phase is represented by the arrow between the tractor-implement and the cloud, and the storage phase is in the cloud environment at the top of the Figure. As can be seen in Fig. 2, the data transfer infrastructure described above is only small part of the overall digital environment of a modern farm. The rest of the farm digital environment is described in the original publication in Nikander et al. (2015).

2.2. Tractor and implement data

Currently, data management and transfer in tractors and
implements are, in practice, based on one of three methods. The first is the ISO 11783 (ISOBUS) standard (ISO, 2017a,b), the second are proprietary systems, typically used in older machinery that do not conform to ISOBUS, and the third are analog systems, where the data first needs to be converted to a digital signal in order to be transferred for permanent storage.

ISOBUS defines the communication protocol used between different components of the tractor-implement system. Data transfer is in specified form, and therefore the data created by an ISOBUS system can easily be read and understood by any interface conforming to the standard. The majority of machine-related data (such as implement status, engine performance, or speed) are covered by the standard. There is also a part where manufacturers can add data not currently included in ISOBUS. Oksanen et al. (2005) has published an informative
article about the standard in its use.

In addition, ISOBUS defines a set of functionalities and devices required from the machinery. The simplest ISOBUS system consists of a tractor ECU (T-ECU), an implement ECU (I-ECU) and a universal user interface called universal terminal (UT) or virtual terminal (VT) (AEF, 2019b). These devices form the basic structure used to control the tractor and implement combination. There can be additional devices attached to the system, such as a positioning device (GNSS), or a task controller (TC). The task controller can be used to control the implement, as well as to store the data logged in an executed field operation (ISO, 2015). A TC that is capable of location-specific control and logging is called TC-GEO and a TC that is capable only of data logging is called TC-LOG.

Not all the machines in the agriculture are compatible with the ISOBUS standard. There are also numerous proprietary solutions for implement control. There is no single solution for capturing data from these systems. Instead, each system needs an interface to adapt the data provided by the proprietary system to the farm software infrastructure. Should the data be used to control the device, the conversion needs to be done the other way around. However, it can be difficult to use farm data as input into a proprietary system unless the control software of the system has an open input method.

In the systems where digital controllers are not used, signals between controls, sensors and actuators use only voltage or current signals. In this case, the system needs a setup similar to proprietary systems, where the signal is first converted to a digital form, and then adapted to the needs of the larger farm infrastructure. As analog systems seldom are able to take input except through the analog controllers, it can be difficult to use data to control such a system.

2.3. Back-end technologies in cloud interfaces

Although the ISO 11783 standard will eventually cover the data transfer from the tractor to cloud services (AEF, 2019b), these technologies have already been used elsewhere. In essence, there are two different technologies needed: a data transfer method and a data format. Possible data transfer technologies are, for example, REST and MQTT. JSON and Protobuf are potential candidates for data format.

2.3.1. REST

REST (Representational State Transfer) is an architectural style, not a standard, that describes a set of constrains over the HTTP (Leach, 1999a) application protocol for designing distributed client/server related information systems that provide interoperability between computer systems on the Internet. Web services following the REST architectural style are called RESTful web services. The objects provided by the RESTful web services (sometimes referred as RESTful API) are called resources. The resources are typically identified by using URIs. The role of the application server is to provide access to resources and the client accesses, and in some cases it modifies the resources (Fielding, 2000). A client uses HTTP methods for reading a resource (GET), for creating a new resource (POST), to update an existing resource or create a new resource (PUT) and for removing a resource (DELETE) (IETF Trust, 2007).

2.3.2. MQTT

MQTT (Message Queuing Telemetry Transport) is a simple and lightweight publish/subscribe-based messaging protocol running over TCP/IP, or other similar protocols like Bluetooth (ISO/IEC 20922, 2016). MQTT is designed for devices in Machine-to-Machine communication and IoT context with limited capabilities and networks with low-bandwidth, high-latency and possibly unreliable connectivity. In attempt to assure some degree of delivery in constrained environments the standard describes three qualities of service: at most once, where occasional message loss is allowed, at least once, where message delivery is guaranteed but duplicates can occur and exactly once, where messages are assured to arrive and only once.

The underlying publish/subscribe messaging paradigm means that the system requires a message broker. The broker operates between publishers and subscribers taking incoming messages from data provider and routing them to relevant destinations. The information (published data) is organized in a hierarchy of topics and the broker distributes the information to subscribers according to the subscribed topics. Public/subscribe model implements decoupling: on the general level publishers and subscribers have no need to know anything about each other’s systems (ISO/IEC 20922, 2016).

