UNIVERSITÀ DEGLI STUDI DI FOGGIA

DOTTORATO DI RICERCA
IN

"INNOVAZIONE E MANAGEMENT DI ALIMENTI AD ELEVATA VALENZA SALUTISTICA"
(XXX CICLO)

TESI DI DOTTORATO

ICT TOOLS FOR DATA MANAGEMENT AND ANALYSIS TO SUPPORT DECISIONAL PROCESS ORIENTED TO SUSTAINABLE AGRI-FOOD CHAINS

Dottorando: Claudio Zaza

Tutor: Prof. Francesco Contò

ANNO ACCADEMICO: 2016-2017

Index

1. Introduction

The Agri-Food sector is facing global challenges. The first issue concerns feeding a world population that in 2050, according to United Nations projections, will reach 9.3 billion people (United Nations, 2013). The second challenge is the request by consumers for high quality products obtained by more *sustainable*, safely and clear agri-food chains (Grunert, 2005; Seuring and Müller, 2008). In particular, the Sustainable agriculture is a management strategy able to preserve the biological diversity, productivity, regeneration capacity, vitality and ability to function of an agricultural ecosystem, ensuring, today and in the future, significant ecological, economic and social functions at the local, national and global scales, without harming other ecosystems (Lewandowski et al., 1999). The concept of sustainable agriculture derives from the worldwide-accepted definition of sustainable development, introduced by the World Commission on Environment and Development in the Brundtland Report (WCED, 1987) as "development that meets the needs of the present without compromising the ability of future generations to meet their own needs". Therefore, to face the challenge of the sustainable agriculture, farmers need to increase quality and quantity of the production, reducing the environmental impact through new management strategies and tools.

Moreover, the farming enterprises needs to cope with a level of complexity that grows constantly over the time, moving farms from simple food production units, to agricultural businesses with multifunctional service sectors (Parafos et al., 2017). Therefore, to survive and to be sustainable at financial level, a farm must be well managed (Husemann and Novković, 2014), which is a challenging and time-consuming task (Doyle et al., 2000). For the monitoring and managing of the field data collected, farmers needs appropriate information and technologies, even in the Big Data area (Wolfert et al., 2017). The access to time-related information coupled to a careful decision-making is the base for a successful agricultural business (Singh et al., 2008). Nowadays, the farmer's decision-making based on the management of available information, increased in quantity and quality, is facilitated by the huge technological growth in computer hardware and software, that has allowed an effective computer-based support process (Lewis, 1998).

In this context, Farm Management Information Systems (FMISs) play an important role. FMIS has been defined as a ‘‘planned system for collecting, processing, storing, and disseminating data in the form needed to carry out a farm’s operations and functions.” (Sorensen et al., 2010). According to (Fountas et al., 2015), FMISs provide functionalities for field operations management, best practice tools[1], finance, inventory, traceability, reporting, site-specific tools, sales, machinery management, human resource management, and quality assurance. In particular, incorporating the Integrated Pest Management (IPM) framework into FMISs appears as a mandatory instrument to help farmers facing the challenges of sustainable agriculture. The IPM requires the simultaneous use of different crop protection techniques for the control of insects, pathogens, weeds and vertebrates, through an ecological and economic approach (Prokopy, 2003). The aim is to combine different techniques to control pest populations below an economic damage threshold (Chandler et al., 2011). The relevance of the IPM is underlined by the EU that has recognized to IPM a central role to reduce the reliance on the use of conventional pesticides, in the context of the Framework Directive 2009/128/EC (European Parliament, 2009) on Sustainable Use of Pesticides.

Unfortunately, the IPM is manually fulfilled by farmers, and data regarding IPMs of different campaigns and farms are not shared and stored. Therefore, it appears difficult to investigate for the best sustainable practices according to past IPMs. This is an important limitation in the support of the farmers in the decisional process for improving environmental and production performances. In fact, to draw benefits from available databases, farmers needs to collect, process, provide and use data in an efficient way (Parafos et al., 2017).

[1] Best practices tools: based on (complex) agronomy models suggest to farmers some particular technical operations (such as respecting the organic standards) (Fountas et al., 2015).

The needs to acquire data is important not only at farm level, but it regards the whole agri-food sector. In particular, the introduction of new technologies able to monitor the food products along the entire life cycle, from farm to market, could represents an opportunity to meet the growing interest shown by consumers in relation to what happens to the product over the entire supply chain (Brewster et al., 2012). In fact, the consumers consider traceability as a guarantee of quality on the goodness of all treatments performed on the product, such as harvesting, processing, etc. (Costa et al., 2013). Moreover, quality and food safety are parameters that accompany the whole life of the product, since the primary production, in all stages of processing, storage and packaging. According to the definition provided by International Organization for Standardization ISO 9000:2015, the quality adjective could be applied to objects (any entity either conceivable or perceivable) when a set of inherent characteristics (feature that exists in an object) fulfills with a set of requirements. Therefore, the quality of an object depends on a set of characteristics and a set of requirements and how well the former complies with the latter (ISO, 2015a). In addition, other quality certifications were provided by ISO 14001:2015 (ISO, 2015b) and ISO 22000:2005 (ISO, 2005) to address specific topics related to environment and food quality. Finally, it should be mentioned the EC Regulation 178/2002 (European Commission, 2002) on food safety in which, according to the "farm to fork" approach, all stakeholders are responsible for the traceability of products at all stages of production, processing and distribution.

In the stages of post-harvest and post-production is more critical to perform controls due to the risk of a deterioration that affects the work done upstream and degrades a product which initially was of high quality. Keep under control and continuously monitor certain parameters essential for the quality of the product is an activity which, however, still today, is not carried out in a totally efficient manner. Although the current procedures provide to set the best conditions for the processing and to perform a test on individual lots, this system does not provide the certainty that uncontrolled lots are in accordance. The potential economic saving should not be underestimated because sometimes the individual lots are so broad as to cover a working day or a whole order of the customer. To reach this goal is necessary to move completely the traceability management system of the product: not affecting only the origin but involving detailed information on what happened at each stage of the life cycle, consistently and without interruption. Currently, the management systems of traceability and quality are concerned with collecting data at various stages of production, but these data are acquired in discontinuous procedures and grouped in batches. In other words, the data available are restricted to individual production lots which represent quantities greater than those referred to the individual production units. The quality is considered as the average value of more detections, in the best of cases, or as the single value of the significant sample on which are performed the analysis. There is no way to know what happens at the level of single production unit. The information is then distributed and not centralized and "follows" the product along the way acquiring and transmitting the parameters that identify the preservation of quality.

At this regards, it is useful to cite the experience of the CESAR (Certification and Food Safety by RFID) project, which has proposed a monitoring system for food products that allows preserving quality and safety of foods at all stages, from primary production to processing, packaging and storage, through the use of the Radio Frequency Identification (RFID) technology. Often, the existing procedures are weak on identifying and checking quality parameters, since they are generally carried out on random or ex-post. According to these considerations, the project has proposed a new management system of the agri-food traceability, preventive and distributed. It consists in the use of miniaturized and intelligent sensors that can track the single production/batch units, during all stages of the agri-food chain. In particular, it can be constantly acquired the value of environmental and chemico-physical parameters (temperature, humidity, light, pH, O_2, CO_2, ethylene) of the product throughout its lifetime. In this way, it allows to verify if these parameters comply with the production specifications, and to intervene with corrective actions before quality loss. The described traceability system is able to be adapted to different product and supply chains. These innovative technologies prevent from market compliance, improving logistic management and generating costs reduction.

This work is organized as follows: in the background section an overview is given regarding all the studied topics; the section three is dedicated to related works presented in literature; then the objectives have been proposed; the section five is dedicated to the materials and methods adopted; in the following sections the results are exposed, while the section nine concludes the thesis.

For the redaction of this thesis have been used two papers published on international peer reviewed journals and one proceeding paper presented to an international conference, listed below:

1) Zaza, C., Bimonte, S., Faccilongo, N., La Sala, P., Contò, F., & Gallo, C. (2018). BI4IPM: A Business Intelligence System for the Analysis of Olive Tree's Integrated Pest Management. *International Journal of Agricultural and Environmental Information Systems (IJAEIS)*, *9*(1), 16-38.
2) Faccilongo, N., Contò, F., Dicecca, R., Zaza, C., & Sala, P. L. (2016). RFID sensor for agri-food supply chain management and control. *International Journal of Sustainable Agricultural Management and Informatics*, *2*(2-4), 206-221.
3) De Pascale, G., La Sala, P., Faccilongo N., & Zaza, C., (2017) Adopting ICT Tools by Farms in Lucania Region. Proceedings of the 8th International Conference on Information and Communication Technologies in Agriculture, Food and Environment (HAICTA 2017), Chania, Greece, 21-24 September, 2017.

Bibliography

Brewster C., Wolfert S., Sundmaeker H., 2012, Identifying the ICT Challengesof the Agri-Food Sector to Define the Architectural Requirements for a Future Internet Core Platform, eChallenges e-2012 Conference Proceedings Paul Cunningham and Miriam Cunningham (Eds) IMC International Information Management Corporation, 2012 ISBN: 978-1-905824-35-9.

Chandler, D., Bailey, A. S., Tatchell, G. M., Davidson, G., Greaves, J., & Grant, W. P. (2011). The development, regulation and use of biopesticides for integrated pest management. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, *366*(1573), 1987-1998.

Costa, C., Antonucci, F., Pallottino, F., Aguzzi, J., Sarriá, D., & Menesatti, P. (2013). A review on agri-food supply chain traceability by means of RFID technology. *Food and Bioprocess Technology*, *6*(2), 353-366.

Doye, D., Jolly, R., Hornbaker, R., Cross, T., King, R. P., Lazarus, W. F., & Yeboah, A. (2000). Case studies of farmers' use of information systems. *Review of Agricultural Economics*, *22*(2), 566-585.

European Commission. (2002). Regulation (EC) No 178/2002 of the European Parliament and of the Council of 28 January 2002 laying down the general principles and requirements of food law, establishing the European Food Safety Authority and laying down procedures in matters of food safety. *Official Journal of the European Communities, L31,* 1–24.

European Parliament. (2009). Directive 2009/128/EC of the European Parliament and of the Council of 21 October 2009 establishing a framework for Community action to achieve the sustainable use of pesticides. *Off. J. Eur. Union*, *309*, 71-86.

Fountas, S., Carli, G., Sørensen, C. G., Tsiropoulos, Z., Cavalaris, C., Vatsanidou, A., Liakos, B., Canavari, M., Wiebensohn, J., & Tisserye, B. (2015). Farm management information systems: Current situation and future perspectives. *Computers and Electronics in Agriculture*, *115*, 40-50.

Grunert, K. G. (2005). Food quality and safety: consumer perception and demand. *European review of agricultural economics*, *32*(3), 369-391.

Husemann, C., & Novkovic, N. (2014). Farm management information systems: a case study on a german multifunctional farm. *Ekonomika Poljoprivrede*, *61*(2), 441.

ISO (2005) *ISO 22000:2005. Food Safety Management Systems—Requirements for any Organization in the Food Chain*.

ISO (2015a) *ISO 9000:2015. Quality management systems–Fundamentals and Vocabulary*.

ISO (2015b) *ISO 14001:2015 Environmental Management Systems-Requirements with Guidance for Use*.

Lewandowski, I., Härdtlein, M., & Kaltschmitt, M. (1999). Sustainable crop production: definition and methodological approach for assessing and implementing sustainability. *Crop science*, *39*(1), 184-193.

Lewis, T. (1998). Evolution of farm management information systems. *Computers and Electronics in Agriculture*, *19*(3), 233-248.

Paraforos, D. S., Vassiliadis, V., Kortenbruck, D., Stamkopoulos, K., Ziogas, V., Sapounas, A. A., & Griepentrog, H. W. (2017). Multi-level automation of farm management information systems. *Computers and Electronics in Agriculture*.

Prokopy, R. J. (2003). Two decades of bottom-up, ecologically based pest management in a small commercial apple orchard in Massachusetts. *Agriculture, ecosystems & environment*, *94*(3), 299-309.

Seuring, S., & Müller, M. (2008). From a literature review to a conceptual framework for sustainable supply chain management. *Journal of cleaner production*, *16*(15), 1699-1710.

Singh, M., Singh, P., & Singh, S. B. (2008). Decision support system for farm management. *World Academy of Science, Engineering and Technology*, *39*, 346-349.

Sørensen, C. G., Fountas, S., Nash, E., Pesonen, L., Bochtis, D., Pedersen, S. M., Basso, B., & Blackmore, S. B. (2010). Conceptual model of a future farm management information system. *Computers and electronics in agriculture*, *72*(1), 37-47.

United Nations, D. (2013). World population prospects: The 2012 revision.

WCED U (1987) Our common future. World Commission on Environment and Development Oxford University Press.

Wolfert, S., Ge, L., Verdouw, C., & Bogaardt, M. J. (2017). Big Data in Smart Farming–A review. *Agricultural Systems*, *153*, 69-80.

2. Background

This section explores the background related to the topics treated in this work. In particular, the first subsection is dedicated to the innovation system and innovation processes, while the second presents the technology used for the data acquisition, management and analysis introducing RFID technology, FMIS, DW and OLAP and finally the complementarity between OLAP and FMIS. The third subsection is dedicated to the olive crop, treating the economic and the nutraceutical aspects of this crop, while the fourth subsection describes the background information related to IPM approach.

2.1 Innovation system and innovation processes

Innovation is one of the key strategies proposed in literature and economic government policies as a crucial driver of the agri-food sector. Especially in the agricultural sector, it is not only intended as adoption of new technologies, but also the introduction of new practices, techniques and alternative ways of organization and management (markets, labor, land tenure, distribution of benefits, etc.). An innovation system is a network of organizations, companies and individuals with the aim of bringing to market new products, processes and forms of organization; moreover, together with government institutions and policies, they influence the action methods of the supply chain agents (World Bank, 2006). Beyond researchers, extension agents and farmers, an agricultural innovation system consists of all types of public and private actors, such as industry actors, agricultural traders, retailers, policymakers, consumers and NGOs. The system approach recognizes the influential role of institutions (i.e. laws, regulations, attitudes, habits, practices, incentives) in shaping how actors interact (World Bank, 2006). Innovation cannot be seen as a linear approach, in which research only delivers new technology to the agricultural sector, but is the result of a process of networking, interactive learning and negotiation among a heterogeneous set of actors (Leeuwis, 2004; Röling, 2009).

The European Commission in its Europe 2020 Strategy (European Commission, 2010a), attributes to innovation and research a central role to face future challenges. According to the orientations for the "CAP towards 2020” (European Commission, 2010b) innovation plays also as main leader in the European Union agriculture. Analysis of systems of innovation (Freeman, 1995; Lundvall, 1992; Nelson, 1993), and scientific networks (Callon, 1994; Freeman, 1991; Häusler et al., 1994), triple helix model (Etzkowitz and Leydesdorff, 1997, 2000) state that the innovation is becoming more open or distributed over time (Coombs et al., 2003), associated with increasing levels of collaboration and outsourcing (Chatterjee, 1996; Howells, 1999). This phenomenon has led the analysis to investigate more closely the role of the nodes and links in this process (Howells, 2006). The concept is considered to have great potential on adding value to previous agricultural research systems by drawing attention to the actors needed for innovation and growth. This consolidates the role of the private sector and the importance of interactions within a sector, emphasizing the outcomes of technology and knowledge generation (World Bank, 2006). The central role of research and innovation is developed further in one of the seven EU 2020 flagship initiative "Innovation Union" (European Commission, 2010c) which introduces the concept of European Innovation Partnerships (EIP) as a new way to improve innovation.

The EIP aim to foster a competitive and sustainable agriculture and forestry that achieves more from less and works in harmony with the environment (Contò et al., 2012). They help building a competitive primary sector that secures global food availability, diversified products and production, long-term supply of various raw-materials for food and non-food uses, as well as a better allocation of added value across the food chain. Under these conditions, the EIP identify two main objectives: as an indicator to promote the productivity and efficiency of the agricultural sector, aiming at reverse by 2020 the recent downward trend in the increase of productivity; and as an indicator of sustainability of agriculture, aiming at ensure the achievement of a satisfactory level of functionality of soils in Europe by 2020.

Regarding the innovation transfer in agricultural practices, the EIP make use of a number of existing policies as the Common Agricultural Policy (CAP) and the Rural Development Programs (RDP). CAP, in the field of Union research and innovation, finances innovative actions as rural development policy. The RDP are implemented especially at the local, regional or national levels, where innovative actions must be co-financed by the Union policy in the area of research and innovation. Synergies are sought with the opportunities offered by cohesion policy, in particular through regional strategies for innovation and transnational and interregional cooperation programmes (Materia, 2012).

Other key concepts related to the innovation as system and knowledge diffusion can be found on social dynamics and the so-called open innovations. In line with the communication of the European Commission, the rural sectors in Europe call for a review of the links between knowledge production and its use to foster innovation. The new agricultural knowledge and innovation system (AKIS) must adapt to the emerging economic, social and environmental challenges by making the best use of diversity in technologies and innovations that can achieve more with less while respecting the environment. Public innovation stresses the need for social and political changes in the context of rural development and producer-consumer relationships. Social innovation includes collective and creative learning processes, in which actors form different social groups and rural and urban contexts participate (European Commission, 2013). They develop new skills, products and/or practices, as well as new attitudes and values that make a difference in addressing the sustainability challenge in rural societies. Moreover, the necessary skills to achieve new forms of competitive advantage for small and medium-sized enterprises (SMEs) in the agri-food sector as "dynamic capabilities" emphasize the key role of strategic management in adapting and integrating organizational abilities, resources and competences to match with a changing environment (Teece et al., 1997). The role assigned to the social innovation passed thus through the concept of open innovation and its relation with the competitive advantage for the SMEs. The theory of Open Innovation and the business model that derives from it are particularly adaptable to the configuration management of SMEs. Many small companies manage to be innovative only in the moment in which they are able to define, support and give continuity to its competitive advantage. Nevertheless, the likelihood of such optimal conditions is reduced to the high level of risk related to innovation, the high degree of uncertainty about the possible economic returns and the lack of a coherent model of innovation management (Edgett and Kleinschmidt, 2003). However, we must emphasize that economic conditions are forcing even the most entrepreneurial "closed" to consider the possibility to go beyond their boundaries and explore the outside world. Faced with this situation, recent studies in the field of innovation and technology management, explained the potential benefits related to an innovative process of opening to the outside (Gassman, 2006), usually characterized by reduced bureaucracy and greater inclination to risk by administrators, possession of highly specialized knowledge, increased ability to react to rapid changes in the market (Christensen et al., 2005). In SMEs, even more than in the large corporates, being innovative means knowing how to better manage their "strength" competitive. From this point of view, the rapid changes in technology are certainly not helpful, because induce small businesses to activate processes of product development in an ever more quickly and efficiently manner. One way to stimulate this new evolutionary process consists in emphasizing links with actors in the micro and macro business environment, thus creating a knowledge system useful to acquire the dynamic capabilities and to meet the challenges that agricultural companies have to face.

2.2 Data acquisition, management and analysis

2.2.1 RFID technology

The food products quality depends on all the processes that undergo, starting from primary production to the phases of processing and distribution. Continuous controls are needed in post-harvest and post-production phases, in which there is a real risk to deteriorate product quality obtained with great efforts in the field. Therefore, it is necessary a control system that follows the product on whole supply chain in order to ensure his integrity and quality. RFID technology allows constant monitoring of a product along the entire agri-food chain (Kelepouris et al., 2007; Ngai et al., 2008; Shi et al., 2009).