2.3.3. JSON

JSON (JavaScript Object Notation) is a programming language independent, open-standard data-interchange file format for structuring and sharing data. It is based on subset of the JavaScript programming language and it can represent any serializable data. The format can be generated and parsed programmatically, and is human-readable at the same time. JSON is used as a data exchange format in asynchronous web applications and as a natural data exchange format in interfaces that are bound to document oriented database management systems (Bray, 2015, 2017).

2.3.4. PROTOBUF

Protocol buffers offer a programming language neutral and platform-neutral mechanism for serializing structured data as well as an interface description language in the inter-machine communication domain (Google, 2018). A data structure in Protobuf is defined using three types of variables: required fields, repeated fields, and optional fields. Data is compiled into data access classes for automated reading and writing for a variety of data streams. For data transfer the Protobuf text format messages are encoded into a binary format, which reduces the message size up to 40% and speeds up the parsing time up to 50 times when compared to the size and processing time of a text based format like XML. Protocol buffers are well suited to the environments with time critical computing requirements. (Google, 2018)

3. Results

Fig. 3 depicts the overall structure and the different devices and services of the Cropinfra research data collection system on a conceptual level. At the bottom of the picture there is the measurement level, which contains the physical sensors that are either commercial ones, or reference sensors added for research purposes. Sensors are connected to ECUs on the data capturing level that transform the analog signals and transmit those to ISOBUS communication (CAN) bus. Above the ISOBUS CAN in Fig. 3, there are control and logging devices that capture, utilize and store the measurements locally. Finally, at the top of the Figure are the data transfer and storage means. The data itself is stored either directly in the file system, or in a database. The different phases of the data collection process are explained in further detail in the following sections. In the following, we will first discuss how legacy machines that provide only analog readings need to be modified in order to fit into the Cropinfra framework. Then we will discuss data capturing, data transfer in both bulk and real-time modes, and data storage. After that, we will discuss the data capturing and transfer results using examples where the Cropinfra framework has been successfully used.

3.1. Measuring legacy machines and using reference measurements

To support all different implements in Cropinfra, analog sensors and I/O device (Axiomatic, 2008) are used to provide measurements from non-ISOBUS implements to the CAN-bus. The I/O device sends the machine identifier that is the same as the globally unique NAME-identifier used in ISOBUS (ISO, 2019). The identification and the measurements messages from I/O device can be read similarly as from
the ISOBUS machines. In another words, the old and proprietary tractors and implements are modified to provide limited ISOBUS-like functionality (see Fig. 4).

For example Valtra 8950 tractor measurements were transmitted to ISOBUS using the I/O device which measured the signals from ISO 11786 signal connector. The ISO 11786 defines a signal connector where the ground speed, linkage position and PTO speed are provided using pulses and analog signals. The fuel consumption was measured using two FLOWMATE OVAL M-III (Oval, 2016) flowmeters (Fig. 5).

The implements without proprietary ECU were also equipped with Axiomatic I/O device. For example the Potila Magnum 540 harrow does not have any electric control (see Fig. 6). All the functions are directly hydraulically controlled. The I/O device was used to measure whether the implement is in working position or not. In other implements the I/O device was also used to provide reference measurements to validate correct operation of the commercial controllers.

### 3.2. Data capturing

As explained in Section 2.2, the data can be captured from the ISOBUS CAN-bus using standard ISOBUS TC-GEO or TC-LOG device. However, if such device is used for research purposes, it restricts the use of a parallel device with the same functionality. By the TC protocol definition in the standard, the implement can communicate with only one TC at a time. These kinds of conflicting situations may occur when technological research projects are conducted and ISOBUS prototype
devices are tested. Thus, a data capturing device that only listens to the other ECUs and doesn’t send anything to the can-bus was developed.

Furthermore, the requirements for data capturing in different research projects and in different machines can vary. To meet the requirements, the data capturing software had to modify in different research project. The National Instruments LabVIEW system engineering software (National Instruments, 2018) was selected to be used to program the data capturing software. LabVIEW based software are fast to program, flexible and self-documenting.

The data capturing software was designed to be user friendly and reliable. First, the program identifies all the devices in the ISOBUS network using the address claim functionality (ISO, 2019). All the devices have to claim their own address in the bus using globally unique ISOBUS NAME. The data capturing program is able to identify the machines that are connected based on this NAME. Next, the user is asked to complete the rest of the information before the data capturing is started. The user interface is described more detailed in the next subsection. When all the initial information is completed and user has started the data capturing, the program stores all essential data from the ISOBUS network. The data capturing system does not send any request message to the can bus or use the task controller protocol to communicate with the implement. However, the software stores the process data messages between the TC and the Implement as well as all other messages that are in the bus.

### 3.2.1. User interface

The user interface of the data capturing software is shown in Fig. 7 and in Fig. 8. The first picture is from the initial data input tab, which is used at the beginning of the work. The second picture is from the driving tab, which is used during the work.

As described above, the data capturing software automatically detects the machines that are connected in the bus. Those are set to be default values in the dropdown boxes in Fig. 7 (“Traktori” [tractor in English] and “Työkone” [implement]). The user can also manually select the machinery. In addition, the user has to select the correct worker (“Suorittaja”), operation (“Työtehtävä”), parcel (“Lohko”), plant type (“Viljelykasvi”) and variety (“Siemenlajike”) and also the targeted seed (“Siemenmäärä”) and fertilization rate (“Lannoitemäärä”) as well as corresponding calibration values (“Kiertokoe”). In the bottom of the page, there are also place for free comment (“kommentti”) that will be added in the log file. The comment can be added any time during the operation. The user interface is slightly different when different types of implements are used. For example when capturing data from a ploughing task, there is no need to select the plant or rate values and in plant protection tasks there is a selection menu for pesticides instead of seeds and fertilizers.

Fig. 8 shows the tab used during a field operation when data capturing from the sensors has been started. During the work, the user interface is used to monitor that all ECUs in the ISOBUS network are working and sensors produce the correct data. For example, the GPS-box and TC-box are green if those produce data to the bus, otherwise they are red. The current field parcel based on the GPS-location is also show in the text box to verify that the worker is in a correct parcel. The tractor and implement measurements are show as meters for visual verification of correct measurement values.

### 3.2.2. Hardware for data capturing

The data capturing software was made using the National Instruments LabVIEW system engineering software, so the executable-version of the software could be run in any device that has a LabVIEW runtime. In Cropinfra, Panasonic Toughbook CF-19 rugged laptop computer was used. The system was used in tractors and in a combine harvester. A docking station was mounted in each machine to which power supply, CAN adapter and 3G/4G router (ASUS 4G-N12) was connected. The basic setup of the hardware is shown in Figs. 9 and 10. The initial device setup was selected at the beginning of the Cropinfra-
related projects in 2003. However, additional devices were used in different research projects.

3.3. Bulk data transfer

The Cropinfra data capturing software stores all the collected information internally in a csv-type text file. Another option would have been to use the ISO 11783-10 XML format. The example of the structure of the data file is in the attachments (Table 3). The data file consists of two parts: header and data. The header stores the information that describes the agricultural operation and is constant during the process (initial data input tab in user interface). In the data section, the process data produced by the machinery and sensors are stored using 5 Hz logging frequency. The file is created when the user starts the data capturing and the data is flushed to the disk in every iteration cycle to ensure that there will not be any loss of data even if the system crashes.

Fig. 8. Data capturing user interface during the work.

Fig. 9. Rugged computer for data capturing is mounted in the tractor.

Fig. 10. Wireless 4G router is used to cloud connection and national instruments CAN-adapter for ISOBUS connection.
during a field operation.

The final log file is transferred to the database either manually using an USB-stick or automatically using FTP. When using FTP, a file transfer daemon monitors the local folder that is used to store the log files. If the daemon notices that a new log file is created or data capturing software has been closed, it is concluded that previous log file is completed and the transfer of the previous log file can be started. The file transfer daemon keeps track of all the log files that are created and transferred. If the network is disconnected or the computer shutdown, the data transfer is tried again after the computer is started and connection resumed.

3.4. Real-time data transfer

The Real-time data transfer is integrated to the data capturing software. Unlike the bulk data transfer, the real-time data transfer attempts to follow the ISO 11783-10/11 standard. The ISO 11783-10 XML schema for the data is used in the system (ISO, 2015).
The REST (W3C, 2011)/JSON (ECMA International, 2017) API (Application Programming Interface) was selected to be used, because the JSON is human readable and XML is easy to convert to JSON objects. There are at least two different ways to implement the online data transfer using REST/JSON communication protocol adapting the ISO 11783-10 standard:

1. Using ISO 11783-10 XML structure as it is and converting all the root elements to own JSON objects. These objects are requested and sent to and from the cloud database, in order to get all the needed ids. New task is started by sending a new TSK element (element structure is explained later) with correct references.

2. Simplify the ISO 11783-10 XML structure by removing the references inside the XML and moving all the elements inside the TSK element. All information is sent once inside the TSK JSON, which starts the new task.

As examples of a TSK document, the following JSON objects are shown:

```
A TSK document
```

As an example of a TLG document, the following JSON objects are shown:

```
A TLG document
```
Both approaches have advantages and disadvantages, but the first option was chosen because it follows the existing standard more closely.