RFID systems consist of three elements: Tag or transponder that is the label to be applied on food products, Reader or transceiver that reads and writes data output from the tag, and Computer that contains the database and analysis software of the data collected (Suhong et al., 2006). RFID active Tags containing sensors able to detect environmental parameters such as pH (Steinberg and Steinberg, 2009), temperature (Jedermann et al., 2009; Amador et al., 2008), humidity (Chang et al., 2007; Abad et al., 2009), shock/vibration (Todd et al., 2009), light (Abad et al., 2009; Cho et al., 2005) and concentrations of gases such as acetaldehyde or ethylene (Vergara et al., 2006). The constant control of these parameters is essential to avoid the loss of product's intrinsic characteristics. This allows, also, to automatically perform a series of checks on each individual product, such as storage time or temperature (Wang et al., 2010), decreasing significantly the noncompliance (Golan et al., 2004). In addition, constant control determines not only an economic advantage, but also a reduction of food losses.

The sensors are in the network and communicate the data in a wireless manner using RFID technique that allows the transmission in two different ways. In the first approach the sensor are queried when they cross special gates, equipped with readers, places corresponding to certain stages of production (at the entrance of the storage area, at the entrance of cold storage, at the beginning of line processing, packaging, shipping, etc.). The second way provides the spontaneous transmission by the sensor when one of the detected parameters exceed certain threshold values. In this phase, the sensor sends a signal to indicate that the product is in an early stage of degradation and action must be taken promptly.

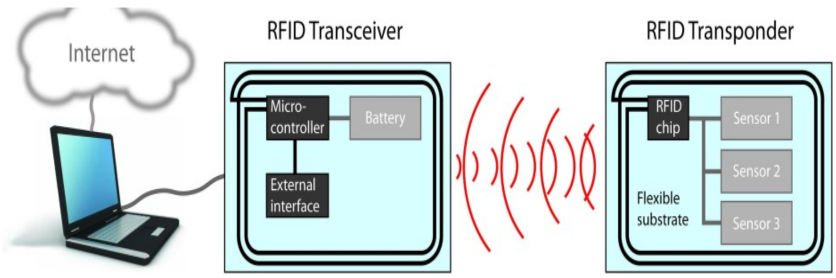


Figure 1: RFID technology

This is very important for the monitoring of food organoleptic parameters and perishables in general, where it is necessary to ensure controlled operational regime. The RFID active tag can monitor the conservation state of product, or report an alarm when one of the detected parameters is not in the desired range without opening the packages and managing the data acquired in order to take the appropriate decisions: remove the product or accelerate the correction treatment. The software platform, by interfacing with the appropriate RFID detectors, will deal with the acquisition of the data collected by the sensors and processing them in order to ensure constant products monitoring during each phase of the production cycle. In each phase, the system will be able to determine the product condition and to identify promptly the situations that may lead to reduction of the quality and healthiness. The technology therefore consists in a system of distributed monitoring of perishable food products using intelligent micro sensors placed on the individual units of product. The system consists, briefly, of two basic elements: a network of distributed sensors and a software platform for the collection and analysis of data and may also be implemented in an integrated form (RFID tags) for the monitoring of the individual products and as Mini-Card (black box) for the monitoring of batches.

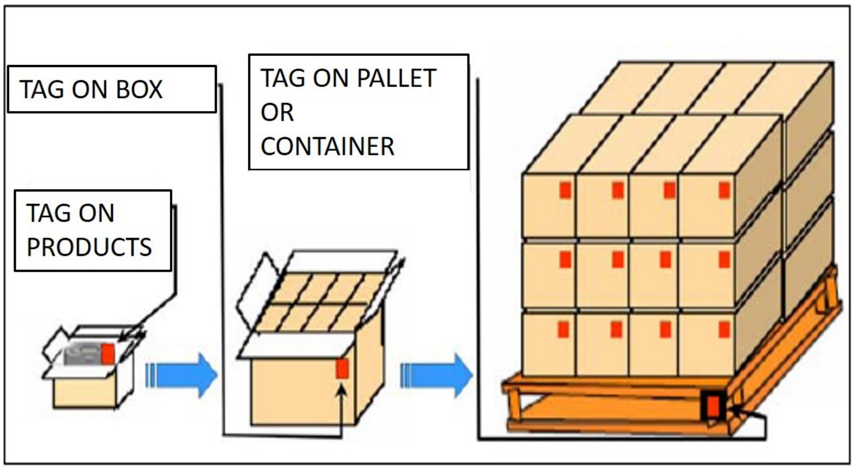


Figure 2: Tag positions

2.2.2 Farm Management Information Systems

The need for keeping data and information permanently, as it may be useful at later times, is a very obvious problem in the modern world and is now facing a large number of people and economic agents. Retention and subsequent use of data can be a valuable resource for making decisions or increasing the company's economic activity. This target has been initially achieved by means of files on computer systems. The limits of traditional data files derive from: redundancy (the same data appear two or more times); incongruence (when information is updated in a store and not in another, or there are different values for the same data); inconsistency (available data is no longer reliable, because no one knows with certainty which of several values is correct). Then these limits have been superseded through the advent of databases, which organize data in an integrated manner through modeling techniques (Codd, 1990). Data are managed on a storage system through dedicated software, with the aim of achieving high efficiency in processing and retrieval of data. While files are anchored to the physical media, databases are independent of the location and the physical structure of the data (Date, 2006).

A database is managed by the Data Base Management System (DBMS), software providing:

- Consistency. The data contained in the records must be meaningful and be usable in user applications;
- Safety. Prevent data from being damaged by accidental and unauthorized operations on database;
- Integrity. Ensure that database transactions performed by authorized users do not cause a loss of data consistency.

Data management performed by a DBMS has these key features:

- Independence from the physical structure of data;
- Independence from the logical structure of data;
- Multiple concurrent users;
- Elimination of redundancy;
- Elimination of inconsistency;
- Ease of access;
- Data integrity;
- Data security;
- Use of languages for database management and query.

Databases used for daily operations are typically named “operational DB” in which data are inserted/updated/deleted with high frequency (Färber et al., 2012). This leads to the need for a sophisticated transaction processing system (being a transaction a batch of operations to be done ALL and in the EXACT sequence defined) which must be always available (otherwise the risk is to block the entire Management Information System). From this comes the term OLTP = On-Line Transactional Processing, in which queries execute transactions that read and write a small number of records from different tables (Harizopoulos et al., 2008). An OLTP system must so guarantee reliable transactions on data, recovery from every possible data fault and data consistency, all within a high level of of "competition" (parallelism/concurrency of accesses) (Schaffner et al., 2008).

FMISs are systems designed for collecting, processing, storing, and disseminating data in the form required to perform operations and functions of farms (Sørensen et al., 2010). These operations including strategic, tactical and operational planning; implementation and documentation; assessment and optimization of the work performed in the fields or on the farms (Kaloxylos et al., 2014). FMISs can be viewed as OLTP systems to analyse spatial data, considering that are used to store and process farm data for everyday farm management (Fountas et al., 2015).

2.2.3 Data Warehouse and OLAP

Data Warehouse (DW) and OLAP systems are Business Intelligence technologies allowing for online analysis of a massive volume of multidimensional data. Moreover, Chaudhuri and Dayal, (1997) have defined the Data warehousing as a pool of *decision support* technologies, aimed at enabling the *knowledge workers* to make better and faster decisions. Warehoused data are stored according to the multidimensional model (Gallo et al., 2010; Kimball and Ross, 2013). Data are organized in dimensions and facts. Dimensions are represented by the analysis axes and are organized into hierarchies (for example, cities, departments and regions). Facts are represented by the analysis subjects and are described by numerical attributes called measures (for example, quantity of sold products). Measures are explored with the OLAP operators, which allow for navigating into the DW. Codd et al., (1993) have defined that OLAP approach consist of several, speculative "what-if" and/or "why" data model scenarios performed within the context of some specific historical basis and perspective. Within these scenarios, the values of key variables or factors are frequently transformed, to reflect potential variances based on the firm application sector (supply, production, sales, marketplace, etc.) and/or other environmental and internal aspects. Then, the gained information is synthesized through animation of the data model, including the consolidation of projected enterprise data according to the several dimensions considered. Therefore, the OLAP approach allows to discriminate new or unanticipated relationships between variables, to detect the parameters essential to manage large amounts of data, to generate an unlimited number of dimensions (based on the firm needs), and to define cross-dimensional conditions and expressions (Codd et al., 1993). Common OLAP operators include Slice, which allows selection of a subset of warehoused data, and Drill, which allows for navigating into hierarchies aggregating measure values with SQL aggregation operators (i.e., MIN, MAX, SUM, AVG, etc.). For example, in a retail application, a OLAP query could be "*What is the total quantity of sold products by year, product and region?*".

2.2.4 Complementarity between OLAP and FMIS

As described by Boulil et al., (2014), several differences exist between OLAP and OLTP systems and, therefore, OLAP and FMISs. OLTP can handle very detailed, non-historical and small amounts of data, whereas OLAP can handle aggregated, historical and large amounts of data. Transactional queries can access only a small number of records, whereas OLAP queries can analyse massive amounts of data. Contrary to OLTP systems, which define data models according to data characteristics and provide queries on the top of these models, in OLAP systems, to grant good performance, data are modelled according to analysis needs and using ad-hoc physical (for example join indexes) and logical data techniques (such as table denormalization).

To conclude, OLAP and FMIS technologies address different categories of problems:

- FMISs are well-adapted for operational tasks and complex analysis on small datasets, and
- OLAP technology is suitable for historical analysis based on the exploration and summarization of massive datasets.

2.3 Olive crop

2.3.1 Economic aspects of olive crops

Olive tree (*Olea europea* L.) is one of the most important crop at global scale. The cultivated surface at world level is equal to 10.27 Mha with a production of 15.4 Mt (FAOSTAT, 2014). The first three countries for Olive crop invested area are Spain (2.52 Mha), Tunisia (1.59 Mha) and Italy (1.15 Mha), while at production level there are Spain (4.56 Mt), Italy (1.96 Mt) and Greece (1.78 Mt) (FAOSTAT, 2014). Therefore, in the Mediterranean Area, Olive crops represents an important element of the cultural heritage with a vital role in the economy and significant social and environmental effects (Clodoveo et al., 2016; Camposeo et al., 2013).

In this work, the attention is focused on the Apulia Region's technical specification for the cultivation of the Olive Tree, which is the most important crop in the selected area. According to the Italian Institute of Statistics (ISTAT, 2016), Apulia is the first olive-producing region in Italy. In particular, in 2016 the agricultural surface invested in Apulia for olive production was equal to 0.379 Mha, thus representing 33% of the corresponding national rate (1.14 Mha). The obtained production was 0.99 Mt, corresponding to 35% of the national amount (2.81 Mt) (ISTAT, 2016). Therefore, in the Apulia region, a large number (227245) of farms work on Olive crop (ISTAT, 2010), which generate a huge amount of Integrated Pest Management (IPM) data. Actually, these data are quite unexploited.

2.3.2 Nutraceutical aspects of olive oil

Extra virgin olive oil (EVOO) is the principal source of fat in the Mediterranean diet (Willet et al, 1995).

Recently, many authors have focused their efforts on the nutraceutical aspects of food, like EVOO, able to preserve a proper state of health, reducing the incidence of some diseases, which are partly caused by the assumption of processed foods that contain high levels of unhealthy nutrients (Marotta et al., 2014). The principal nutraceutical aspects of EVOO are mainly related to the high amount of oleic acid, a balanced content of polyunsaturated fatty acids and the abundance in phenolic substances acting as natural antioxidants (Volpe et al., 2015). Furthermore, others important physiological properties are associated to the phenolic compounds, such as antiallergenic, anti-atherogenic, anti-inflammatory, antimicrobial and the prevention of several cardiovascular and oncologic diseases. Moreover, EVOO consumption have highlighted positive effects for the antioxidant capacity and antioxidant enzyme activity (Oliveras-López et al., 2013, 2014), improving oxidative stress and health-related psychological status (Rus et al., 2017). The high percentage of triglycerides (from 97 to 99%) contained in the EVOO improves sensory and health-promoting properties much more than any other edible oils (Mariotti and Peri, 2014). Finally, an antimicrobial activity on food borne pathogens has demonstrated by EVOO (Hatami et al., 2016).

2.4 Integrated Pest Management

The application of the IPM principles ensures to the farms a better performance, by sustainable and economic points of view. Pretty and Bharucha (2014) have reported the results of 62 IPM project implemented in several countries, developed and not (Pretty & Weibel, 2005). In more than 60% of these projects, the application of the IPM approach allows to reduce the use of pesticides and, at the same time, to increase the yields. The pesticides reduction is predictable, considering that is one of the aim of IPM programmes, while the yield increase could be explained by the improvement of others input, such as higher quality seeds and inorganic fertilizers, obtained using the money saved from phytosanitary products (Pretty and Bharucha, 2014).

Furthermore, in some regions, like Apulia Region, with the Measure 10.1.1 of the RDP 2014 – 2020, the diffusion of certified integrated production was stimulated by an economic incentive. To get incentives, farmers must meet all the requirements included in the crop specific disciplinary provided by the Region. Moreover, producers are also checked by an independent control authority, who verifies that all the farm’s operations are compliant with the IPM technical specifications.

In the Apulia Region, each crop’s specific IPM disciplinary is composed of two parts:

1. *General:* a set of recommendations regarding all the aspects of the crop management, such as fertilization restrictions, soil management, irrigation frequency, etc.

2. *Defense rules:* all the allowed operations (agronomic and chemical) to protect the crop from the pests.

2.4.1 General Part and Defense rules

The General part of the Apulian technical specifications contains all the agronomic indications that the farmers must respect during the whole production cycle, starting from planting to harvesting. In particular, in the general part of the technical specifications regards the requirements described in table 1.

Table 1: Requirements of the IPM general part

Name	Description
Choice of cultivation environment and vocational attitude	It concerns the evaluation of the pedoclimatic characteristics of the cultivation area related to the crop needs. For example, the choice of the cultivation area must be particularly accurate in the case of introduction of new cultivation and/or cultivars in the selected environment.
Preservation of the natural agroecosystem	The biodiversity is considered one of the best resources to reduce the use of chemical substances in the agricultural environment. Some indications included in the technical specifications concern the protection or creation of hedges composed by native species or dry stone walls, etc. In case of perennial crops, like Olive trees, these areas must be at least the 3% of the Utilized Agricultural Area (UAA).
Choice of cultivars and propagation materials	The introduction of verified and safety genetic materials from the phytosanitary point of view, is the best way to control plant diseases caused by viruses, bacteria, fungi, etc. Therefore, all the selected plants must be certified by the phytosanitary aspect, choosing the resistant or tolerant cultivars, taking into account the market requests. Finally, the choice must be done in function of the pedoclimatic conditions of the cultivation area.

Preparation of soil to plant and sowing	All the technical operations performed must be finalized to improve the soil fertility, reducing the soil erosion and degradation.
Crop rotation	This agronomic technique allows preserving the soil fertility, preventing pests and improving the productivity. For the tree crops there is a set of indications in case of re-planting, such as to remove the roots residues, to leave the soil on rest with green manure, etc.
Planting	The timing, the distance and the density of these technical must be chosen in order to maximize the production, and, at the same time, reduce the impact of pests, the use of chemical substances, etc.
Soil management and agronomic practices for weeds control	The different agronomic practices must maintain the soil in good conditions, improve the efficiency of nutrients reducing the leakage losses, prevent the soil erosion, etc. The slope value influences the soil management. With slope more than 30%, it is mandatory for the farmers to apply the cover crop technique, also with spontaneous plant, for the entire year and the whole area, while for plots with inclination between 10 and 30%, the cover crop is mandatory only between the rows. Finally, with slope less than 10% the cover crop are mandatory between the rows, but only for the autumn-winter period
Tree and fructification management	The technical operations for the management of trees, like pruning, pollination, etc., must ensure a correct balance between production and phytosanitary conditions.
Fertilization	The fertilization practices must ensure high qualitative and quantity production, without exceeds, in order to maintain the soil fertility, protect the environment and preventing pests. Therefore, the fertilization plans change in function of the physico-chemical conditions of soil. Finally, each crop has different requirements.

	For Olive tree the quantity of fertilizers change in function of the high or low production and the age of orchard.
Irrigation	This practice must satisfy the water need of crop without exceeds, in order to reduce waste of water, leakage of nutrients and pests development. According to these indications, the farmers must use pluviometric data collected in the fields with agro-meteorological instruments or taken by external sources. Moreover, it is necessary to use irrigation systems that minimize the wastes of water, like drip irrigation. Finally, the irrigation volume changes in function of the soil texture, for example with clay soils the maximum volumes admitted is 450 cubic meters per hectare.
Harvesting	All the adopted techniques must preserve the high quality characteristics of the products. For Olive tree another important parameter is the time of harvesting, because it influences the oil organoleptic characteristics.

The defense rules included in the IPM technical specifications specify all the allowed treatments to protect the crop against the listed pests (Table 2).

Table 2: Excerpt of the IPM defense rules for the protection of the Olive crop in the Apulia region

Pest	Intervention criteria	Active substance and auxiliary	Maximum numb. of intervention	Usage limitations and notes
Olive fruit fly (*Bactrocera oleae*)	Intervention threshold: • Table olives: presence of the first punctures • Oil olives: 10–15% presence of active	*Beauveria bassiana* Dimethoate Imidacloprid	/ 2 1*	

	infestation (sum of eggs and larvae)			* allowed only after flowering

Table 2 shows an excerpt of the IPM defense rules for the protection of the Olive crop in the Apulia region. In the following, each column was described:

- *Pest*: This is a list of the pests of the Olive crop. In the example, the pest is the *Bactrocera oleae* (Rossi), which is most dangerous pest of the Olive tree worldwide (Navrozidis et al., 2000).
- *Intervention criteria*: The intervention threshold represents the limit before the pest causes economic damage to the cultivated crop. In the example, the value is different between table and oil olive cultivars. In the first case, the damage is higher because a single puncture strongly reduces the value of the product, so it is necessary to control the pest immediately. On the other hand, in olives used for oil, a fly presence lower than 10–15%, as a function of different cultivars, is economically acceptable. Therefore, the control starts only when the pest pressure exceeds the aforementioned threshold.
- *Active substance and auxiliary:* This reports the list of authorized Active Substances (ASs) and auxiliaries that could be used to control the pest. In the example, phytosanitary products containing Beauveria bassiana, an entomopathogenic fungus, Dimethoate or Imidacloprid are allowed.
- *Maximum number of treatments*: This is the maximum number of treatments admitted for a specific molecule. For example, products based on *B. bassiana* are not limited, because it is a natural organism without environmental impact. In contrast, Dimethoate and Imidacloprid are limited to two and one treatment per year, respectively, due to the toxicity of the molecules.
- *Usage limitations and notes*: This represents the list of eventual other restriction for specific molecules, such as a temporal restriction. For example, Imidacloprid products are allowed only after the phenological phase of flowering.

2.4.2 Environmental aspects on plants and pests

Taking into account the environmental (i.e. climatic) factors is compulsory to create a real useful instrument to help farmers to analyse the management choices. In fact, the environmental variables, such as air temperature, relative humidity, solar radiation, amount of rain, etc., interact constantly with the plants and the pests, as reported by many authors in literature. In particular, radiation, temperature, and water have a great impact on the growth of plants (Hay and Porter, 2006). As described by Easterling et al., (2007) higher temperatures and altered patterns of precipitation probably can reduce yields of many crops. For example, Gregory et al., (1999) have reported yields reduction in rice of about 5% per °C rise overhead 32 °C.