### 3.4.1. Backend

The backend application stack is built on REST/JSON API with MongoDB document oriented Data Base Management System (MongoDB, 2008). The interface is implemented using Java and deployed on GlassFish application server, the open source Java EE (Oracle, 2018) reference implementation, and published using representational state transfer architectural style that provides a set of operations that a remote client can invoke over a network using the HTTP protocol (Leach, 1999a). The backend data storage solution is built on MongoDB, a document oriented non-relational database management system. MongoDB implements a flexible data model that supports storing the JSON data as it is, enabling agile prototyping and efficient client-to-backend data processing.

### 3.4.2. Data operations

The prototype REST/JSON API provides simple resources for online data transfer from mobile unit to data storage and for requesting data from that storage. The API implements HTTP Basic Authentication (Leach, 1999b) for authentication and authorization.

The data transfer sequence starts with authentication handshake between mobile unit and back-end (Fig. 11). The client accesses the authenticate resource using POST HTTP method and sends BASE64 (Joseffson, 2006) encoded credentials (username and password) as Authorization header. Also a valid client id is required. In case of successful authentication sequence the authorization is granted by assigning an access token to be used in further requests.

After successful authentication and authorization process the client requests a task identifier using GET/taskid resource (Fig. 12). The task identifier is used by the back-end system as a data set join key to identify the task specific documents in the database. The task identifier is required as a path parameter whenever the data is posted to the resource.

Data is sent to POST/data/[taskid] resource using the HTTP POST method (see Fig. 13). The resource accepts application/json mime type and the payload JSON array can contain TSK or TLG type documents, or both. The back-end stores each document in received array as an individual object into database (Table 1).

### Table 2

The sizes of the transmission buffers in different tests.

<table>
<thead>
<tr>
<th>Name of the parcel</th>
<th>Area (ha)</th>
<th>number of TLG messages</th>
<th>Buffer size (number of TLG messages)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Min</td>
</tr>
<tr>
<td>Pumppulohko A</td>
<td>2.99</td>
<td>23,939</td>
<td>0</td>
</tr>
<tr>
<td>Heikkilä A</td>
<td>2.86</td>
<td>45,614</td>
<td>0</td>
</tr>
<tr>
<td>Uutela A</td>
<td>11.29</td>
<td>112,268</td>
<td>0</td>
</tr>
<tr>
<td>Kaupanpelto A</td>
<td>1.84</td>
<td>25,614</td>
<td>0</td>
</tr>
<tr>
<td>Tientaus A</td>
<td>4.09</td>
<td>44,537</td>
<td>0</td>
</tr>
<tr>
<td>Uutela A</td>
<td>11.29</td>
<td>27,443</td>
<td>0</td>
</tr>
<tr>
<td>Kirjava Pohj. A</td>
<td>3.09</td>
<td>37,734</td>
<td>0</td>
</tr>
<tr>
<td>Kirjava Et. A</td>
<td>7.56</td>
<td>81,124</td>
<td>0</td>
</tr>
<tr>
<td>Pello A (part 1)</td>
<td>12.30</td>
<td>134,545</td>
<td>0</td>
</tr>
<tr>
<td>Pello A (part 2)</td>
<td>12.30</td>
<td>185,123</td>
<td>0</td>
</tr>
<tr>
<td>Pello A (part 3)</td>
<td>12.30</td>
<td>166,510</td>
<td>0</td>
</tr>
<tr>
<td>Pello A (part 4)</td>
<td>12.30</td>
<td>52,942</td>
<td>0</td>
</tr>
<tr>
<td>Hovin Luhta A (part 1)</td>
<td>4.66</td>
<td>70,303</td>
<td>0</td>
</tr>
<tr>
<td>Hovin Luhta A (part 2)</td>
<td>4.66</td>
<td>32,846</td>
<td>0</td>
</tr>
<tr>
<td>Luhta A</td>
<td>8.02</td>
<td>69,209</td>
<td>0</td>
</tr>
<tr>
<td>Riihipelto A</td>
<td>5.26</td>
<td>37,251</td>
<td>0</td>
</tr>
<tr>
<td>Kirjava Länt. A</td>
<td>5.85</td>
<td>55,441</td>
<td>0</td>
</tr>
<tr>
<td>Kirjava Länt. A</td>
<td>5.85</td>
<td>55,441</td>
<td>0</td>
</tr>
</tbody>
</table>

Both approaches have advantages and disadvantages, but the first option was chosen because it follows the existing standard more closely.
3.4.3. Beginning of the work

The work is started by creating a new task, including the header information of the work that is meant to be done. The database returns the task id that is used as a reference in communication with the database after the beginning of the work.

**URI:** POST {BASE URL}/data/[taskid]

**Send:**

```java
("TK":{"B":<TaskDesignator>}, "D":"FRM<FarmIdRef>", "E":"PPD<ParcelIdRef>", "F":"WKR<WorkerIdRef>", "G":2,"TIM":{"A":<StartTime>}, "OTP":{"A":"CPC<OperationIdRef>", "B":"OTQ<OperationSpecificationIdRef>", "CTP":{"A":"CTP<PlantIdRef>", "B":<PlantDesignator>}, "CVT":{"A":"CVT<PlantVarietyIdRef>", "B":<PlantVarietyDesignator>}), "PDT":{"A":"PDT<FertilizerIdRef>", "B":<FertilizerDesignator>}, "DAN":{"A":<TractorNAME>}, "C":"DVC<TractorIdRef>",("A":<ImplementNAME>,"C":"DVC<ImplementIdRef>"))

Receive:

HTTP response 200/401
```

3.4.4. Definition of captured data

It is supposed that each machine has a device description according to ISO 11783 standard. The device description defines the structure of the TLG elements. In this example two different device descriptions are used: one for the tractor and one for the implement. In the attachments, device description of the ISOBUS compatible Valtra T-163 tractor and also ISOBUS compatible Junkkari Maestro 4000 seed drill are presented. Both produce measurement messages to the ISOBUS which can be logged by the data capturing software. The tractor T-ECU sends the wheel and ground speed, engine RPM, PTO RPM, Rear hitch position and diesel consumption. For the urea consumption, a separate ECU is used to transfer the message from the tractor bus to the ISOBUS. The seed drill sends fertilizer and seed rates as a response to the TC message. An external I/O device was used to send the reference measurements from the fertilizer and seed rates. Position, speed, direction and GNSS quality information is received from the GPS.

3.4.5. Calibration data

The calibration data is sent only when the task is started or when the calibration is changed. This kind of data is for example seed and fertilizer rates when prescription control is not used.

Template for the calibration data is:

**URI:** POST {BASE URL}/data/[taskid]

**Send:**

```java
("TLG":{"TIM":{"A":<time>"}, "DLV":{"A":<DDI>, "B":<value>, "C":<DTE>, "G":<DVC>}, ... }})
Receive:

HTTP response 200/401
```

Note! In the standard, there is not G-attribute to specify the device; instead it is supposed that device element numbers (DET) are unique. For example, the seed rate and fertilizer rate are sent using:

**URI:** POST {BASE URL}/data/[taskid]

**Send:**

```java
("TLG":{"TIM":{"A":<time>"}, "DLV":{"A":0006, "B":<seed rate>"}, "C":<DTE78>, "G":<DVC2>}, 
("A":0006, "B":<fertilizer rate>, "C":<DTE75>, "G":<DVC2>)})
Receive:

HTTP response 200/401
```

3.4.6. Process data

Process data is sent similarly to calibration data. However, there can be multiple TLG-elements in the same data frame to allow burst sending of measurements from different time stamps. In addition to that, the TLG-element also includes the PTN-element which defines the geographical location of the machine when the measurement has taken place.