However, at the same time, the abovementioned environmental factors affect the growth and diffusion of pathogens and pests, such as fungi, insects, bacteria and so on. For example, the presence of the olive fruit fly, one of the most dangerous pest for olive trees in the Mediterranean Basin and the Middle East (Tzanakakis, 2006), is strictly correlated to abiotic environmental conditions (McFadden et al., 1977; Tzanakakis, 2006; Burrack and Zalom, 2008). Absence of olives through the spring and the beginning of summer combined with adverse environmental conditions like high or low temperatures and short photoperiods could cause a facultative reproductive dormancy in *B. oleae* adults (Crovetti et al., 1982; Fletcher et al., 1978). Wang et al., (2009) have highlighted the effects of high temperature on *B. oleae*. In particular, when temperatures rise above 30 °C, the *B. oleae* adults are frenetically active and oviposition is inhibited, while at 35 °C the activity stops. In addition, the larval mortality increases at temperature higher than 30 °C, particularly throughout young larvae (Tsitsipis, 1977), while at 35 °C no flies develop to adult (Genç and Nation, 2008).

Bibliography

Abad, E., Palacio, F., Nuin, M., De Zarate, A. G., Juarros, A., Gómez, J. M., & Marco, S. (2009). RFID smart tag for traceability and cold chain monitoring of foods: Demonstration in an intercontinental fresh fish logistic chain. *Journal of food engineering*, *93*(4), 394-399.

Amador, C., Emond, J. P., & do Nascimento Nunes, M. C. (2009). Application of RFID technologies in the temperature mapping of the pineapple supply chain. *Sensing and Instrumentation for Food Quality and Safety*, *3*(1), 26-33.

Bimonte, S. (2016). Current Approaches, Challenges, and Perspectives on Spatial OLAP for Agri-Environmental Analysis. *International Journal of Agricultural and Environmental Information Systems (IJAEIS)*, *7*(4), 32-49.

Bimonte, S., Pradel, M., Boffety, D., Tailleur, A., André, G., Bzikha, R., & Chanet, J. P. (2013). A New sensor-based spatial OLAP architecture centered on an agricultural farm energy-use diagnosis tool. *International Journal of Decision Support System Technology (IJDSST)*, *5*(4), 1-20.

Burrack, H. J., & Zalom, F. G. (2008). Olive fruit fly (Diptera: Tephritidae) ovipositional preference and larval performance in several commercially important olive varieties in California. *Journal of Economic Entomology*, *101*(3), 750-758.

Callon, M. (1994). Is science a public good? fifth mullins lecture, virginia polytechnic institute, 23 march 1993. *Science, Technology, & Human Values*, *19*(4), 395-424.

Camposeo, S., Vivaldi, G. A., & Gattullo, C. E. (2013). Ripening indices and harvesting times of different olive cultivars for continuous harvest. *Scientia Horticulturae*, *151*, 1-10.

Chang, K., Kim, Y. H., Kim, Y. J., & Yoon, Y. J. (2007). Functional antenna integrated with relative humidity sensor using synthesised polyimide for passive RFID sensing. *Electronics letters*, *43*(5), 259-260.

Chatterji, D. (1996). Accessing external sources of technology. *Research-Technology Management*, *39*(2), 48-56.

Chaudhuri, S., & Dayal, U. (1997). An overview of data warehousing and OLAP technology. *ACM Sigmod record*, *26*(1), 65-74.

Cho, N., Song, S. J., Kim, S., Kim, S., & Yoo, H. J. (2005, September). A 5.1-/spl mu/W UHF RFID tag chip integrated with sensors for wireless environmental monitoring. In *Solid-State Circuits Conference, 2005. ESSCIRC 2005. Proceedings of the 31st European* (pp. 279-282). IEEE.

Christensen, J. F., Olesen, M. H., & Kjær, J. S. (2005). The industrial dynamics of Open Innovation—Evidence from the transformation of consumer electronics. *Research policy*, *34*(10), 1533-1549.

Clodoveo M.L., Dipalmo T., Crupi P., Durante V., Pesce V., Maiellaro I., Lovece, A., Mercurio, A., Laghezza, A., Corbo, F., Franchini, C. (2016). *Comparison Between Different Flavored Olive Oil Production Techniques: Healthy Value and Process Efficiency*. Plant Foods for Human Nutrition, 71 (1), pp. 81-87.

Codd, E. F. (1990). *The relational model for database management: version 2*. Addison-Wesley Longman Publishing Co., Inc..

Codd, E. F., Codd, S. B., & Salley, C. T. (1993). Providing OLAP (on-line analytical processing) to user-analysts: An IT mandate. *Codd and Date*, *32*.

Communication from the European Commission, Europe 2020 Flagship Initiative, Innovation Union. COM (2010) 546.

Contò, F., Fiore, M., La Sala, P., (2012) The role of innovation in the integration processes of the supply chain and in the New Economics of Food, University of Foggia, Italy.

Coombs, R., Harvey, M., & Tether, B. S. (2003). Analysing distributed processes of provision and innovation. *Industrial and Corporate Change*, *12*(6), 1125-1155.

Crovetti, A., Quaglia, F., Loi, G., Rossi, E., Malfatti, P., Chesi, F., Conti, B., Belcari, A., Raspi, A. & Paparatti, B. (1982). Influence of temperature and humidity on the development of the immature stages of Dacus oleae (Gmelin). *Frustula Entomol*, *5*, 133-166.

Date, C. J. (2006). *An introduction to database systems*. Pearson Education India.

Easterling, W. E., Aggarwal, P. K., Batima, P., Brander, K. M., Erda, L., Howden, S. M., Kirilenko, A., Morton, J., Pingali, P., Soussana, J. F., & Tubiello, F. N. (2007). Food, fibre and forest products. *Climate change*, 273-313.

Edgett, S. J., & Kleinschmidt, E. J. (2003). *Best practices in product innovation: What distinguishes top performers*. Ancaster, ON: Stage-Gate.

Etzkowitz, H., & Leydesdorff, L. (1998). THE ENDLESS TRANSITION: A" TRIPLE HELIX" OF UNIVERSITY-INDUSTRY-GOVERNMENT RELATIONS: Introduction. *Minerva*, 203-208.

Etzkowitz, H., & Leydesdorff, L. (2000). The dynamics of innovation: from National Systems and "Mode 2" to a Triple Helix of university–industry–government relations. *Research policy*, *29*(2), 109-123.

European Commission (2010a) Communication from the European Commission, Europe 2020: A strategy for smart, sustainable and inclusive growth, COM (2010) 2020.

European Commission (2010b) Communication from the European Commission, The CAP towards 2020. COM (2010) 672.

European Commission (2013) 1st Meeting of the High Level Steering Board of the EIP Agricultural Productivity and Sustainability, Brussels, 21 February 2013.

FAOSTAT (2014) available at http://www.fao.org/faostat/en/#data/QC (accessed on 20.11.2017)

Färber, F., May, N., Lehner, W., Große, P., Müller, I., Rauhe, H., & Dees, J. (2012). The SAP HANA Database--An Architecture Overview. *IEEE Data Eng. Bull.*, *35*(1), 28-33.

Fletcher, B. S., Pappas, S., & Kapatos, E. (1978). Changes in the ovaries of olive flies (Dacus oleae (Gmelin)) during the summer, and their relationship to temperature, humidity and fruit availability. *Ecological Entomology*, *3*(2), 99-107.

Fountas, S., Carli, G., Sørensen, C. G., Tsiropoulos, Z., Cavalaris, C., Vatsanidou, A., Liakos, B., Canavari, M., Wiebensohn, J., & Tisserye, B. (2015). Farm management information systems: Current situation and future perspectives. *Computers and Electronics in Agriculture*, *115*, 40-50.

Fountas, S., Wulfsohn, D., Blackmore, B. S., Jacobsen, H. L., & Pedersen, S. M. (2006). A model of decision-making and information flows for information-intensive agriculture. *Agricultural Systems*, *87*(2), 192-210.

Freeman, C. (1991). Networks of innovators: a synthesis of research issues. *Research policy*, *20*(5), 499-514.

Freeman, C. (1995). The national innovation systems in historical perspective. *Cambridge Journal of Economics*, *19*(1), 5-24.

Gallo, C., De Bonis, M., & Perilli, M. (2010). Data Warehouse Design and Management: Theory and Practice. *IEEE MEMBERS. Dipartimento di Scienze Economiche, Matematiche e Statistiche Università Di Foggia Largo Papa Giovanni Paolo II*, 1-71121.

Gassmann, O. (2006). Opening up the innovation process: towards an agenda. *R&d Management*, *36*(3), 223-228.

Genç, H., & Nation, J. L. (2008). Survival and development of Bactrocera oleae Gmelin (Diptera: Tephritidae) immature stages at four temperatures in the laboratory. *African Journal of Biotechnology*, *7*(14).

Golan, E., Krissoff, B., Kuchler, F., Calvin, L., Nelson, K., & Price, G. (2004). Traceability in the US food supply: economic theory and industry studies. *Agricultural economic report*, *830*(3), 1-48.

Gregory, P. J., Ingram, J. S., Campbell, B., Goudriaan, J., Hunt, L. A., Landsberg, J. J., Linder S., Stafford Smith M., Sutherst R. W. & Valentin, C. (1999). Managed production systems.

Harizopoulos, S., Abadi, D. J., Madden, S., & Stonebraker, M. (2008). OLTP through the looking glass, and what we found there. In *Proceedings of the 2008 ACM SIGMOD international conference on Management of data* (pp. 981-992). ACM.

Hatami, S., Hatami, S., Mohamadi Sani, A., Mohamadi Sani, A., Yavarmanesh, M., & Yavarmanesh, M. (2016). Chemical composition and antibacterial activity of organic extra virgin olive oil from Iran. *Nutrition & Food Science*, *46*(3), 388-395.

Häusler, J., Hohn, H. W., & Lütz, S. (1994). Contingencies of innovative networks: A case study of successful interfirm R & D collaboration. *Research policy*, *23*(1), 47-66.

Hay, R. K. M. , Porter, J. R. (2006). The physiology of crop yield, 2nd edn. Oxford UK: Blackwell Publishing.

Howells, J. (1999). Research and technology outsourcing. *Technology Analysis & Strategic Management*, *11*(1), 17-29.

Howells, J. (2006). Intermediation and the role of intermediaries in innovation. *Research policy*, *35*(5), 715-728.

Italian Institute of Statics. (2010). *Agricultural Census*. Available at: http://dati-censimentoagricoltura.istat.it/ (accessed 20.03.17).

Italian Institute of Statics. (2016). *Area and Production*. Available at: www.istat.it/it/ (accessed 20.03.17.).

Jedermann, R., Ruiz-Garcia, L., & Lang, W. (2009). Spatial temperature profiling by semi-passive RFID loggers for perishable food transportation. *Computers and Electronics in Agriculture*, *65*(2), 145-154.

Kaloxylos, A., Groumas, A., Sarris, V., Katsikas, L., Magdalinos, P., Antoniou, E., Politopoulou, Z., Wolfert, S., Brewster C., Eigenmann, R., & Terol, C. M. (2014). A cloud-based Farm Management System: Architecture and implementation. *Computers and Electronics in Agriculture*, *100*, 168-179.

Kelepouris, T., Pramatari, K., & Doukidis, G. (2007). RFID-enabled traceability in the food supply chain. *Industrial Management & data systems*, *107*(2), 183-200.

Kimball, R., & Ross, M. (2013). *The data warehouse toolkit: The definitive guide to dimensional modeling*. John Wiley & Sons.

Leeuwis, C. (2004). Communication for Rural innovation: Rethinking Agricultural Extension Blackwell Science Inc.

Lundvall, B. (1992) National Systems of Innovation: Towards a Theory of Innovation and Interactive Learning, London: Pinter Publishers.

Mariotti, M., & Peri, C. (2014). The composition and nutritional properties of extra-virgin olive oil. *The extra-virgin olive oil handbook*, 21-34.

Marotta, G., Simeone, M., & Nazzaro, C. (2014). Product reformulation in the food system to improve food safety. Evaluation of policy interventions. *Appetite*, *74*, 107-115.

Materia, V.C. (2012) European Innovation Partnership (EIP) and opportunities provided to the agri-food SMEs by the network approach, ISCI, Bari.

McFadden, M. W., Kapatos, E., Pappas, S., & Carvounis, G. (1977). Ecological studies on the olive fly Dacus oleae Gmel. in Corfu. I. The yearly life cycle. *Bollettino*.

Nelson, R.R. (1993) National Innovation Systems: A Comparative Analysis, Oxford, Oxford University Press.

Ngai, E. W. T., Moon, K. K., Riggins, F. J., & Candace, Y. Y. (2008). RFID research: An academic literature review (1995–2005) and future research directions. *International Journal of Production Economics*, *112*(2), 510-520.

Oliveras-López M., Berná G., Jurado-Ruiz E., López-García de la Serrana H., Martín F. (2014). Consumption of extra-virgin olive oil rich in phenolic compounds has beneficial antioxidant effects in healthy human adults. *Journal of Functional Foods*, *10*, 475-484.

Oliveras-López M., Molina J. J. M., Mir M. V., Rey E. F., Martín F., De la Serrana H. L. G. (2013). Extra virgin olive oil (EVOO) consumption and antioxidant status in healthy institutionalized elderly humans. *Archives of Gerontology and Geriatrics*, *57*(2), 234-242.

Paraforos, D. S., Vassiliadis, V., Kortenbruck, D., Stamkopoulos, K., Ziogas, V., Sapounas, A. A., & Griepentrog, H. W. (2016). A Farm Management Information System Using Future Internet Technologies. *IFAC-PapersOnLine*, *49*(16), 324-329.

Pretty, J., & Bharucha, Z. P. (2014). Sustainable intensification in agricultural systems. *Annals of botany*, *114*(8), 1571-1596.

Pretty, J., & Waibel, H. (2005). Paying the price: the full cost of pesticides. *The Pesticide Detox. Earthscan, London*, 39-54.

Röling, N. (2009). Pathways for impact: scientists' different perspectives on agricultural innovation. *International journal of agricultural sustainability*, *7*(2), 83-94.

Rus A., Molina F., Ramos M. M., Martínez-Ramírez M. J., del Moral M. L. (2017). Extra virgin olive oil improves oxidative stress, functional capacity, and health-related psychological status in patients with fibromyalgia: A preliminary study. *Biological Research for Nursing, 19*(1), 106-115.

Schaffner, J., Bog, A., Krüger, J., & Zeier, A. (2008). A hybrid row-column OLTP database architecture for operational reporting. In *International Workshop on Business Intelligence for the Real-Time Enterprise* (pp. 61-74). Springer Berlin Heidelberg.

Shi, Y-D., Pan, Y-Y. and Lang, W-M. (2009) The RFID application in logistics and supply chain management, *Research Journal of Applied Sciences*, Vol. 4, No. 1, pp.57–61.

Sørensen, C. G., Fountas, S., Nash, E., Pesonen, L., Bochtis, D., Pedersen, S. M., Basso, B., & Blackmore, S. B. (2010). Conceptual model of a future farm management information system. *Computers and electronics in agriculture*, *72*(1), 37-47.

Steinberg, I. M., & Steinberg, M. D. (2009). Radio-frequency tag with optoelectronic interface for distributed wireless chemical and biological sensor applications. *Sensors and Actuators B: Chemical*, *138*(1), 120-125.

Suhong, L., Visich, J.K., Basheer, M.K. and Zhang, C. (2006) Radio frequency identification technology: applications, technical challenges and strategies, *Sensor Review*, Vol. 23, pp.193–202.

Teece, D. J., Pisano, G., & Shuen, A. (1997). Dynamic capabilities and strategic management. *Strategic management journal*, 509-533.

Todd, B., Phillips, M., Schultz, S. M., Hawkins, A. R., & Jensen, B. D. (2009). Low-cost RFID threshold shock sensors. *IEEE Sensors Journal*, *9*(4), 464-469.

Tsitsipis, J. A. (1977). Effect of constant temperature on the eggs of the olive fruit fly, Dacus oleae (Diptera, Tephritidae). In *Annales de Zoologie Ecologie Animale*.

Tzanakakis, M. E. (2006). *Insects and mites feeding on olive: distribution, importance, habits, seasonal development and dormancy*. Brill Academic Publishers.

Vergara, A., Llobet, E., Ramírez, J. L., Ivanov, P., Fonseca, L., Zampolli, S., Scorzoni A., Becker T., Marco S., & Wöllenstein, J. (2007). An RFID reader with onboard sensing capability for monitoring fruit quality. *Sensors and Actuators B: Chemical*, *127*(1), 143-149.

Volpe R., Stefano P., Massimiliano M., Francesca M., Gianluca S., Federica R. (2015). Healthy fats for healthy nutrition. An educational approach in the workplace to regulate food choices and improve prevention of non-communicable diseases. *High Blood Pressure and Cardiovascular Prevention*, 22 (4), pp. 395-401.

Wang, L., Kwok, S. K., & Ip, W. H. (2010). A radio frequency identification and sensor-based system for the transportation of food. *Journal of Food Engineering*, *101*(1), 120-129.

Wang, X. G., Johnson, M. W., Daane, K. M., & Opp, S. (2009). Combined effects of heat stress and food supply on flight performance of olive fruit fly (Diptera: Tephritidae). *Annals of the Entomological Society of America*, *102*(4), 727-734.

Willett, W. C., Sacks, F., Trichopoulou, A., Drescher, G., Ferro-Luzzi, A., Helsing, E., & Trichopoulos, D. (1995). Mediterranean diet pyramid: a cultural model for healthy eating. *The American journal of clinical nutrition*, *61*(6), 1402S-1406S.

World Bank. (2006). *Enhancing agricultural innovation: how to go beyond the strengthening of research systems*. World Bank.

3. Related Work

This section is dedicated to the analysis of the related works presented in the literature. In particular, the first subsection is an overview of the works focused on the use of the RFID technology for traceability purpose. The second subsection is dedicated to the evolution of FMISs. Finally, in the third subsection an overview of the OLAP solutions adopted in the agricultural field is presented.

3.1 RFID technology for traceability purpose

One of the most important issue in the agri-food sector is the traceability of the products, however the ability to trace an animal, a commodity, a food product or an ingredient and monitor its history in the supply chain either forward (from source to consumer) or backward (from consumer to source) (Hobbs et al., 2007).

The RFID technology has been adopted for the monitoring of several types of food products. Regarding the horticultural sector, many authors have reported several RFID solutions for the smart traceability of the food products. In particular, Vergara et al., (2007) have integrated into RFID readers some gas sensors able to monitor the climacteric conditions during the supply chain phases. These sensors can operate only in a defined temperature range, working as an alarm and checking in an efficient way the conservation stage of several fruits, like apples.

The RFID technology developed by Vergara et al., (2007), was tested by Gandino et al., (2009) to build a traceability system in fruit warehouses.

Ampatzidis et al., (2009) reported two RFID methods to match bins containing harvested fruits with corresponding trees, in order to record the yield for manual fresh fruit harvesting. The first method consist of a long-range RFID reader and a Global Positioning System in differential mode (DGPS) applied to an orchard tractor combined to a passive low-cost RFID tags fixed to the bins. In the second method, the RFID tags were applied to individual trees and to bins, while the DGPS was not used.

Moreover, in Hertog et al., (2008) active RFID labels equipped with temperature sensors were used for the monitoring of Belgian tomatoes from farmers to supply chain. Different simulations on quality change model were performed based on the observed temperature scenarios. Furthermore, several batches of tomatoes were identified and characterised at farm level, and the product was monitored using RFID tags along all the supply chain, including the phases of transport and marketing.

Amador et al., (2009) adopted active tags RFID to monitor the temperature across the pineapple supply chain. The performances of the temperature tags were compared with the results obtained by others conventional technologies, such as probes and sensors, showing a comparable accuracy between the two adopted methods. Nevertheless, at data recovery level, the RFID temperature tags showed better performance than the conventional methods.