Template for the process data:
In this example the process data consist of TIME_PC, Seed_r, Fertilizer_r, Position, PTO, Hitch, Diesel, Urea, Tr_rpm, Tr_dir, Tr_W_Speed, Tr_R_Speed, LAT, LON, GPS_Speed, Dir, Alt, Quality, Sat_num. Those values are originally used in the proprietary data format and mapped to corresponding device description elements. The JSON frame is hence:

```json
3.4.7. Requesting data
Prototype provides simple resources for requesting data from process data collection. Client can request all the TSK documents in collection or alternatively a TSK document or a TLG document matching a given task identifier. The resources are:
– GET/data/tsk
– GET/data/tsk/[taskid]
– GET/data/tlg/[taskid]
If successful, the response mime type is application/json and the body is formatted as JSON array with 1−n JSON documents ([{}(...)]).
All the POST and GET methods require client id and access token fields in HTTP request header (Fig. 14).

3.5. Results of data capturing and real-time data transfer
In the presented examples in Section 3.4.6 Process data, the payload size for the process data (TLG message) is 570 bytes if only measurements from one time instance are sent. The measurements are updated 5 times per second, which means that the bandwidth requirement is 2850 bps plus 10% overhead from HTTP-frames producing totally 3.1 kbps bandwidth requirement. The upload data rate in 2G EGDE is 60 kbps, in 3G 2 Mbps and in 4G LTE 75 Mbps. Even though the upload data rate varies according to the load of the network, the required bandwidth is only a fraction of the available capacity. However, the latency of the mobile network and HTTP server makes it impossible to post messages with 5 Hz data rate. The solution was to pack several TLG messages together and send those in one http message. In the Cropinfra, the default was to pack 5 TLG messages together, and if the network is disconnected, all messages after the last successful sent are packed together.

The data transfer was tested during sowing operations in Luke’s experiment farm in Vihti that is located in Southern Finland. There are no big cities near the farm, so the mobile network is typical to countryside. The sizes of the transmission buffers (TLG message queue length) in different tests are listed in Table 2 together with the field area and total number of TLG messages.

<table>
<thead>
<tr>
<th>Buffer Size</th>
<th>Field Area (ha)</th>
<th>Total Number of TLG Messages</th>
</tr>
</thead>
<tbody>
<tr>
<td>5 TLG</td>
<td>81.1</td>
<td>1,202,443</td>
</tr>
<tr>
<td>10 TLG</td>
<td>100.0</td>
<td>2,404,886</td>
</tr>
<tr>
<td>15 TLG</td>
<td>150.0</td>
<td>3,607,328</td>
</tr>
</tbody>
</table>

4. Discussion and conclusions
This paper has presented how data collection from field operations can be implemented to support research. The system supports both state-of-the-art ISOBUS (ISO 11783) compatible tractors and implementations as well as old proprietary machinery. The provided examples verify that the solution can be used in real scale farm.

In the presented examples in Section 3.4.6 Process data, the payload size for the process data (TLG message) is 570 bytes if only measurements from one time instance are sent. The measurements are updated 5 times per second, which means that the bandwidth requirement is 2850 bps plus 10% overhead from HTTP-frames producing totally 3.1 kbps bandwidth requirement. The upload data rate in 2G EGDE is 60 kbps, in 3G 2 Mbps and in 4G LTE 75 Mbps. Even though the upload data rate varies according to the load of the network, the required bandwidth is only a fraction of the available capacity. However, the latency of the mobile network and HTTP server makes it impossible to post messages with 5 Hz data rate. The solution was to pack several TLG messages together and send those in one http message. In the Cropinfra, the default was to pack 5 TLG messages together, and if the network is disconnected, all messages after the last successful sent are packed together.

Totally, 1,202,443 TLG messages were sent during the data transfer tests, which means 653 Mbytes payload or 18,036,645 individual sensor readings. The total area of the parcels were 81.1 ha, which means that the average raw data production density is 8 Mbytes/ha. The buffer size was typically less than five TLG messages in queue and in average about two, which means that the TLS messages were usually sent successfully once per second. However, the network connection was lost occasionally and in every parcel the maximum size of the buffer is at least 14. The biggest buffer size is in last part of “Pelo A”, when the buffer size was 11,767. The connection was lost twice in that test. In the first time, the buffer size grew up to 8125 before the TLG messages were successfully sent. In the second time, the connection did not recover before the computer was shut down. That was the only occasion when data was lost during the tests.

This paper has presented how data collection from field operations can be implemented to support research. The system supports both state-of-the-art ISOBUS (ISO 11783) compatible tractors and implementations as well as old proprietary machinery. The provided examples verify that the solution can be used in real scale farm.

The data captured using the system has been used for example in following research; yield maps and in the combined seed drilling applied fertilizer rates were used as source data for generating new fertilizer application task (Kaivosoja et al., 2013, 2017); data from different field operations to determine the spatial overlapping of working widths (Kaivosoja and Linkolehto, 2016), where 140 different complete field operations were analysed to found out that in driving lines are overlapping 10% in average; the positioning data was also used to analyse GNSS positioning error (Kaivosoja and Linkolehto, 2015). In the work of Kaivosoja et al. (2014), application of real time web services were
adapted to the platform. In the future the collected data is even more valuable when machine learning and other artificial intelligence methods will be utilized (Kamilaris and Prenafeta-Boldú, 2018; Liakos et al., 2018).