Chunxia et al., (2009) discussed the implementation of RFID technology during all the phases of the vegetable supply chain, such as cultivating, processing, storage, transporting and retailing.

Yang and Wang (2012) suggested that the quality traceability of vegetable products might be performed by RFID technology, introducing the basic working principle and the technical characteristics of the proposed method, and analysing the difficulties in the application carrying out different measurements.

The RFID technology could be conveniently applied for the monitoring of the meat products, as argued by many authors in literature. Particularly, Tomeš et al., (2009) proposed to adopt the RFID systems to identify cattle samples associated to a Biotrack database. This solution allows matching each sample ID with all set of biometric identifiers, recorded in the Biotrack database, in order to associate the meat sold units with the animal of origin. Moreover, he argued that the solution based on integrated RFID-Biotrack database system could be adopted instead of barcodes to create a full traceability between the participants.

In Kong et al., (2009) a dedicated RFID architecture was developed, taking into account the safety control required for the meat supply chain, identifying every pig in farm and structuring the information recorded in a farm database.

An RFID tag for carcass was reported in Luo et al., (2011) working with an online reading and writing system to be adopted in meat production lines, able to collect and send the critical information necessary for a good traceability. Moreover, the authors provide commercial meat RFID tags to be applied on carcass and cuttings in the sales stores.

The RFID technology is used also for traceability purpose for dairy products. In particular, an RFID-based technology was adopted by Cai and Liang, (2011) to track the movement of milk packages in China, ensuring information flux from producers to consumers. Moreover, the authors have provided Internet applications usable by the consumers to recover real-time information on the dairy item's production and processing.

Two different types of RFID tags were tested by Pérez-Aloe et al., (2007) for the cheese traceability. In particular, the authors performed several lecture trials in different conditions (i.e., temperature, humidity, corrosive or saline solutions immersions, etc.). The tags readability resulted not affected by the different conditions tested, excepted when metallic materials occurred in the range of the reader.

Later, Varese et al., (2008) explored the possibility to use the RFID technology in dairy cheese production to avoid cases of imitations of Protected Designation of Origin (PDO). Therefore, two solutions based on small-sized tags were provided, comparing their efficiency. The first embedded tag was directly applied on the side of the cheese at the end of the forming process, while the second tag was external applied in a casein plate after the first or second turning over of the cheese. The efficiency test showed that the tags positioning did not affect readability. Moreover, the first solution resulted to be more resistant than the second, during the different processing phases.

3.2 FMIS evolution

In the literature, many publications focus on FMISs. In the 1970s, the first FMIS, specialized in recordkeeping and operation planning, was introduced (Blackie, 1976; Thompson, 1976). Later, Kok and Gauthier (1986) incorporated decision support algorithms for recordkeeping and planning into an FMIS. In California, the Calex system was presented, which combined decision support tools with recordkeeping and planning, including irrigation, pest management, and fertilization applications (Plant, 1989).

The application of precision agriculture (PA) coupled with the advance of more accurate global positioning systems (GPS) has generated a large amount of data and the necessity to create FMISs able to manage this information. To cope with these needs, many authors proposed solutions and architectures for the development of FMISs focused on PA (Fountas et al., 2006; Nikkila et al., 2010). At this regards, Nash et al., (2009a) represented the domain knowledge through the entity-relationship approach, related to the data flows of the crop production system and their planning levels.

Moreover, several FMISs have been designed for the management of machinery, using for the data collection the standard ISOBUS protocol for the tractors and the implements. The difficulties of the data acquisition related to the absence of compatibility between hardware and software was highlighted by Steinberger et al., (2009), who proposed a prototype implementation of an agricultural process. In particular, the task formulation is the result of analysis performed on the agricultural process data acquired from the ISOBUS.

Nash et al., (2009b) proposed to build a geospatial web services to face the compatibility problems.

Tsiropoulos et al., (2013a,b) presented web-based applications for farm machinery with real-time data acquisition in order to record the sub-field spatial variability within field operations and to ensure the communication with autonomous mobile vehicles.

Carli et al., (2014) explore the possibilities of including in the FMISs modules able to analyse the costs of the obtained products based on resources consumed and to suggest more profitable behaviours to the farmers.

Parafos et al., (2016) introduce an FMIS that integrates a farm financial tool based on future internet technologies.

Kaloxylos et al., (2014) propose the implementation of a cloud-based FMIS. The architecture is composed of a local FMIS, connected with the agro-environmental sensors for data collection and a cloud FMIS dedicated to data storage and analysis. The system has been tested for the management of a greenhouse in Greece.

Therefore, Lantzos et al., (2013) presented an FMIS developed for the Android operating system, optimized to manage small farms.

3.3 OLAP in the agricultural field

Recently, the OLAP approach have been used to analyse agricultural data.

Chaudhary et al., (2004) presented a system to analyse the economic factors affecting the yield of cereal crops. The dimensions included in the system are the spatial dimension, the temporal dimension, the type of fertilizers and cereals.

Chaturvedi et al., (2008) investigated factors affecting cotton production, defining 13 data cubes related to the analysis of crops taking into account several aspects, such as soil and socioeconomic resources, using a hybrid design methodology. In this work, the authors have faced the problem related to the multi-granular data, which consists of data sources that could be not all available at the same dimension levels, disaggregating data, when possible, using complex interpolation functions. The provided data cubes could be explored using not only the classical SQL, but also with some advanced aggregations, such as a weighted average for the market price.

A model to estimate the crop performance for a specific area, weather and environmental domain was developed by Gupta et al., (2013). The authors defined a data cube able to investigate the success of a specific crop as a function of irrigation, weather, seeding and the soil type. Moreover, the schema of the data warehouse is detailed, providing some SQL example decisional queries.

In the work of Pestana et al., (2005) a data model to face the problem of integrating spatial and temporal data providing information is proposed to simplify semantic interoperability and data analysis in a Spatial Data Warehouse (SDW) that uniformly manages all data types to study the vineyard production. The model provides a temporal dimension, a spatial dimension, which represents the plots, and others thematic dimensions. The number of grafts and the number of changes of the geometric shape of the spatial phenomenon are the measures included in the model. Finally, the presented model allows investigating the evolution of the production and the parcel.

Deggau et al., (2010) also focused on farm activities and agricultural production. The proposed system enriches SDWs with ontologies enabling to find the more appropriate SDW by a keyword search.

The support of macro-level planning activities is the aim of the OLAP system developed by Nilakanta et al., (2008), which could be used for the analysis of animal and crop resources. For the design of the multidimensional model, the authors taken into account the available data and the decision-maker needs, validating it through a usability test. Therefore, applying a hybrid methodology and considering the data sources and the feedback obtained with the usability test, the authors provided several cubes organized in a constellation schema. The model dimensions are classical, i.e. the product dimension, and complex, such as temporal and spatial dimensions, facing the multi-granular problems as previously described.

The integration of forestry data in a spatiotemporal DW has treated by Miquel et al., (2003), providing a complex multidimensional model, composed by the temporal and spatial dimensions to measure the forest surface. In this work, the authors reported that the main problem concerns the spatial hierarchy, which is dependant over time.

Pinet et al., (2010) adopted the OLAP approach to analyse the usage of agricultural fertilizers. The presented spatio-multidimensional model, based on a UML profile, allows investigating the spread amount per commune and fertilizer.

Bimonte et al., (2013a) during the DispEAU project, which aims at monitoring meteorological data collected from French vineyards, evaluated the effects of humidity and temperature on vineyard production. To reach the objective of the project was designed a Spatial-OLAP (SOLAP) model presenting a spatial dimension, representing plots and farms, and a temporal dimension. The measures, temperature and moisture, are recorded through a wireless sensors network. The model is implemented in an OLAP-GIS architecture using Map4Decision as Web SOLAP client. Adopting a similar SOLAP architecture, the energy consumption of farms was analysed for the wheat crop (Bimonte et al., 2013b) and the milk production (Bimonte et al., 2014).

Bibliography

Amador, C., Emond, J. P., & do Nascimento Nunes, M. C. (2009). Application of RFID technologies in the temperature mapping of the pineapple supply chain. *Sensing and Instrumentation for Food Quality and Safety*, *3*(1), 26-33.

Ampatzidis, Y. G., Vougioukas, S. G., Bochtis, D. D., & Tsatsarelis, C. A. (2009). A yield mapping system for hand-harvested fruits based on RFID and GPS location technologies: field testing. *Precision agriculture*, *10*(1), 63-72.

Bimonte, S., Boulil, K., Pinet, F., & Kang, M. A. (2013a). Design of Complex Spatio-multidimensional Models with the ICSOLAP UML Profile-An Implementation in MagicDraw. In *ICEIS (1)* (pp. 310-315).

Bimonte, S., Pradel, M., Boffety, D., Tailleur, A., André, G., Bzikha, R., & Chanet, J. P. (2013b). A New sensor-based spatial OLAP architecture centered on an agricultural farm energy-use diagnosis tool. *International Journal of Decision Support System Technology (IJDSST)*, *5*(4), 1-20.

Bimonte, S., Chanet, J. P., Capdeville, J., & Lefrileux, Y. (2014). Energetic assessment of dairy activities using OLAP systems. In *Ageng 2014 International conference on agricultural engineering* (Vol. 6, No. 07, pp. 2014-10).

Blackie, M. J. (1976). Management information systems for the individual farm firm. *Agricultural systems*, *1*(1), 23-36.

Cai, W. G., & Liang, B. (2011). Research and implementation of the safety of raw milk traceability management perspective based on RFID technology. *Hubei Agricultural Sciences*, *7*, 053.

Carli, G., Canavari, M., & Grandi, A. (2014). Introducing Activity-Based Costing in Farm Management: The Design of the FarmBO System. *International Journal of Agricultural and Environmental Information Systems (IJAEIS)*, *5*(4), 69-84.

Chaturvedi, K., Rai, A., Dubey, V., & Malhotra, P. (2008). On-line analytical processing in agriculture using multidimensional cubes. *J. Ind. Soc. Agril. Statist*, *62*(1), 56-64.

Chaudhary, S., Sorathia, V., & Laliwala, Z. (2004). Architecture of sensor based agricultural information system for effective planning of farm activities. In *Services Computing, 2004.(SCC 2004). Proceedings. 2004 IEEE International Conference on* (pp. 93-100). IEEE.

Chunxia, W., Weimin, Y., Yingyan, P., & Huiyuan, X. (2009). Research on RFID technology adoption of vegetable supply chain in China. In *Information Technology and Applications, 2009. IFITA'09. International Forum on* (Vol. 2, pp. 669-672). IEEE.

Deggau, R., Fileto, R., Pereira, D., & Merino, E. (2010). Interacting With Spatial Data Warehouses Through Semantic Descriptions. *GeoInfo*, *2010*, 122-133.

Gandino, F., Sanchez, E. R., Montrucchio, B., & Rebaudengo, M. (2009). Opportunities and constraints for wide adoption of RFID in agri-food. *International Journal of Advanced Pervasive and Ubiquitous Computing (IJAPUC)*, *1*(2), 49-67.

Gupta, A. K., & Mazumdar, B. D. (2013). Multidimensional Schema For Agricultural Data Warehouse. *International Journal of Research in Engineering and Technology*, *2*(3), 245-253.

Hertog, M. L. A. T. M., Yudhakusuma, R. F., Snoekx, P., De Baerdemaeker, J., & Nicolaï, B. M. (2006). Smart traceability systems to satisfy consumer expectations. In *XXVII International Horticultural Congress-IHC2006: International Symposium on The Role of Postharvest Technology in the 768*(pp. 407-415).

Hobbs, J. E. (2007). *Identification and analysis of the current and potential benefits of a national livestock traceability system in Canada*. Agriculture and Agri-Food Canada.

Kaloxylos, A., Groumas, A., Sarris, V., Katsikas, L., Magdalinos, P., Antoniou, E., Politopoulou, Z., Wolfert, S., Brewster C., Eigenmann, R., & Terol, C. M. (2014). A cloud-based Farm Management System: Architecture and implementation. *Computers and Electronics in Agriculture*, *100*, 168-179.

Kok, R., & Gauthier, L. (1986). Development of a prototype farm information management system. *Computers and Electronics in Agriculture*, *1*(2), 125-141.

Kong, Q., Zhao, L., Sun, S., Wang, X., & Zhang, M. (2009). Safety control meat supply chain: A case study of SUTAI pigs. In *Innovative Computing, Information and Control (ICICIC), 2009 Fourth International Conference on* (pp. 709-712). IEEE.

Lantzos, T., Koykoyris, G., & Salampasis, M. (2013). FarmManager: an Android application for the management of small farms. *Procedia Technology*, *8*, 587-592.

Luo, Q., Xiong, B., Geng, Z., Yang, L., & Pan, J. (2010). A study on pig slaughter traceability solution based on RFID. In *International conference on computer and computing technologies in agriculture* (pp. 710-720). Springer, Berlin, Heidelberg.

Miquel, M., Bédard, Y., Brisebois, A., Pouliot, J., Marchand, P., & Brodeur, J. (2002). Modeling multi-dimensional spatio-temporal data werehouses in a context of evolving specifications. *International Archives Of Photogrammetry Remote Sensing And Spatial Information Sciences*, *34*(4), 142-147.

Nash, E., Dreger, F., Schwarz, J., Bill, R., & Werner, A. (2009a). Development of a model of data-flows for precision agriculture based on a collaborative research project. *Computers and electronics in Agriculture*, *66*(1), 25-37.

Nash, E., Korduan, P., & Bill, R. (2009b). Applications of open geospatial web services in precision agriculture: a review. *Precision agriculture*, *10*(6), 546.

Nikkilä, R., Seilonen, I., & Koskinen, K. (2010). Software architecture for farm management information systems in precision agriculture. *Computers and electronics in agriculture*, *70*(2), 328-336.

Nilakanta, S., Scheibe, K., & Rai, A. (2008). Dimensional issues in agricultural data warehouse designs. *Computers and electronics in agriculture*, *60*(2), 263-278.

Paraforos, D. S., Vassiliadis, V., Kortenbruck, D., Stamkopoulos, K., Ziogas, V., Sapounas, A. A., & Griepentrog, H. W. (2016). A Farm Management Information System Using Future Internet Technologies. *IFAC-PapersOnLine*, *49*(16), 324-329

Perez-Aloe, R., Valverde, J. M., Lara, A., Carrillo, J. M., Roa, I., & Gonzalez, J. (2007). Application of RFID tags for the overall traceability of products in cheese industries. In *RFID Eurasia, 2007 1st Annual* (pp. 1-5). IEEE.

Pestana, G., da Silva, M. M., & Bédard, Y. (2005). Spatial OLAP modeling: an overview base on spatial objects changing over time. In *Computational Cybernetics, 2005. ICCC 2005. IEEE 3rd International Conference on* (pp. 149-154). IEEE.

Pinet, F., & Schneider, M. (2010). Precise design of environmental data warehouses. *Operational Research*, *10*(3), 349-369.

Plant, R. E. (1989). An artificial intelligence based method for scheduling crop management actions. *Agricultural systems*, *31*(1), 127-155.

Steinberger, G., Rothmund, M., & Auernhammer, H. (2009). Mobile farm equipment as a data source in an agricultural service architecture. *Computers and electronics in agriculture*, *65*(2), 238-246.

Thompson, S. C. (1976). Canfarm—A farm management information system. *Agricultural Administration*, *3*(3), 181-192.

Tomeš, J., Lukešová, D., & Machá, J. (2009). Meat traceability from farm to slaughter using global standards and RFID. *Agricultura Tropica et Subtropica*, *42*(3), 98-100.

Tsiropoulos, Z., Fountas, S., Gemtos, T., Gravalos, I., & Paraforos, D. (2013). Management information system for spatial analysis of tractor-implement draft forces. In *Precision agriculture'13* (pp. 349-356). Wageningen Academic Publishers, Wageningen.

Varese, E., Buffagni, S., & Percivale, F. (2008). Application of RFId technology to the agroindustrial sector: Analysis of some case studies. *Journal of Commodity Science, Technology and Quality*, *47*(1-4), 171-190.

Vergara, A., Llobet, E., Ramírez, J. L., Ivanov, P., Fonseca, L., Zampolli, S., Scorzoni, A., Becker, T., Marco, S., & Wöllenstein, J. (2007). An RFID reader with onboard sensing capability for monitoring fruit quality. *Sensors and Actuators B: Chemical*, *127*(1), 143-149.

Yang, F., & Wang, T. (2012). The application of radio frequency identification technology in vegetable quality traceability system. In *Systems and Informatics (ICSAI), 2012 International Conference on* (pp. 1356-1359). IEEE.

4. Objective of the thesis

The main objectives of this PhD thesis regard the integration of ICT tools in the agri-food sector.

In particular, during the first year the RFID technology, to acquire data along the agri-food supply chain, has been tested.

After that, the student moved his attention on tool for the data manipulation and management at farm scale.

To analyse the collected farm information, two data models, OLTP and OLAP, have been developed with techniques of data design. Therefore, the data models developed have been integrate in a system able to evaluate if the farm operations recorded adhere to the Apulia region olive crop IPM technical specification.

In conclusion, a climatic model has added to system, in order to take into account this fundamental aspect for a better support during the decision making phase.

To attempt the aforementioned objectives, the following contributions have been proposed in this work:

1. Evaluation of two proposed RFID models, compared with other traceability technology, such as the barcode, in order to analyse this technology;
2. Definition and implementation of a structured Data Base in order to obtain a FMIS, called Agri-Food Chain – Web Data Management (AFC-WDM), working on a Web 2.0 environment on multiple devices (PCs, tablets, smartphones) in the model SaaS—Software as a Service;
3. Definition and implementation of a Business Intelligence (BI) system, called BI4IPM, which combines the developed FMIS and OLPT approach, for the IPM general indications, with DW and OLAP system, focused on the IPM defense rules;
4. Integration into BI4IPM system a climatic model that allows comparing over different campaigns, and farms, which can be geographically distributed over a large area;
5. Evaluation, means by a survey, of the adoption level of ICT tools by the farmers of Lucania region.

5. Materials and Methods

This section introduces the materials and methods adopted for the PhD research work. The first subsection is dedicated to the ICT design methodologies and technologies adopted, introducing the spiral model, the ICSOLAP tool and the technological materials used for the FMIS implementation. The second subsection presents the agronomic materials that includes the definition of GSI and the data used for the case study. Finally, the third subsection concludes this chapter reporting the survey definition introducing the identified questions and the methods adopted for the data analysis.

5.1 ICT Design Methodologies and Technologies

5.1.1 Design Methodologies

The methodologies used for the modelling activities are the spiral model, for the development of the structured database dedicated to the collection of the farm management data, and a UML profile, called ICSOLAP, for the OLAP data model.

5.1.1.1 The spiral model

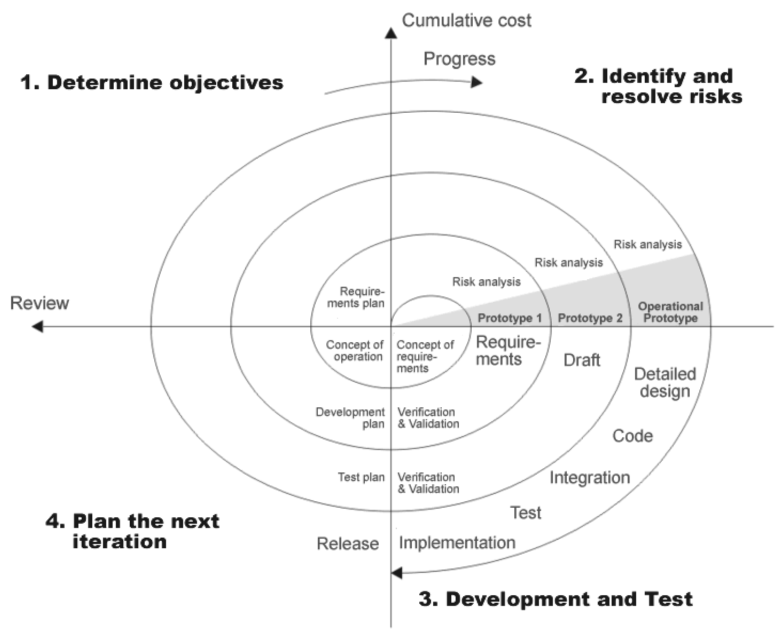


Figure 3: The spiral model (from Boehm, 1988)

The spiral model (Boehm, 1988) (Figure 3) are structured in four phases:

1. Determine objectives: the objectives and the needs to be reached by the development of the new DB have been defined, based on previous analysis. In this stage is necessary a close collaboration between the ICT and agronomic experts.
2. Evaluate alternatives, and identify and resolve risks: the ICT expert analyse all the objectives and the request and release a prototype.
3. Develop and test: the agronomic domain expert and the ICT expert work together to evaluate and test the prototype. A list of new requirements has been obtained in this stage.
4. Plan the next iteration: the two experts analyse the new requirements, derived from the previous stage and start with the design of the new iteration of the cycle.

The spiral model allows a rapid development of more and more complete versions of the software due to the implementation of a new prototype in each cycle (Atzeni and De Antonellis, 1993). Furthermore, the spiral method follows the entire product lifetime. In fact, every single change to the software “awakens" the spiral, starting a new cycle with a new prototype, and so on (Abiteboul et al., 1995).

To improve the efficiency of the model, a panel composed by two ICT experts, two domain experts, an agronomist and an agricultural economist, and two farmers, have strictly collaborated in each phases of the cycle, evaluating all the prototypes developed.

5.1.1.2 The ICSOLAP tool

ICSOLAP (Boulil et al., 2015) is a UML profile for the conceptual design of spatial DW (SDW) that allows the conceptual representation of classical and advanced aspects of spatio-multidimensional modeling, such as multiple and complex hierarchies, etc. The profile defines a stereotype or a tagged value for each multidimensional element, for example, <<Fact>> for the fact and <<TemporalAggLevel>> for temporal dimension levels.

5.1.2 Technological Materials

The technological materials adopted for AFC-WDM are:

- The open source relational DBMS MySQL used for the data storage;
- The AJAX (Asynchronous JavaScript and XML) (Garrett, 2005) technology adopted to manage a Web 2.0 flexible and user-friendly web interface accessible from any browser and operating system. The AJAX technology is a combination of JavaScript and Document Object Model manipulation, along with asynchronous server communication able to ensure a high level of user interactivity.

For the OLAP system, integrated in BI4IPM, the technological materials used are:

- The open source DBMS PostgreSQL is adopted for data storage.
- The OLAP server Mondrian is adopted for the implementation of the OLAP operators (Drill-Down, Roll-Up, etc.), and access grants. Mondrian is an Open Source Business Analytics engine written in JAVA used to perform multidimensional queries against warehoused data, using the MDX (multidimensional expressions) query language. It reads from SQL and other data sources and aggregates data in a memory cache. Mondrian is also adopted for high performance, interactive analysis of large or small volumes of information, parsing the MDX language into SQL to retrieve answers to dimensional queries, advanced calculations using the calculation expressions of the MDX language, etc.
- The OLAP Client JRubik is used to visualize and explore the warehoused data.

5.2 Agronomic Materials

5.2.1 Growing Season Index

The Growing Season Index (GSI) is adopted to compare the farms indirectly by a climatic point of view.

The (GSI) is a model developed for the prediction of plant phenology in response to low temperatures, evaporative demands and photoperiod, applicable at global scale (Jolly et al., 2005). The GSI can predict not only the beginning and the ending of growing season, but also the canopy status during the year without *a priori* knowledge of the vegetation or climate. In particular, for each variables minimum and maximum threshold limits have been defined, assuming that phenological activity varied linearly from inactive (0) to unconstrained (1).

The low temperatures influence many biochemical processes (Levitt, 1980). For example, water in the xylem of some trees can freeze at temperatures below -2 °C (Zimmerman, 1964). The selected range varied between a lower minimum temperature threshold of -2 °C (TMMin) and an upper threshold of 5 °C (TMMax), according with Larcher & Bauer (1981). The minimum temperature index (iTMin), has been created as follows:

$$iT_{Min} = \begin{cases} 0 & if\ T_{Min} \leq T_{MMin} \\ \frac{T_{Min}-T_{MMin}}{T_{MMax}-T_{MMin}}, & if\ T_{MMax} > T_{Min} > T_{MMin} \\ 1 & if\ T_{Min} \geq T_{MMax} \end{cases} \quad (1)$$

where *iTMin* is the daily indicator for minimum temperature bounded between 0 and 1; *TMin* is the observed daily minimum temperature in degrees Celsius; TM_{Min} = -2 °C and TM_{Max} = 5 °C.

Partial to complete stomatal closure (Mott & Parkhurst, 1991), leaf development rate reduction (Salah & Tardieu, 1996), induction of the shedding of leaves (Childes, 1989), and slowing down or stopping cell division (Granier & Tardieu, 1999) could be caused by water stress. To estimate the water needs, the authors selected an index of the evaporative demand, the atmosphere Vapor Pressure Deficit (VPD). According with the literature, they set the VPD_{min} at 900 Pa, because VPD values less than should exert little effect on stomata, while upper threshold VPD_{max} has been fixed at 4100 Pa, because generally greater values of VPD are sufficient to force complete stomatal closure, even when the soils are moist (Osonubi & Davies, 1980; Tenhunen et al., 1982). The VPD index (iVPD) was derived as follows:

$$iVPD = \begin{cases} 0 & if\ VPD \geq VPD_{Max} \\ 1 - \frac{VPD-VPD_{Min}}{VPD_{Max}-VPD_{Min}}, & if\ VPD_{Max} > VPD > VPD_{Min} \\ 1 & if\ VPD \leq VPD_{Min} \end{cases} \quad (2)$$

where *iVPD* is the daily indicator for VPD limited between 0 and 1; VPD is the observed daily VPD in Pascals; VPD_{Min} = 900 Pa and VPD_{Max} = 4100 Pa.

The photoperiod is a consistent annual climatic indication for plant, because it does not change during the years at a specific location. Many authors have exposed that the photoperiod is strictly connected to both leaf flush and leaf senescence all over the world (Njoku, 1958; Rosenthal and Camm, 1997; White et al., 1997; Hakkinen et al., 1998; Partanen et al., 1998; Borchert and Rivera, 2001). Moreover, a strong interaction between photoperiod and temperature that could limit foliar phenology exist, and temperature changes are ineffective without corresponding photoperiod changes (Partanen et al., 1998). The authors set up the minimum threshold at 10 hours, because they assumed that equal or lesser values completely limited canopy development, while the upper limit was fixed at 11 hours, assuming that equal or higher values allows canopies to develop unconstrained. The photoperiod index (iPhoto) was calculated as follows:

$$iPhoto = \begin{cases} 0 & if\ Photo \leq Photo_{Min} \\ \frac{Photo - Photo_{Min}}{Photo_{Max} - Photo_{Min}}, & if\ Photo_{Max} > \ Photo > Photo_{Min} \\ 1 & if\ Photo \geq Photo_{Max} \end{cases} \quad (3)$$

where *iPhoto* is the daily photoperiod indicator bounded between 0 and 1; Photo is the daily photoperiod in seconds; $Photo_{Min}$ = 36 000 seconds (10 h) and $Photo_{Max}$ = 39 600 sseconds (11 h).

The product of the individual daily indicators for minimum temperature, VPD, and photoperiod forms a single metric, the GSI, which is a daily indicator of the relative constraints to foliar canopy development or maintenance due to climatic limits. GSI is continuous and restricted between 0 (inactive) and 1 (unconstrained). The daily metric (*iGSI*) is calculated as follows:

$$iGSI = \ iT_{Min} \times iVPD \times iPhoto \quad (4)$$

The daily GSI is then calculated as the 21-day moving average of the daily *iGSI*, to avoid reaction to short-term changes in environmental conditions (Lieberman, 1982).

Orlandi et al., (2013) have presented a new version of the GSI proposed by Jolly et al., (2005). The main difference is that the “original” GSI was developed and tested for the assessment of canopy foliar dynamics on different vegetal species at global scale, while the second version is adapted to the olive trees and tested in several locations in the Mediterranean basin.

The variables took into account by Orlandi et al., (2013) to adapt the model are explained in the following section.

As for the original GSI, iT_{min} is derived by equation (1), but T_{MMin} is set up on 0° C and T_{MMax} is fixed at 7 °C (relation 5).

$$iT_{Min} = \begin{cases} 0 & if\ T_{Min} \leq T_{MMin} \\ \frac{T_{Min}-T_{MMin}}{T_{MMax}-T_{MMin}}, & if\ T_{MMax} > T_{Min} > T_{MMin} \\ 1 & if\ T_{Min} \geq T_{MMax} \end{cases} \quad (5)$$

where *iTMin* is the daily indicator for minimum temperature bounded between 0 and 1; *TMin* is the observed daily minimum temperature in degrees Celsius; TM_{Min} = 0 °C and TM_{Max} = 7 °C.

Moreover, *iPhoto* is calculated with the same relation (3) and thresholds proposed by Jolly et al., (2005).

In this version of GSI, the water requirements are estimated with the potential evapotranspiration (ETP) instead of VPD. Therefore, the *iETP*, that substitute *iVPD*, derived by the following relation:

$$iETP = \begin{cases} 0 & if\ ETP \geq ETP_{Max} \\ 1 - \frac{ETP-ETP_{Min}}{ETP_{Max}-ETP_{Min}}, & if\ ETP_{Max} > ETP > ETP_{Min} \\ 1 & if\ ETP \leq ETP_{Min} \end{cases} \quad (6)$$

where *iETP* is the daily index for ETP bounded between 0 and 1; *ETP* is the daily evapotranspiration value (mm d^{-1}) estimated with the Priestley–Taylor method (Priestley and Taylor, 1972); ETP_{Min} = 2 mm d^{-1} and ETP_{Max} = 5 mm d^{-1}.

The daily solar radiation is also included in the Mediterranean GSI. The daily index *iRad* is obtained with the following relation:

$$iRad = \begin{cases} 0 & if\ Rad \leq Rad_{Min} \\ \frac{Rad-Rad_{Min}}{Rad_{Max}-Rad_{Min}}, & if\ Rad_{Max} > Rad > Rad_{Min} \\ 1 & if\ Rad \geq Rad_{Max} \end{cases} \quad (7)$$

where *iRad* is the daily index for the solar radiation included between 0 and 1; *Rad* is the daily solar radiation value (MJ m^{-2} d^{-1}) estimated with the Campbell and Donatelli model; Rad_{Min} = 5 MJ m^{-2} d^{-1} and Rad_{Max} = 12 MJ m^{-2} d^{-1}.

Finally, the daily metric *iGSI* is calculated with the product of the new variables:

$$iGSI = iT_{Min} \times iPhoto \times iETP \times iRad \quad (8)$$

Again, the daily GSI is calculated as the 21-day running average of the daily *iGSI*.

5.2.2 Case study data

The data adopted in this work could be divided in two section:

- Climatic data, used to calculate the GSI distributions of the Apulian city of Bari, for two period, January 2016 to December 2016 and January 2017 to October 2017;
- Agronomic phytosanitary data, used to test the OLAP system integrated in BI4IPM.

5.2.2.1 Climatic data

The climatic data regarding the Apulian city of Bari are extracted from the Apulia Regional Agency for the Prevention and Protection of the Environment (ARPA-Puglia, 2017). The raw data taken from ARPA-Puglia are measures recorded every 30 minutes for Minimum and Maximum Temperature (°C) and the amount of rain in mm (Table 3).

Table 3: ARPA-Puglia raw data related to the 3 January 2017

Date	Time	TMIN (°C)	TMAX (°C)	Rain (mm)
03/01/2017	00:00:00	7.0	7.3	0.0
03/01/2017	00:30:00	6.7	7.1	0.0
03/01/2017	01:00:00	7.0	7.6	0.0
03/01/2017	01:30:00	7.0	7.6	0.0
03/01/2017	02:00:00	6.6	7.2	0.0
03/01/2017	02:30:00	6.2	6.7	0.0
03/01/2017	03:00:00	6.3	7.0	0.0
03/01/2017	03:30:00	6.7	7.7	0.0
03/01/2017	04:00:00	7.7	9.1	0.0
03/01/2017	04:30:00	9.1	9.5	0.0
03/01/2017	05:00:00	9.5	10.0	0.0
03/01/2017	05:30:00	9.7	10.0	0.0
03/01/2017	06:00:00	9.1	9.8	0.0
03/01/2017	06:30:00	8.7	9.2	0.0
03/01/2017	07:00:00	8.6	8.9	0.0
03/01/2017	07:30:00	8.7	9.2	0.0
03/01/2017	08:00:00	9.0	9.2	0.0
03/01/2017	08:30:00	9.0	9.5	0.0
03/01/2017	09:00:00	9.2	9.3	0.0
03/01/2017	09:30:00	9.2	9.7	0.0
03/01/2017	10:00:00	9.6	10.7	0.0
03/01/2017	10:30:00	10.8	11.3	0.0
03/01/2017	11:00:00	11.2	11.6	0.0
03/01/2017	11:30:00	11.5	13.1	0.0
03/01/2017	12:00:00	13.0	13.7	0.0

03/01/2017	12:30:00	13.5	13.9	0.0
03/01/2017	13:00:00	13.5	13.9	0.0
03/01/2017	13:30:00	13.4	13.9	0.0
03/01/2017	14:00:00	13.1	13.5	0.0
03/01/2017	14:30:00	13.0	13.4	0.0
03/01/2017	15:00:00	11.9	13.0	0.0
03/01/2017	15:30:00	11.2	12.0	0.0
03/01/2017	16:00:00	10.2	11.2	0.2
03/01/2017	16:30:00	9.6	10.3	0.2
03/01/2017	17:00:00	9.5	9.9	0.0
03/01/2017	17:30:00	9.4	9.5	0.0
03/01/2017	18:00:00	9.4	9.6	0.0
03/01/2017	18:30:00	9.5	9.7	0.0
03/01/2017	19:00:00	9.5	9.8	0.0
03/01/2017	19:30:00	9.5	9.8	0.0
03/01/2017	20:00:00	9.5	9.6	0.0
03/01/2017	20:30:00	9.6	9.8	0.0
03/01/2017	21:00:00	9.4	9.8	0.0
03/01/2017	21:30:00	9.0	9.5	0.0
03/01/2017	22:00:00	8.8	9.1	0.0
03/01/2017	22:30:00	8.7	8.8	0.0
03/01/2017	23:00:00	8.5	8.7	0.0
03/01/2017	23:30:00	7.9	8.4	0.0

In Table 3 is presented an excerpt of the dataset taken from ARPA-Puglia, related to the 3 January 2017. For each day, from the 1st January 2016 to 31st October 2017, the lowest (T_{min}) and the highest values of temperature were extracted, while for the amount of rain all the single measurement were summed. Therefore, in the provided example (Table 3), the data extracted are T_{min} = 6.2 °C, T_{max} = 13.9 °C and *Rain* = 0.4 mm. The mean imputation method (Acuna and Rodriguez, 2004), at month scale, was used for the treatment of the missing value. The obtained data T_{min}, T_{max} and *Rain* per day are used to feed the software RadEst version 3.00 (Donatelli et al., 2003), in order to estimate the daily solar radiation (*Rad*) and potential evapotranspiration (*ETP*), according with Orlandi et al., (2013). In particular, the solar radiation data were estimated by the Campbell-Donatelli model (Donatelli and Campbell, 1998), while the Priestley–Taylor method (Priestley and Taylor, 1972) was used to estimate the ETP values. Finally, the daily photoperiod (*Photo*) was estimated by the latitude and the yearday (Monteith & Unsworth, 1990) for the city of Bari.

5.2.2.2 Agronomic phytosanitary data

Data used to test the system concerns two real Apulian farms, involved by the Department of Economics of University of Foggia during the development of the AFC-WDM (Agri-Food Chain Web Data Management) FMIS. Then, the farms were renamed respectively *"Bimonte"* and *"Zaza"* for privacy reasons.

Moreover, in order to test the system after the integration of the GSI model are used the phytosanitary data derived from others four real Apulian farms located in the Bari department and renamed for privacy reason. The data extracted from the farms 1 and 3 are related to the 2016-2017 campaign, while for the farm 4 the data are related to the campaign 2017-2018. Finally, the data taken from the farm 2 are related to the two analysed campaigns, 2016-2017 and 2017-2018. In Table 4 is showed an example of phytosanitary data taken from the farm 2 and related to the campaign 2017-2018. In particular, the information recovered and loaded in the OLAP system are the date of a phytosanitary treatment, the location at plot level, the active substance used to control the pest specified in the last column.

Table 4: farm 2 phytosanitary data related to the campaign 2017-2018

Date	Plot	A. S.	Pest
15/03/2017	P1_2	Dodine	*S. oleagina*
28/04/2017	P1_2	Trifloxystrobin - Tebuconazolo	*C. gloesporioides*
13/06/2017	P1_2	Dimethoate	*P. oleae*
12/09/2017	P1_2	Copper	*S. oleagina*

5.3 Survey definition and analysis

5.3.1 Identified Questions

The questionnaire is composed by twenty-two Questions (Qs) in total. The survey is divided in two parts: i) *General Information*, starting from the Q1 to the Q7, regarding the general aspects of the farms involved in the survey. Based on the Q7 reply, regarding the use of ICT tools, the questionnaire foresees the second section dedicated to the ii) *Farms using ICT* (from Q8 to Q22), or it ends in case of negative answer. In the second section there are set of Qs dedicated to analyse what are the most used ICT tools applied for the farm management and the impact that these technologies could have on the decrement of agronomic input and manpower employed and on the production increasing.

Following the Qs are described:

Q1) Legal status: possible answers (*partnership*; *capital company*; *others*).

Q2) Time of Constitution: possible answers (*less than five years; between five and ten years; more than ten years*).

Q3) Farmer's Age: possible answers (*less than thirty-five years; between thirty-five and fifty years; more than fifty years*).

Q4) Utilized Agriculture Area (UAA): possible answers (*less than ten hectares; between ten and fifty hectares; more than fifty hectares*).

Q5) Crop Type: possible answers (*Tree crops, herbaceous crops, mixture crops*).

Q6) Income: possible answers (*between 0 and 50000€; between 50001 and 120000€; between 120001 and 250000€; between 250001 and 500000€; between 500001 and 1000000€; more than 1000000€*).

Q7) Do you use ICT tools? Possible answers (*yes or no*). If the reply is positive, the farmer answers the demands from Q8 to Q22, in contrary the questionnaire ends.

Q8) What type of Management Tools do you use? Possible answers (*none; tools for Farm's notebook; tools for warehouses' management; tools for management of balance sheet; tools for management of invoicing; Enterprise Resource Management; others*). Multiple answers are allowed.

Q9) What type of Software for Data Management do you use? Possible answers (*none; software for data storing; software for market analysis; Decision Support System software; software to analyse the costs; others*). Multiple answers are allowed.

Q10) Do you use tools for Precision Agriculture? Possible answers (*yes or no*). If the reply is positive, the farmer answers the demand Q11, in contrary the Q17.

Q11) Do you use environmental sensors? Possible answers (*yes or no*). If the reply is positive, the farmer answers the demand Q12, in contrary the Q13.

Q12) Why do you use environmental sensors? Possible answers (*Fertilization, Phytosanitary treatments, Weeding, Irrigation, Sowing, Soil management*). Multiple answers are allowed.