This paper has also presented how cloud connection can be implemented in the data collection system. The solution was designed to be compatible with the existing ISO 11783-10 standard. The connection was tested and found to be working in the sowing operations of a real scale farm. There was one case of significant data loss, which happened when the computer used for data capturing was shut down before all data was transferred to the cloud service. The loss could have been prevented by saving the buffer to non-volatile memory. It was also found out that the payload size does not matter. The latency of the mobile network and the response time of the HTTP server as well as the reliability of the connection are the primary restrictions in the data transfer.

The FMIS working group in AEF is preparing the guideline to implement the cloud connection manufacturer independently. After the AEF has published the guideline, the work is started in the ISO to standardise the solution. Based on the prototypes implemented so far among the AEF community, it seems that the solution will be based on the MQTT publish-subscribe-based messaging protocol and protobuf data serializing solution developed by the Google. The solutions presented in this paper are alternative but still not conflicting solutions. The POST message used in this paper could be changed to send messages in MQTT with the topic being the same as the URL in POST. The JSON format used in this paper is conducted from the ISO 11783 standard and it is possible to convert the JSON to the protocol buffer.

### Appendix A: CSV-file format

See Table 3.

**Table 3**

Example from the beginning of the datafile.

<table>
<thead>
<tr>
<th>Data (“27.5.2015 Kirjava Lånt. A 5.85 ha Vehnä Kylvä”)</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;CROPINFORDATA&gt;&lt;DATE&gt;27.5.2015&lt;/DATE&gt;&lt;WRK&gt; Esko Kaskioja&lt;/WRK&gt;&lt;PDF&gt; Kirjava Lånt. A 5.85 ha&lt;/PDF&gt;&lt;CPC&gt; Kylvä&lt;/CPC&gt;&lt;CTR&gt; Vehnä&lt;/CTR&gt;&lt;CTP&gt; Junkari kylvökone 400 Hinattava&lt;/CTP&gt;&lt;CTR&gt;Valtra T-163 Urea&lt;/CTR&gt;&lt;CAN&gt;---------------------------------------------</td>
</tr>
<tr>
<td>Worker: Esko Kaskioja</td>
</tr>
<tr>
<td>Time: 27.5.2015 14:40:53</td>
</tr>
<tr>
<td>Task: Kylvä</td>
</tr>
<tr>
<td>Plant type: Vehnä</td>
</tr>
<tr>
<td>Plant variety: wellamo</td>
</tr>
<tr>
<td>Seedrate kg/ha: 200</td>
</tr>
<tr>
<td>Pesticide: , Water 1:</td>
</tr>
<tr>
<td>Fertilizer: NP lannos</td>
</tr>
<tr>
<td>Fertilizer rate kg/ha: 370</td>
</tr>
<tr>
<td>Parcel: Kirjava Lånt. A 5.85 ha</td>
</tr>
<tr>
<td>Tractor: Valtra T-163 Urea</td>
</tr>
<tr>
<td>Implement: Junkari kylvökone 400 Hinattava</td>
</tr>
<tr>
<td>Calibration kg Fertilizer: 3,02</td>
</tr>
<tr>
<td>Ratio Fertilizer: 3,02</td>
</tr>
<tr>
<td>Calibration kg Seed: 1,52</td>
</tr>
<tr>
<td>Ratio Seed: 1,52</td>
</tr>
<tr>
<td>Project: BIONET</td>
</tr>
<tr>
<td>Comment:</td>
</tr>
</tbody>
</table>

**Appendix A**

### CSV-file format

**Table 3**

Example from the beginning of the datafile.

<table>
<thead>
<tr>
<th>TIME_PC</th>
<th>Parcel_N</th>
<th>Mark</th>
<th>Seed_r</th>
<th>Fertilizer_r</th>
<th>Position</th>
<th>PTO</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hitche</td>
<td>Diesel</td>
<td>Urea</td>
<td>Tr_rpm</td>
<td>Tr_dir</td>
<td>Tr_W_speed</td>
<td></td>
</tr>
<tr>
<td>Qality</td>
<td>Pos_Time</td>
<td>Sat_num</td>
<td>Empty</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>134232.3000000</td>
<td>66.00000000</td>
<td>0.00000000</td>
<td>0.00000000</td>
<td>0.00000000</td>
<td>0.00000000</td>
<td></td>
</tr>
<tr>
<td>0.00000000</td>
<td>26.00000000</td>
<td>3.05000000</td>
<td>0.00000000</td>
<td>647.50000000</td>
<td>1.00000000</td>
<td></td>
</tr>
<tr>
<td>0.00000000</td>
<td>60.44967910</td>
<td>24.35707830</td>
<td>0.07200000</td>
<td>209.66310000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>68.60372900</td>
<td>1.00000000</td>
<td>38554.000000</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11.00000000</td>
<td>0.00000000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>134232.5000000</td>
<td>66.00000000</td>
<td>0.00000000</td>
<td>0.00000000</td>
<td>0.00000000</td>
<td>0.00000000</td>
<td></td>
</tr>
<tr>
<td>0.00000000</td>
<td>26.00000000</td>
<td>3.05000000</td>
<td>0.00000000</td>
<td>648.00000000</td>
<td>1.00000000</td>
<td></td>
</tr>
<tr>
<td>0.00000000</td>
<td>60.44967910</td>
<td>24.35707830</td>
<td>0.07200000</td>
<td>209.66310000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>68.60372900</td>
<td>1.00000000</td>
<td>38554.000000</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11.00000000</td>
<td>0.00000000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>134232.7000000</td>
<td>66.00000000</td>
<td>0.00000000</td>
<td>0.00000000</td>
<td>0.00000000</td>
<td>0.00000000</td>
<td></td>
</tr>
<tr>
<td>0.00000000</td>
<td>26.00000000</td>
<td>3.05000000</td>
<td>0.00000000</td>
<td>648.00000000</td>
<td>1.00000000</td>
<td></td>
</tr>
<tr>
<td>0.00000000</td>
<td>60.44967930</td>
<td>24.35707990</td>
<td>0.07200000</td>
<td>209.66310000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>68.60372900</td>
<td>1.00000000</td>
<td>38554.000000</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11.00000000</td>
<td>0.00000000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>134232.9000000</td>
<td>66.00000000</td>
<td>0.00000000</td>
<td>0.00000000</td>
<td>0.00000000</td>
<td>0.00000000</td>
<td></td>
</tr>
<tr>
<td>0.00000000</td>
<td>26.00000000</td>
<td>3.05000000</td>
<td>0.00000000</td>
<td>647.50000000</td>
<td>1.00000000</td>
<td></td>
</tr>
<tr>
<td>0.00000000</td>
<td>60.44967930</td>
<td>24.35707990</td>
<td>0.07200000</td>
<td>209.66310000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>68.60372900</td>
<td>1.00000000</td>
<td>38554.000000</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11.00000000</td>
<td>0.00000000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Appendix B: DDOP

See Tables 4 and 5.

Table 4
Device description of the tractor.

<table>
<thead>
<tr>
<th>Device</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>DVC1: Valtra T-163 Urea</td>
<td></td>
</tr>
</tbody>
</table>

<Device Description>

Table 5  
Device description of the implement.  
**DVC2**: Junkkari Maestro 4000