Q13) Do you use Unmanned Aerial Vehicle (UAV or drones)? Possible answers (*yes or no*). If the reply is positive, the farmer answers the demand Q14, in contrary the Q15.

Q14) Why do you use UAV? Possible answers (*Fertilization, Phytosanitary treatments, Weeding, Irrigation, Sowing, Soil management*). Multiple answers are allowed.

Q15) Do you use Satellite Data? Possible answers (*yes or no*). If the reply is positive, the farmer answers the demand Q16, in contrary the Q17.

Q16) Why do you use Satellite Data? Possible answers (*Fertilization, Phytosanitary treatments, Weeding, Irrigation, Sowing, Soil management*). Multiple answers are allowed.

Q17) Do you use External Data Sources? Possible answers (*yes or no*). If the reply is positive, the farmer answers the demand Q18, in contrary the Q19.

Q18) What types of Data do you research? Possible answers (*Agro-Meteorological, Market, Legal aspects, Phytosanitary bulletin, Others*). Multiple answers are allowed.

Q19) What type of tools do you think is the most useful? Possible answers (*External Data Sources, Enterprise Resource Planning, Software for Data Management, Precision Agriculture tools*).

Q20) Since you started to use ICT tools, do you have detected a reduction in the use of agronomic inputs (pesticides, fertilizers, water, etc.)? To what extent? Possible answers (*None; between 0 and 5%; between 6 and 10%; between 11 and 20%; more than 20%*).

Q21) Since you started to use ICT tools, do you have detected a reduction of employed manpower? To what extent? Possible answers (*None; between 0 and 5%; between 6 and 10%; between 11 and 20%; more than 20%*).

Q22) Since you started to use ICT tools, do you have detected an increment of production? To what extent? Possible answers (*None; between 0 and 5%; between 6 and 10%; between 11 and 20%; more than 20%*).

For the cluster analysis presented in the next section a subset of the total variables was used (Table 5).

Table 5: Variables used in the cluster analysis. The code is associated to a single answer in the cluster analysis.

Question	Name and abbreviation	Answers	Code
Q1	Legal Status	*Partnership*	1
		Capital company	2
Q3	Farmer's Age (Age)	*Less than thirty-five years*	1
		Between thirty-five and fifty years	2
		More than fifty years	3
Q4	Utilized Agriculture Area (UAA)	*Less than ten hectares*	1
		Between ten and fifty hectares	2
		More than fifty hectares	3
Q5	Crop Type	*Tree crops*	1
		Herbaceous crops	2
		Mixture crops	3
Q6	Income	*Between 0 and 50000€*	1
		Between 50001 and 120000€	2
		Between 120001 and 250000€	3
		Between 250001 and 500000€	4
		Between 500001 and 1000000€	5
Q7	Do you use ICT tools? (ICT)	*Yes*	1
		No	2

5.3.2 Data analysis methods

The collected data through the survey have been analysed using clustering analysis. For obtaining groups featured by homogeneous parameters, it has been resorted to considering k-means cluster method. The analysis returned acceptable results setting two clusters. The choice of selecting two clusters it was possible due to:

- the k-means clustering method can be applied with both supervised and un-supervised methodology (Wagstaff et al., 2001);
- three clusters not returned acceptable results.

In general, k-means is a method born as un-supervised. Therefore, processing machine automatically calculates the least distances, respecting the set threshold between features (Zhang et al., 1996). The goal aims to evaluate if the distances are such to consider the minimum sum of the squared error (SSE) within each groups (Likas et al., 2003). The formula of the SSE is the following:

$$SSE = \sum_{k=1}^{K} \sum_{\forall x_i \in C_k} \|x_i - \mu_k\|^2$$

where C_k is the set of grouped data in cluster k; μ_k is the vector mean of cluster k. Using un-supervised method, it was found that the clusters were two. Nevertheless, for being in line with Diedern et al., (2003), the scope was to find three groups to be labelled as innovators, early adopters and laggards. The test not achieved the goal and it was tried with a supervised method setting three clusters. In turn, the test not succeeded due to the cluster two and three presented identical features (the reason why un-supervised method returned two clusters). Hence, it has been chosen to apply a supervised method selecting two clusters. At this stage, the test succeeded and results were accepted to fulfill the two groups theory (innovators and remaining groups) issued by Dimara and Skuras, (2003), Sauer and Zilberman, (2012), Stefanides and Tauer, (1999).

Then again, it has been attempted to go through the data, analyzing data through a boxplot to summarize the frequencies. The analysis comes from intersection of selected variables. It has chosen to fix variables for creating groups and, to this extent, UAA and age have been selected. Within each group, it has been investigated how the presence of ICT tools is bridged to the incomes and legal status.

Bibliography

Abiteboul S., Hull R., Vianlu V. (1995). Foundations of Database, Addison-Wesley, Reading, Massachussetts.

Acuna, E., & Rodriguez, C. (2004). The treatment of missing values and its effect in the classifier accuracy. In *Proceedings of the Meeting of the International Federation of Classification Societies (IFCS)* (pp. 639-47).

ARPA-Puglia (2017) available at http://www.arpa.puglia.it/web/guest/serviziometeo (accessed on 10.11.2017).

Atzeni P., De Antonellis V. (1993). Relational Database Theory, Benjamin-Cummings, Menlo Park, California.

Boehm, B. W. (1988). A spiral model of software development and enhancement. *Computer*, *21*(5), 61-72.

Borchert, R., & Rivera, G. (2001). Photoperiodic control of seasonal development and dormancy in tropical stem-succulent trees. *Tree Physiology*, *21*(4), 213-221.

Boulil, K., Le Ber, F., Bimonte, S., Grac, C., & Cernesson, F. (2014). Multidimensional modeling and analysis of large and complex watercourse data: an OLAP-based solution. *Ecological Informatics*, *24*, 90-106.

Diederen, P., Van Meijl, H., Wolters, A., & Bijak, K. (2003). Innovation adoption in agriculture: innovators, early adopters and laggards. *Cahiers d'Economie et de Sociologie Rurales*, *67*, 29-50.

Dimara, E., & Skuras, D. (2003). Adoption of agricultural innovations as a two-stage partial observability process. *Agricultural Economics*, *28*(3), 187-196.

Donatelli, M., & Campbell, C. S. (1998). A simple model to estimate global solar radiation. In *Proc. of the 5th ESA Congress-Nitra, The Slovak Republic* (pp. 133-134).

Donatelli, M., Bellocchi, G., & Fontana, F. (2003). RadEst3. 00: software to estimate daily radiation data from commonly available meteorological variables. *European Journal of Agronomy*, *18*(3), 363-367.

Garrett, J. J. (2005). Ajax: A new approach to web applications.

Häkkinen, R., Linkosalo, T., & Hari, P. (1998). Effects of dormancy and environmental factors on timing of bud burst in Betula pendula. *Tree Physiology*, *18*(10), 707-712.

Jolly, W. M., Nemani, R., & Running, S. W. (2005). A generalized, bioclimatic index to predict foliar phenology in response to climate. *Global Change Biology, 11*(4), 619-632.

Larcher, W., & Bauer, H. 1. (1981). Ecological significance of resistance to low temperature. In *Physiological plant ecology I* (pp. 403-437). Springer Berlin Heidelberg.

Levitt, J. (1980). *Responses of Plants to Environmental Stress, Volume 1: Chilling, Freezing, and High Temperature Stresses*. Academic Press..

Lieberman, D. (1982). Seasonality and phenology in a dry tropical forest in Ghana. *The Journal of Ecology*, 791-806.

Likas, A., Vlassis, N., & Verbeek, J. J. (2003). The global k-means clustering algorithm. *Pattern recognition*, *36*(2), 451-461.

Mann, H. B., & Whitney, D. R. (1947). On a test of whether one of two random variables is stochastically larger than the other. *The annals of mathematical statistics*, 50-60.

Monteith J. L., Unsworth M. H. (1990). *Principles of Environmental Physics*. Edward Arnold, New York, NY.

Mott, K. A., & Parkhurst, D. F. (1991). Stomatal responses to humidity in air and helox. *Plant, Cell & Environment*, *14*(5), 509-515.

Njoku, E. (1958). The photoperiodic response of some Nigerian plants. *Journal of the West African Science Association*, *4*, 99-111.

Orlandi, F., Garcia-Mozo, H., Dhiab, A. B., Galán, C., Msallem, M., Romano, B., Abichou, M., Dominguez-Vilches, E. & Fornaciari, M. (2013). Climatic indices in the interpretation of the phenological phases of the olive in mediterranean areas during its biological cycle. *Climatic change*, *116*(2), 263-284.

Osonubi, O., & Davies, W. J. (1980). The influence of plant water stress on stomatal control of gas exchange at different levels of atmospheric humidity. *Oecologia*, *46*(1), 1-6.

Partanen, J., Koski, V., & Hänninen, H. (1998). Effects of photoperiod and temperature on the timing of bud burst in Norway spruce (Picea abies). *Tree physiology*, *18*(12), 811-816.

Priestley, C. H. B., & Taylor, R. J. (1972). On the assessment of surface heat flux and evaporation using large-scale parameters. *Monthly weather review*, *100*(2), 81-92.

Rosenthal, S. I., & Camm, E. L. (1997). Photosynthetic decline and pigment loss during autumn foliar senescence in western larch (Larix occidentalis). *Tree Physiology*, *17*(12), 767-775.

Salah, H. B. H., & Tardieu, F. (1996). Quantitative analysis of the combined effects of temperature, evaporative demand and light on leaf elongation rate in well-watered field and laboratory-grown maize plants. *Journal of Experimental Botany*, *47*(11), 1689-1698.

Sauer, J., & Zilberman, D. (2012). Sequential technology implementation, network externalities, and risk: the case of automatic milking systems. *Agricultural Economics*, *43*(3), 233-252.

Sautot, L., Bimonte, S., Journaux, L., & Faivre, B. (2015, April). Dimension Enrichment with Factual Data During the Design of Multidimensional Models: Application to Bird Biodiversity. In *International Conference on Enterprise Information Systems* (pp. 280-299). Springer, Cham.

Stefanides, Z., & Tauer, L. W. (1999). The empirical impact of bovine somatotropin on a group of New York dairy farms. *American Journal of Agricultural Economics*, *81*(1), 95-102.

Tenhunen, J. D., Lange, O. L., & Jahner, D. (1982). The control by atmospheric factors and water stress of midday stomatal closure in Arbutus unedo growing in a natural macchia. *Oecologia*, *55*(2), 165-169.

Wagstaff, K., Cardie, C., Rogers, S., & Schrödl, S. (2001, June). Constrained k-means clustering with background knowledge. In *ICML* (Vol. 1, pp. 577-584).

White, M. A., Thornton, P. E., & Running, S. W. (1997). A continental phenology model for monitoring vegetation responses to interannual climatic variability. *Global biogeochemical cycles*, *11*(2), 217-234.

Zhang, T., Ramakrishnan, R., & Livny, M. (1997). BIRCH: A new data clustering algorithm and its applications. *Data Mining and Knowledge Discovery*, *1*(2), 141-182.

Zimmermann, M. H. (1964). The effect of low temperature on ascent of sap in trees. *Plant physiology*, *39*(4), 568.

6. Results on RFID data acquisition

The CESAR project aim is to create a constant control during all the supply-chain, using miniaturized and smart sensors (RFID technology) able to tracking the individual product/batch unit, in order to improve the quality and safety of the same products. Data are collected through the distributed detection of all the interested parameters on the single product unit, such as individual packages, boxes, bin or pallet. The specifically monitored parameters are environmental (temperature, humidity, light) and physical (pH, O_2, CO_2, ethylene). The real time detection of the aforementioned values during the entire supply chain allows to constantly verifying the quality and safety level of the monitored products.

In order to reach the objectives of the project, two options are provided: I) active Tag; II) Black box.

The first is composed by a chip, on which the sensors are integrate, and an antenna able to receive radio-frequency waves. The data are captured by specific reader and transmitted to PC for the processing phase. The Tag is an adhesive made in different shapes and sizes depending on the nature of the product to be traced. The RFID tag is design to be thrown away together with the package of the monitored product, because it is integrated in the packaging. RFID tags are compose by:

- The sensors network, each with its own interface (front-end) and conditioning electronics. For each of them the emitted signal will be amplified and filtered;
- The multiplexer queries the sensors and stores the received data in a memory that will be read, on command via the antenna by the reading apparatus;
- The SAR ADC12-bit analog to digital converter;
- A microcontroller that controls the multiplexer;
- The memory.

Active tags are powered by an integrate 1.2V paper battery. The charge ensures the monitoring of the product for two years, two weeks of continuous observation or 1350 logging events. The active RFID tag is in the form of small adhesive label, allowing the direct application on the product, the package or the transport pallet, including uncomfortable points. The information contained in the tag are extracted by a fixed or portable reader, which does not need an optical contact with the label and works in extremely short time. Finally, an antenna designed in function of the tag reading distance and the characteristics of the transponder, performs the collection and transmission of the radio signals, to and from the reader.

The Black Box is the discrete configuration of the technology that uses electronic "off the shelf" and not integrated on a single chip. It is made as a mini-card used for the batch monitoring, improving convenience and management. The Black box, at the end of product life, is not eliminated by the packaging, but it is subsequently reconfigured and reused. It is characterised by the flexible structure and the wide variety of sensors can be equipped. The card with on-board sensors records abnormal events and send the data captured to the PC via Bluetooth protocol. The Black box configuration consist of:

- Battery to power sensors and to record all the events in memory with an operating capacity of ten years;
- Ultra-compact dimensions (12x12x1.4 mm) MSP430F1612 microcontroller from Texas Instruments, which works in power-saving mode;
- Sensors of temperature, acceleration in 3 axes, of light, of humidity, pH and gases concentration;
- EEPROM memory with a digital timer and calendar mode connected to an I2C bus;
- Wireless communication with the reader via an integrated Bluetooth module to interface with any device;
- Low-power components that activate the microcontroller only when an event occurs (through interrupt mode);
- Reprogrammable through JTAG interface for adding new sensors;

Current size (85x144x40 mm), considering the use of two AA alkaline batteries; It is possible through an integrated implementation of the main circuits reach the size of 32x80x15 mm, making the Black Box compact and suitable for multiple purposes.

7. Results on Data modelling and analysis

This section is dedicated to the results obtained through the data modelling and analysis. Particularly, the first subsection introduces the FMIS data model for the general part of IPM, the second explores the OLAP model for IPM defense rules, while the third regards the interactions between the two data models described in the previous sections. Therefore, the fourth subsection is dedicated to implementation of BI4IPM tool providing a set of example queries performed. Finally, the fifth subsection explores the integration of the GSI model in the OLAP system.

7.1 FMIS data model for the General part of IPM

The data model used to verify adherence to the IPM general requirements from Table 1 is shown in Figure 4. The proposed data model is an OLTP database conceptual model, included in a UML package (i.e., OLTP package), since (i) requirements concerning technical operation need to be checked daily (for example *Fertilization*), which is well-adapted to FMIS OLTP-based technologies and (ii) the other requirements concern data that do not change daily or that exist in other terms that represent static information (such as slope data). Therefore, these data are not suitable for OLAP models and so are not used for historical analysis.

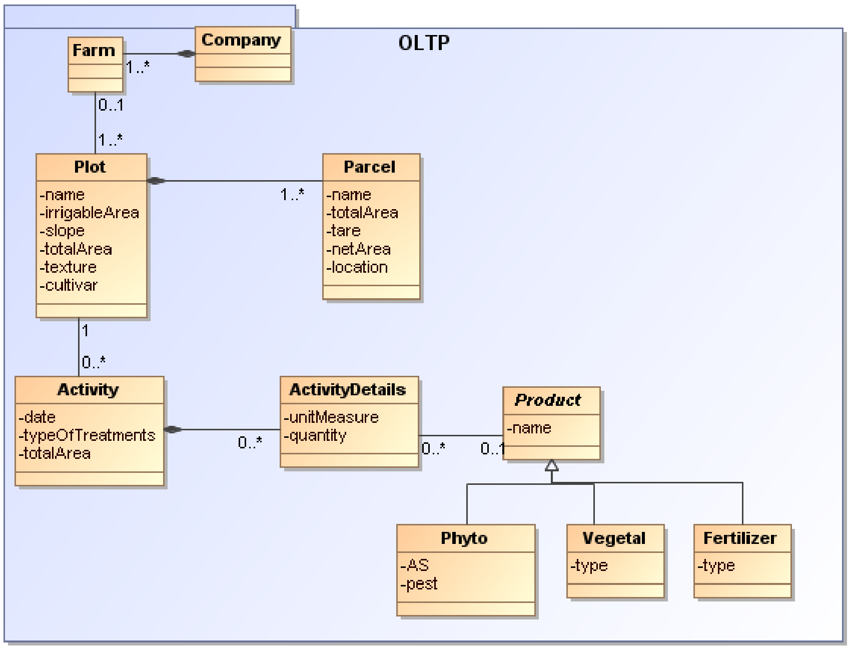


Figure 4: Database model for the IPM general part

The data model is composed of some entities described as follows.

"*Parcel*" represents the farm's parcels. It is described by the following attributes:

- Name: This is the name given by the user.
- Location: This represents the official identification codes provided by the Italian cadastre, divided in Sheet, Parcel and Subdivision.
- Total_Area: This is the total area of the parcel.
- Net_Area: This value is calculated by subtracting Tare from the Total_Area.
- **Tare**: This represents the area that is not able to be cultivated.

This attribute is used to verify adherence to the requirement ***Preservation of the natural agroecosystem*** (Table 1), verifying that the sum of these areas at the plot level is at least the 3% of the UAA.

The "*Plot*" entity is constituted of the following attributes:

- Name: This is the name that the user gives to the plot.
- Irrigable_Area: This is the total area declared irrigable by the user.
- Total_Net_Area: This is calculated summing the values related to the attribute Net_Area of the parcels assigned to the plot (UAA).
- **Slope**: This represents the slope of the plot.

This field is important because, as described in the requirement ***Soil management and agronomic practices for weeds control*** (Table 1), the slope value influences the soil management and agronomic practices for weed control.

- **Texture**: This represents the type of soil texture.

This value is related to the ***Fertilization*** and ***Irrigation*** requirements (Table 1).

- **Cultivar**: This represents the variety of Olive cultivated.

This attribute is used to verify adherence to the ***Choice of cultivars and propagation materials*** and ***Crop rotation*** requirements (Table 1), in cases of new implants.

As shown in Figure 1, a plot is composed of one or more parcels.

A plot belongs to zero or one farm, while one or more farms belong to one company.

The entity "*Activity*" represents all the technical operations provided on a plot (i.e., an activity is associated with one plot). This entity is described by the following attributes:

- Date: This is the date of the operation.
- Total Area: This is the surface involved in the treatment.
- Type of treatments performed:
  - *Soil management operation*: This represents the different soil management operations—such as plowing, harrowing, etc.—involved in the ***Preparation of soil to plant and sowing*** and ***Soil management and agronomic practices for weeds control*** requirements (Table 1).
  - *Harvesting*: Related, in particular, to the time of picking, this attribute is related to the ***Harvesting*** requirement (Table 1).
  - *Pruning*: This concerns data used to check that the ***Tree and fructification management*** requirement is satisfied (Table 1).
  - *Irrigation*: This regards the irrigation strategy.
  - *Planting*: This operation is recorded only in case of a new implant.
  - *Fertilization*: This represents all the performed fertilizations.
  - *Phytosanitary treatments*: This represents all the performed phytosanitary operations.

Apart from the soil management practices, pruning and harvesting activities do not need additional information, an activity is described (composition association of Figure 4) by a set of particular operations ("*Activity details*").