```xml
<DVC B="Seed drill controller" A="DVC-22" G="2F00000000E6E65" F="0000000103E6B8" B="000000000001" D="A0884802D40000E1" C="3.74">
  <DPT B="009A" A="10" D="Physical Object Length 1" C="6000" />
  <DPT B="009C" A="11" D="Physical Object Height" C="2200" />
  <DPT B="009A" A="12" D="Physical Object Length 2" C="500" />
  <DPT B="0093" A="14" D="Actual Cultural Practice 1" C="1" />
  <DPT B="009A" A="15" D="Physical Object Length 3" C="500" />
  <DPT B="0093" A="17" D="Actual Cultural Practice 2" C="2" />
  <DPT B="0086" A="20" D="Device Element Offset X" C="4500" />
  <DPT B="0087" A="21" D="Device Element Offset Y" C="0" />
  <DPT B="0088" A="22" D="Device Element Offset Z" C="0" />
  <DPT B="009D" A="23" D="Connector type" C="0" />
  <DPT B="009B" A="24" D="Physical Object Width" C="4000" />
  <DPT B="008D" A="25" D="Working state" C="y" />
  <DPT B="006E" A="30" D="Fertilizer setting" C="0" />
  <DPT B="0007" A="31" D="Fertilizer measurement" C="15" />
  <DPT B="0006" A="40" D="Seed setting" C="8" />
  <DPT B="0007" A="41" D="Seed measurement" C="15" />
  <DET B="1" A="DET21" F="0" E="0" D="Seed drill body" C="1" />
  <DOR A="25" />
  <DOR A="10" />
  <DOR A="74" />
  <DOR A="11" />
</DET>

<DET B="75" A="DET75" F="1" E="1" D="Fertilizer rate controller" C="2">
  <DOR A="74" />
  <DOR A="12" />
  <DOR A="14" />
  <DOR A="25" />
  <DOR A="30" />
  <DOR A="31" />
</DET>

<DET B="78" A="DET78" F="1" E="4" D="Seed rate controller" C="2">
  <DOR A="74" />
  <DOR A="15" />
  <DOR A="17" />
  <DOR A="25" />
  <DOR A="40" />
  <DOR A="41" />
</DET>

<DET B="81" A="DET81" F="1" E="7" D="Connector" C="6">
  <DOR A="20" />
  <DOR A="21" />
  <DOR A="22" />
  <DOR A="23" />
</DET>
</DVC>
```

References