Therefore, the entity "*Activity_details*" is composed of the following:

- *Unit_Measurement*: This represents the unit of measurement, for example liters, kilograms, number of plants, etc.
- *Quantity*: This represents the amount of water, plant, fertilizer or phytosanitary product used for the treatment.

An "*Activity_details*" is associated with zero or one "*Product*".

For example, no further information is needed for irrigation practices, and the registered data are used to verify adherence to the ***Irrigation*** requirement (Table 1).

For planting operations, the abstract entity "*Product*" is made up by:

- ***Name***: This is the name of the chosen cultivar, for example, Coratina or Frantoio.

In the entity "*Vegetal*" the data recorded is as follows:

- ***Type***: This is categorized as grafted or self-propagated.

These attributes regard the ***Planting*** requirement (Table 1).

Furthermore, in case of fertilization, the information registered in "*Product*" is as follows:

- ***Name***: This is the commercial denomination of fertilizer.

In the entity "*Fertilizer*" the data recorded is as follows:

- ***Type***: This is categorized as organic or mineral.

These data are used to check the compliance with the ***Fertilization*** requirement (Table 1).

Finally, for phytosanitary treatments, the information recorded in the abstract entity "*Product*" is as follows:

- ***Name***: This is the commercial denomination of the pesticide used.

In the entity "*Phyto*" are registered the following data:

- ***Active Substances***: This is the molecule contained in the pesticide, for example Copper, Deltamethrin, etc.

- ***Pest***: This is the treatment's target pest defined using the scientific name, for example *B. oleae*, *S. oleagina*, etc.

A part of these data is used to feed the OliveTreeIPM system, as described in the next section.

7.2 OLAP model for IPM Defense rules

Contrary to the IPM general part, for defense rules, an historical analysis is necessary. Therefore, an OLAP model was proposed (Figure 5) for the analysis of Olive crop IPM defense rules, using the UML Profile ICSOLAP (Boulil et al., 2015).

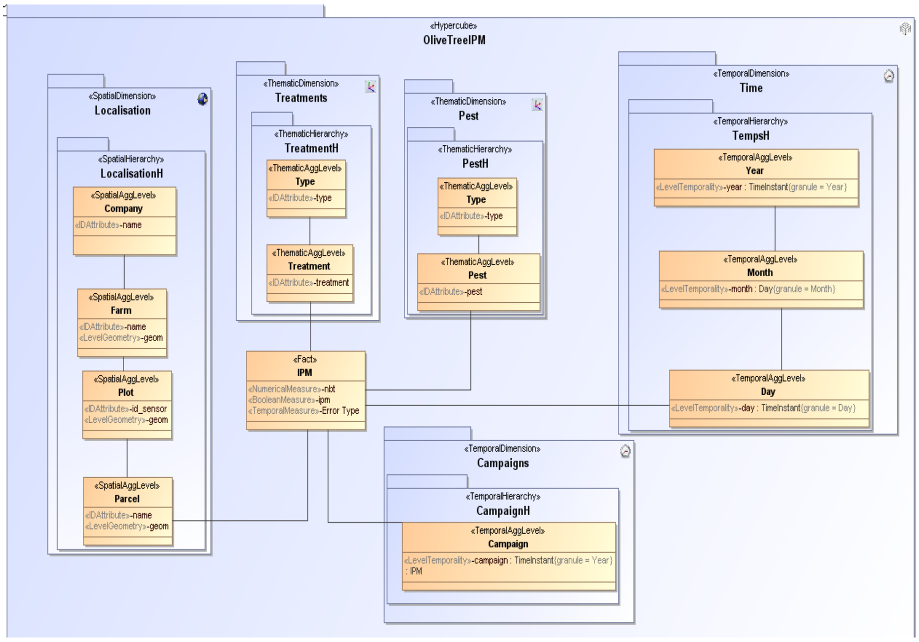


Figure 5: Olive fruit IPM multidimensional model

The multidimensional model of the "OliveTreeIPM" package presents five dimensions:

- *Time*: This is the temporal dimension with the hierarchy day<month<year.
- *Campaign*: This represents the crop campaign (for example 2009–2010).
- *Company*: This describes the organization of the farms. It groups parcels in plots and farms in companies. This hierarchy allows an analysis of all the treatments performed at different spatial scales. For example, the same farm could treat two plots of the same crop in different ways during a single campaign to evaluate the best management strategy for the future years.

- *Pests*: This represents the IPM pest (first column of Table 2). Pests are organized by type, such as weed, fungi, insects, bacteria and so on. Therefore, the pest has a specific scientific name, e.g., *B. oleae* (Table 2), *Spilocaea oleagina*.
- *Treatments*: This represents the different types of interventions allowed in the IPM technical specification. They are grouped by type, such as Agronomic (e.g., Pruning) and Chemical (e.g., Dimethoate, Imidacloprid, etc.).

Measures and their aggregations are as follows:

- *Number of treatments*: This represents the number of performed treatments (for example, the application of Dimethoate during the 2009–2010 campaign). This measure is aggregated along all dimensions using the sum function.
- *Respect of IPM*: This is a Boolean measure describing when the performed action adheres to the IPM rules. For example, the usage of the Glyphosate 30.8% is not allowed. Therefore, the measure value for this treatment is "NO". *Respect of IPM* is aggregated using the AND operator.
- *Error type*: This describes the treatment according to the IPM rules. It is an enumeration value and is described as follows:
    - *Allowed*. The treatment is allowed (for example, the usage of Copper to control the pest *S. oleagina*).
    - *A.S. not Allowed*. The treatment is not allowed (for example, the usage of Glyphosate 30.8%).
    - *Exceeded nb inter*. The number of treatments exceeds the maximum amount of treatment allowed (fourth column, Table 2); for example, a third intervention based on Dimethoate to control *B. oleae* is not allowed.
    - *A.S. not allowed in time*. The treatment performed is not allowed in that time period. For example, Imidacloprid products are forbidden before the phenological phase of flowering.
    - *Errors*. This represents when more errors occur, so it a drill-down operation needed to obtain the specific error type for a single treatment.

For the aggregation of *Error type,* a user-defined aggregation was used that returns *Errors* when more than one of *A.S. not Allowed* OR *Exceeded nb inter* OR *A.S. not allowed in time* are aggregated; otherwise, the model outputs *Allowed.*

Thanks to this model, decision makers can trigger different kinds of questions:

(i) *Technical specification*: Compliance with the Olive Crop IPM technical specifications provided by the Apulia Region. In this case, the model allows investigations into adherence to the technical specification. These queries can be grouped in two classes:

- *"inter-farms"* queries allow treatments among different farms to be compared using historical data. Examples of these kinds of queries are as follows:
    - **Q1:** *"Visualize all chemical treatments of farms"* (Figure 10).
    - *"Visualize the treatments for Bactrocera oleae of farms that adhere to the technical specifications*".
    - *"Visualize all the not allowed chemical treatments of farms that do not adhere to the technical specifications"*.
    - ...

For privacy reasons, deeper information is available only for the user's farm. In this way, the model allows light inter-farm analysis and deeper intra-farm analysis, as shown below.

- *"intra-farm"* queries allow matching treatments between different plots or parcels of the same farm:
    - **Q2:** *"Visualize all the chemical treatments performed for plots on the Bimonte farm"* (Figure 11).
    - *"Visualize the treatments that do not adhere to the technical specification for plots of the farm"*.
    - *"Visualize the allowed chemical treatments for all the parcels"*.
    - ...

In this way, it is possible to analyse compliance with technical specifications at a lower scale, allowing the identification of exactly where a wrong treatment was performed.

(ii) *Management*: detailed analysis of crop protection strategies. In this group of queries, detailed analyses are provided at the intra-farm level. In particular, the different crop protection strategies are investigated from different points of view, such as the number of treatments for plot or parcel, the type of interventions for each pest, etc. Farmers and/or decision makers can answer queries such as the following:

- **Q3:** *"Show the number and type of treatments performed at the farm level for each pest"* (Figure 13).
- *"Show the number of agronomical and chemical treatments performed to control fungi pests"*.
- *"Visualize all the treatment for all the pests at plot scale"*.
- …

This type of query enables the user to carry out different types of analysis, for example, comparing different strategies of crop protection realized in different plots or evaluating which pest caused more damage for better control in the next campaign.

Finally, these queries can be triggered at different spatial, temporal and thematic levels using the dimension hierarchies and the OLAP operators.

For example, it is possible to do all of the following:

- **Q4:** *"Visualize the number and type of treatments performed to control pests grouped per type at the plot scale"* (Figure 14), applying a drill-down operator on the spatial dimension.
- *"Visualize the allowed chemical treatments per parcel and day"*.
- …

7.3 Interactions between data models

As shown in Figure 6, the two presented data models (the OLTP package in Figure 4 and the OliveTreeIPM in Figure 5) interact.

In particular, the OliveTreeIPM uses a part of the data recorded in the OLTP subsystem (*Load*) and analysed them using the OLAP approach.

Moreover, the connection between the two presented models allows feedback to be obtained from the OLAP system to explore the data collected in the general part of the system (*Zoom of OLAP*). For example, if an OLAP technical specification *"intra-*

farm" query highlights some not allowed treatments, the proposed model allows the zooming in on the single instances recorded in the OLTP system to obtain more detailed information regarding the performed phytosanitary operation, such as the dose per hectare, the reason for the intervention, the phenological phase, etc., which are not used in the DW.

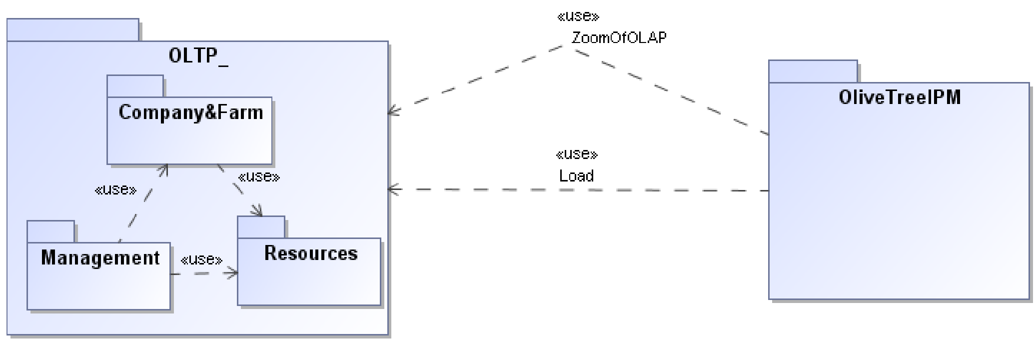


Figure 6: Top-level representation of data models

7.4 BI4IPM Implementation and analysis

BI4IPM is a Business Intelligence system composed of the following (Figure 7):

- AFC-WDM (Agri-Food Chain Web Data Management) that is an OLTP-based FMIS, which allows, among other things, the implementation of the IPM general part database model (OLTP package—Figure 4);
- An OLAP system that allows the multidimensional analysis of IPM defense rules. It implements the OLAP model of Figure 5.

As previously described, the two systems are strictly connected. In fact, the FMIS feeds the OLAP tool with data regarding the phytosanitary operations, while the OLAP system provides useful feedback to zoom in on the detail from information collected in the FMIS.

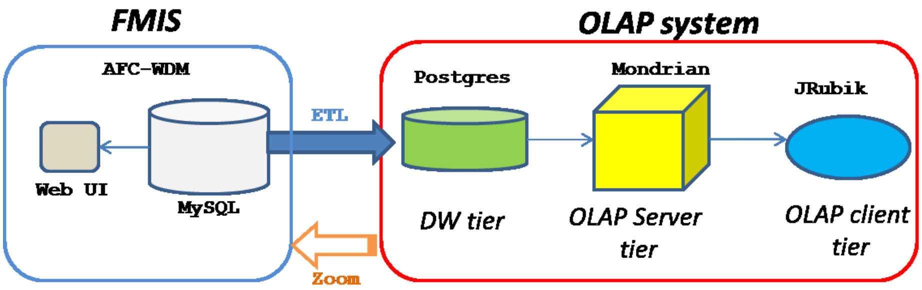


Figure 7: BI4IPM architecture.

7.4.1 BI4IPM: FMIS

AFC-WDM is an integrated cloud platform working on a Web 2.0 environment on multiple devices (PCs, tablets, smartphones) in the model SaaS—Software as a Service.

In the AFC-WDM platform, the boundaries and scope of the system are described in terms of entities and functionalities, where entities are subjects that interface with the system (e.g., managers, database, or physical resources). To fulfill farmers' needs, required information has to be registered into the system through a detailed data-entry phase that starts from the main information sources (company, farm and resources such as plots/parcels, stores/products, buildings, or machines) down to the daily insertion of activities made by the farmers to obtain the correct resource management.

These data-entry activities are all guided, so that each field inserted into the screen forms is matched against the information dimensions present in the database, avoiding mistakes and guaranteeing a formal and substantial correctness of data entered into the system. In particular, for each referenced dimension (e.g., tables containing plots, products) the user can verify and choose the desired item from a predefined list, and the semantic and formal integrity of the database is assured by the user's ability to pick only from the allowed (and thus correct) information related to the activity under management.

The data collected by the platform concerns all the management aspects of the agricultural farms and can be updated every moment by the users. Data are stored in an OLTP database structured in three layers:

1. *Company and Farm*: All the basic information related to the company and the farm are collected (VAT number, address, director, etc.).
2. *Resources*: All the possible resources (buildings, warehouses, machines, plots) that could be assigned to a single farm starting from a specific date are stored. In particular, raw data regarding the protection operations issued from the FMIS are extracted and analysed by an agronomist according to the IPM defense rules. Therefore, adherence to the rules is manually checked and then stored in the FMIS.
3. *Management*: All the technical operations performed using one or more resource are registered. For example, on a single plot (group of parcels), all the agronomic activities (such as irrigations, phytosanitary treatments, tillage operations, etc.) could be managed.

Some data recorded in the FMIS are used to feed the DW tier of the OLAP system. Indeed, as shown in Figure 7, and according to the "use" load association of Figure 6, FMIS data are extracted transformed and loaded (ETL) in the DW tier.

The principal components of AFC-WDM are as follows:

- A complete database based on MySql architecture.
- An Ajax-based Web 2.0 flexible and user-friendly web interface accessible from any browser and operating system.

AFC-WDM provides a complete set of services, from the consultation of publicly available databases (national and international agri-food open data repositories such as ISTAT, EUROSTAT, USD-PSDA, FAOSTAT) to the management of private farm data, with the possibility of obtaining statistics and reports based on farm data and records related to the management of resources such as warehouses, buildings, plots and products.

The platform was designed not only for farmers and agronomists but also for researchers involved in the agricultural and rural sector.

In particular, the platform has three access levels:

1. Super-user: the overall system administrator.
2. Power user: a typical cross-company user, enabled with a series of activities involving multiple farms.

3. User: the standard user, with limited data-management capabilities.

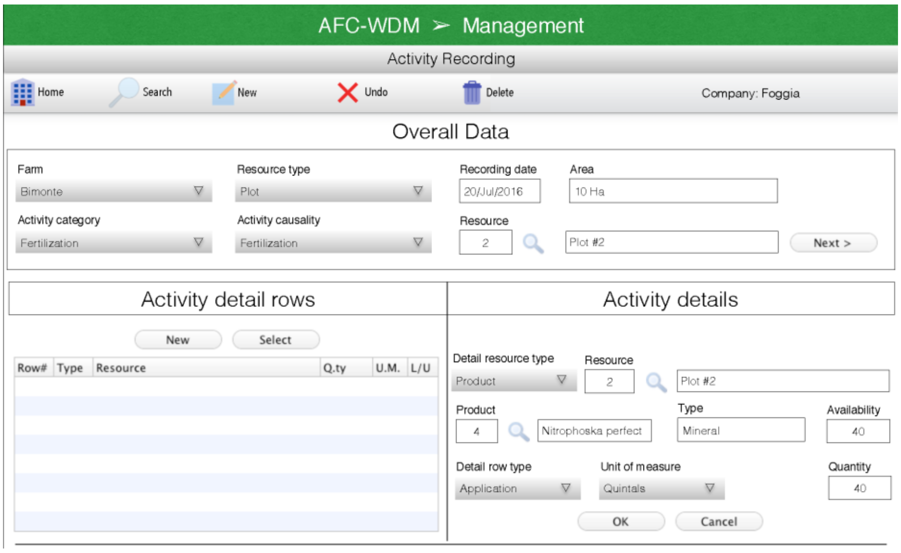


Figure 8: AFC-WDM screenshot related to fertilization activity recording

Figure 8 shows an example of fertilization activity recorded with AFC-WDM and the implementation of the database model described in Figure 4. The system allows the recording of the general information related to the farm, Bimonte in this case, including the plot, the area and the type of activity. Therefore, all other information related to the used product, such as the name, the quantity and the type, are stored. In particular, in the panel titled "Overall Data", are inserted the information related to the entities "*Farm*", "*Plot*" and "*Activity*", while in the panel called "Activity details" are loaded the data regarding the entities "*Activity_details*" and "*Product*". This information is used to verify compliance with the requirement ***Fertilization*** (Table 1).

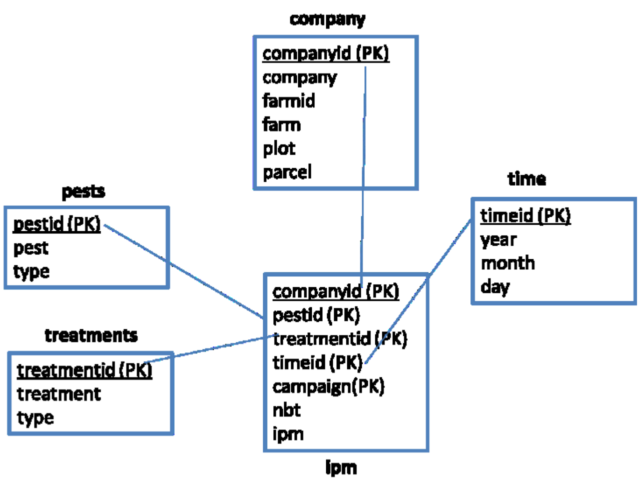
7.4.2 BI4IPM: OLAP system

The OLAP system, based on the multidimensional model described in Figure 5, is a three-tier relational OLAP system composed of (Figure 7):

- The Data Warehouse tier. This tier is responsible for data storage and is implemented using the Postgres DBMS. Data are modeled using the star schema logical model (Kimball & Ross, 2013). Star schema denormalizes dimension tables to avoid expensive join operations.

The star schema model implemented in our case study is shown in Figure 9a. An additional attribute, “farmid” was introduced. It represents farms using simple numerical identifiers. It also allows for an anonymous navigation in the company dimension (see Figure 11 for an example). The campaign dimension is implemented as a degenerated dimension (i.e., an attribute in the fact table).

- The OLAP Server tier. The OLAP server implements the OLAP operators (Drill-Down, Roll-Up, etc.), and access grants. In particular, as shown in Figure 6b, a set of access grants were define to allow only the farm owner to visualize data about his/her plots. Moreover, it calculates the complex measures proposed. "Respect of IPM " is stored in the fact table as a Boolean value (1 or 0). Its aggregation is done using an MDX formula, as shown in Figure 9b. "Error type" is calculated as an MDX formula using two basic values stored in the fact table “code”, which represents "Allowed", "AS not Allowed", "As not allowed in time", and “ipmcount”, which counts the number of errors (>1 when are more errors).
- The OLAP Client tier. It is in charge of the visualization and exploration of warehoused data and is implemented using the OLAP client JRubik.



(a)

```
<HierarchyGrant access="all" hierarchy="[Company].[Company]">
    <MemberGrant access="Bimonte" member="[Company].[Foggia].[Bimonte]"/>
    <MemberGrant access="Zaza" member="[Company].[Foggia].[Zaza]"/>
</HierarchyGrant>
```

```xml
<Measure name="Number of Treatments" column="nbt" visible="true" aggregator="sum"/>
<Measure name="Respect of IPM1" column="ipm" visible="false" aggregator="max"/>
<Measure name="code" column="code" visible="false" aggregator="max"/>
<Measure name="ipmcount" column="ipm" visible="false" aggregator="sum"/>
- <CalculatedMember name="Respect of IPM" visible="true" dimension="Measures">
    <Formula> IIf(([Measures].[Respect of IPM1] = 1), "NO", IIf(([Measures].[Respect of IPM1] = 0), "OK", null))
      </Formula>
  </CalculatedMember>
- <CalculatedMember name="Error type" visible="true" dimension="Measures">
    <Formula> IIf(([Measures].[ipmcount] > 1), "Errors", IIf(([Measures].[code] = 1.0), "A. S. not allowed", IIf
      (([Measures].[code] = 2.0), "Exceeded nb inter.", IIf(([Measures].[code] = 3.0), "A. S. not allowed in time", IIf
      (([Measures].[code] =0.0), "Allowed", null))))) </Formula>
  </CalculatedMember>
```

(b)

Figure 9: Tier implementations: a) star schema and b) Mondrian XML file for access grants and measures calculation.

In the rest of the section, a set of sample OLAP query results are shown using JRubik.

In Figure 10, an example of *Inter-farm Technical specification* query is represented.

Company	Treatments	Time	Error type	Respect of IPM	Number of Treatments
+Bimonte	Copper	+2016	Allowed	OK	2
	Deltamethrin	+2016	Errors	NO	2
	Dimethoate	+2016	Exceeded nb inter.	NO	5
	glyphosate 30,8%	+2016	A. S. not allowed	NO	1
	Pyraclostrobin	+2016	A. S. not allowed in time	NO	1
+2	Copper	+2016	Allowed	OK	5
	Dimethoate	+2016	Allowed	OK	3

Figure 10: Example of light inter-farm technical specification query

In particular, **Q1:** "*Visualize all chemical treatments of farms"* is shown, allowing a comparison between the user's farm (Bimonte) and another farm (FarmID = 2). The user can compare his or her own farm with others for the measures "Error type", "Respect of IPM" and "Number of Treatments" aggregated at farm level, but for privacy reasons, he or she could not access deeper information. As shown in Figure 10, many errors were detected on the Bimonte farm.

Company	Treatments	Time	Error type	Respect of IPM	Number of Treatments
+Plot1	Copper	−2016	Allowed	OK	1
		+09/2016	Allowed	OK	1
	Deltamethrin	−2016	A. S. not allowed	NO	1
		+10/2016	A. S. not allowed	NO	1
	Dimethoate	−2016	Allowed	OK	2
		+06/2016	Allowed	OK	1
		+09/2016	Allowed	OK	1
	glyphosate 30,8%	−2016	A. S. not allowed	NO	1
	Pyraclostrobin	−2016	A. S. not allowed in time	NO	1
+Plot2	Copper	−2016	Allowed	OK	1
		+09/2016	Allowed	OK	1
	Deltamethrin	−2016	A. S. not allowed	NO	1
		+10/2016	A. S. not allowed	NO	1
	Dimethoate	−2016	Exceeded nb inter.	NO	3
		+06/2016	Allowed	OK	1
		+09/2016	Allowed	OK	1
		+10/2016	Exceeded nb inter.	NO	1

Figure 11: Example of deeper intra-farm technical specification query

To obtain more information, an *Intra-farm Technical specification* query is launched (Figure 11). The example provided is the result of the query **Q2:** "*Show all the chemical treatments performed for plots of the Bimonte farm*". In this way, the farmer can obtain detailed information for his/her own farm on where (plot level and deeper) and when (month level and deeper) he or she has not adhered to the IPM. Furthermore, the column "Error type" allows the user to easily understand the explanation of eventual errors for a single treatment. For example, in Figure 11, Plot1 has two A.S. not allowed—Deltamethrin and Glyphosate 30.8%—and one A.S. not allowed in time— Pyraclostrobin. For Plot2, there are only two wrong interventions, Deltamethrin and the third Dimethoate treatment, which exceed the threshold of maximum number of treatments allowed per year (Table 2).

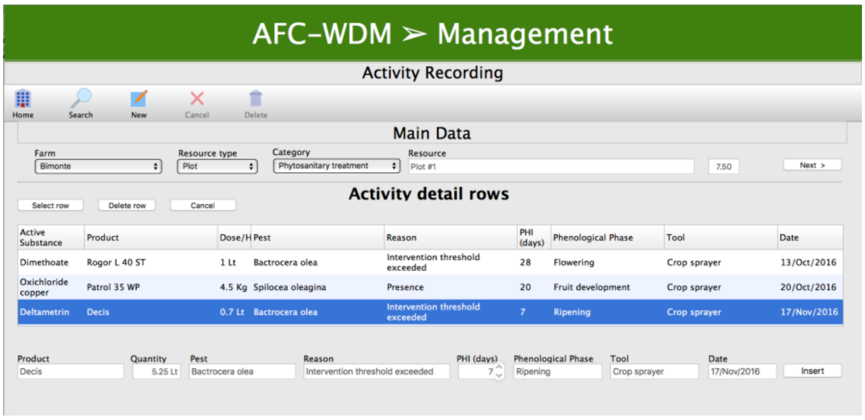


Figure 12: AFC-WDM screenshot related to the recoding of phytosanitary treatments.

Figure 12 shows an example of the *Zoom OLAP* association from Figure 3 being implemented, obtained through the connection between AFC-WDM and the OLAP systems. In particular, based on the feedback provided by the OLAP system through the error type measure—for example, for the Deltamethrin intervention on Plot1 shown in Figure 8—the user can recover additional information related to the single activity instance registered in the FMIS. In fact, as shown in Figure 12, during the activity registration phase in the FMIS, additional information is collected, such as the dose per hectare, the reason for the intervention, the phenological phase, etc., which are not used in the OLAP analyses. From the farmers' point of view, this information is useful because it is necessary to fulfill the farm's notebook, which is mandatory in Italy.

The *Management* queries enable detailed analyses at the intra-farm level. For example, Figure 13 shows the result of following query **Q3**: *"Show the number and type of treatments performed at the farm level for each pest"*. As shown, at the farm level (in this case, the Zaza farm), the type of intervention, Agronomic or Chemical, and detailed information, such as the Active Substance used and the number of treatments were reported for each pest. As in the previous case, this type of query allows the user to visualize an overview of the treatment performed at the farm level.

Company	Treatments	Pests	Number of Treatments
+Zaza	−agronomic	P.syringae pv.savastanoi	1
	Pruning	P.syringae pv.savastanoi	1
	−chemical	Spilocea oleagina	5
		Bactrocera olea	5
	Copper	Spilocea oleagina	5
	Dimethoate	Bactrocera olea	3
	Imidacloprid	Bactrocera olea	2

Figure 13: Example of an intra-farm management query

Figure 14 shows the result of the query **Q4**: *"Visualize the number and type of treatment performed to control pests grouped per type at plot scale"*. This type of query allows a comparison between the different crop-protection strategies carried out by the farmer in different plots on same farm. For example, in Plot11 the farmer executes only one agronomic treatment to control bacteria pest, while in the other plot, s/he does not. The total number of chemical treatments is five in both plots, but in the first one, the farmer has used more fungicides than insecticides and vice versa in the second plot. Based on this information, the farmer could decide the best future crop strategy. For example, in the next campaign s/he will pay particular attention to fungi protection in Plot11, due to the high number of treatments performed.

		Number of Treatments			
		Pests			
Company	Treatments	+ bacteria	+ fungi	+ insects	+ weed
+Plot11	−agronomic	1			
	Pruning	1			
	+chemical		3	2	
+Plot12	+agronomic				
	+chemical		2	3	

Figure 14: Example of a management query at the plot scale

7.5 Integration of GSI in the OLAP system

7.5.1 Local GSI

Attending the indications provided by Orlandi et al., (2013), the GSI related to the Apulian city of Bari, for two period, January 2016 to December 2016 and January 2017 to October 2017, were calculated (Figure 15), using the climatic data described in the Materials and Method section.

Therefore, the obtained daily values *Photo*, T_{min}, *ETP* and *Rad* are used, respectively, to calculate *iPhoto* (3), iT_{min} (5), *iETP* (6) and *iRad* (7). Then, the daily metric *iGSI* is calculated with the relation (8). Finally, the daily GSI is the result of the 21-day running average of the daily *iGSI*. In the Figure 15 are presented the obtained GSI daily distribution for the 2016 and the 2017.

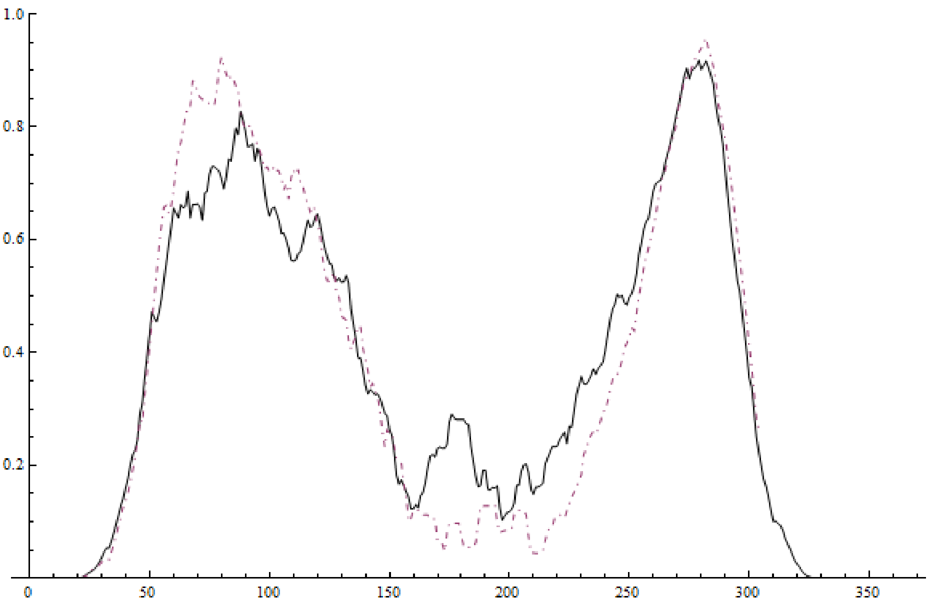


Figure 15: GSI daily distribution for Bari. The continue line represents the GSI for the year 2016, while the dot-dashed line represents the 2017 GSI distribution

7.5.2 Naive OLAP model

To integrate the GSI in BI4IPM, a new version of the OLAP model (Fig. 16) is carried out.

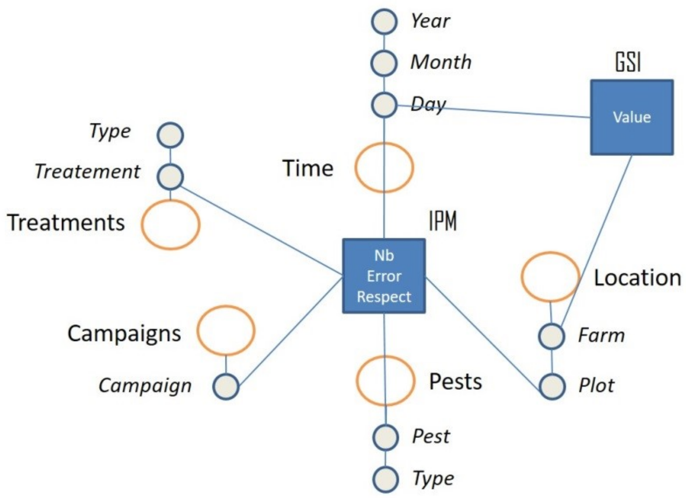


Figure 16: OLAP model with the integration of GSI

For the fact IPM, the measures and the dimension are the same used in the previous OLAP model (Figure 5). The only exception is the dimension *Location*, which substitutes and simplifies the dimension *Company*, grouping the plots in farm. In addition, is introduced the fact GSI which represents the daily value of the GSI (*value* measure), and it is described by the two dimensions *Time* and *Location*. The only difference with the IPM fact is that the *Location* dimension is associated at the Farm level.

Using this constellation model is possible to visually querying each fact for independently exploring IPM and GSI data. An example of these two kinds of queries is shown in Figure 17. Figure 17a shows the OLAP query "*Show the number of performed treatments for Farm2 and Farm4 organized by campaign, pests and treatments*" on the IPM fact. Figure 17b shows the OLAP query "*What is the September daily GSI value calculated for the Farm1?*" on the GSI fact. The comparison of the two data sets must be by hand done by the decision-makers at glance, which is difficult and sometime impossible (when data is huge).

To solve this problem, OLAP systems provide the Drill-Across operator. This operator allows to analyse in the same OLAP client the two facts uniquely using common dimensions, and grouping to "All" other dimensions (using the Drill-Across operator). In the proposed case study, this means that it is possible to analyse the IPM measures with the GSI value according to the Time and the Location (at the Farm level). Other dimensions cannot be used for the OLAP analysis. An example is shown in Figure 17c, results of the OLAP query "*Show the days where one or more treatments occurred and the related daily GSI value*".

Farm	Campaign	Pests	Treatments	Number of Treatments
+Farm2	2016-2017	Spilocea oleagina	Copper	2
		Bactrocera olea	Deltamethrin	1
			Dimethoate	1
			Imidacloprid	1
	2017-2018	Spilocea oleagina	Copper	1
		Prays oleae	Dimethoate	1
−Farm4	2017-2018	Spilocea oleagina	Copper	4
		Prays oleae	Bacillus thuringiensis	2
			Deltamethrin	1
			Fosmet	2
P1_4	2017-2018	Spilocea oleagina	Copper	2
		Prays oleae	Bacillus thuringiensis	1
			Fosmet	1
P2_4	2017-2018	Spilocea oleagina	Copper	2
		Prays oleae	Bacillus thuringiensis	1
			Deltamethrin	1
			Fosmet	1

a)

Time	Farm	GSI
+08/16	Farm1	0,319
-09/16	Farm1	0,665
01/09/2016	Farm1	0,503
02/09/2016	Farm1	0,498
03/09/2016	Farm1	0,501
04/09/2016	Farm1	0,488
05/09/2016	Farm1	0,484
06/09/2016	Farm1	0,498
07/09/2016	Farm1	0,504
08/09/2016	Farm1	0,519
09/09/2016	Farm1	0,54
10/09/2016	Farm1	0,574
11/09/2016	Farm1	0,595
12/09/2016	Farm1	0,619
13/09/2016	Farm1	0,632
14/09/2016	Farm1	0,639
15/09/2016	Farm1	0,662
16/09/2016	Farm1	0,686
17/09/2016	Farm1	0,699
18/09/2016	Farm1	0,703
19/09/2016	Farm1	0,705
20/09/2016	Farm1	0,715
21/09/2016	Farm1	0,737

b)

Time	GSI	Number of Treatments
15/09/2016	0,662	5
16/09/2016	0,686	2
28/09/2016	0,871	1
30/09/2016	0,904	1
11/10/2016	0,876	4
14/10/2016	0,796	2
15/10/2016	0,769	2
25/10/2016	0,397	1
15/03/2017	0,842	1
11/04/2017	0,721	2
28/04/2017	0,66	1
05/05/2017	0,526	2

c)

Figure 17: a) Example of query performed on the IPM fact; b) Example of query performed on the GSI fact; c) Example of query performed thanks to the Drill-Across operator.

This problem is raised from the DW design methodology adopted that does not allow to use factual data to analyse another fact in a constellation model. Therefore, to solve this problem the DW design methodology proposed by Sautot et al., (2015) was adapted as described in the next subsection.

7.5.3 Advanced OLAP model

Sautot et al., (2015) present a mixed multidimensional refinement methodology that transforms constellation schema to define hierarchy level using a hierarchical clustering algorithm. This refinement methodology enriches a dimension with factual data, and considers the context of factual data. In particular, the methodology note "source fact" the fact that is removed and that is used to enrich the "target dimension" with new hierarchies created using a data mining algorithm.

In this work, this methodology was applied using as source fact the GSI fact. The target dimension is the Location dimension. However, to meet the decision-maker analysis needs three modifications were provided to Sautot et al., (2015):

1. The measure of the of the source fact (i.e. *value of GSI*) are added to the other fact (i.e. *IPM*);
2. A new dimension is added with the new hierarchies (i.e. *Similarity*);
3. The data mining algorithm is replace by a statistical method to compare the GSI distributions. The obtained GSI distributions are tested by the Cramér-von Mises normality test. In the presented case the hypothesis that the GSI distributions for Bari 2016 and Bari 2017 (Figure 15) are distributed according to the Normal Distribution is rejected at 5% level (p-value=0). In order to compare the two GSI distributions, the Mann-Withney test (U-test) (Mann and Whitney, 1947) was performed. This test allows to evaluate if two not normal distributions are significant different. In our example, the test returns that the two distributions are significant different at 5% level (p-value=0.019). According to the U-test result, the GSI distributions, expression of the climatic factors, are significant different for the farm in the Bari department for the year 2016 and 2017.

The resulting OLAP model is shown in Figure 18.

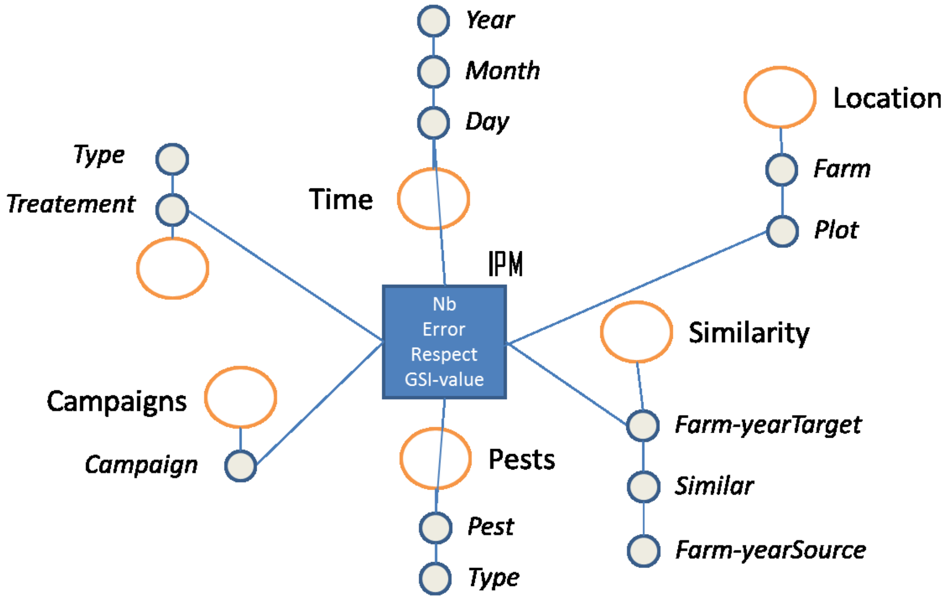


Figure 18: Olive crop IPM advanced multidimensional model

In the following, the new *Similarity* dimension is described.

The Farm-yearSource level represents each agroclimatic condition of farm by year, for example *Farm2-2017* represents the GSI condition of the Farm2 during 2017, *Farm1-2016* represents the GSI condition of the Farm1 during 2016, etc.

The *Similar level*: groups all other farms-year in *similar* and *not similar*, based on the result of the Mann-Whitney test performed on the GSI distribution assigned to farm selected and the others loaded in the system.

Finally, Farm-yearTarget: This represents the farms similar or not, respect to the chosen farm, for example the not similar to Farm1-2016 are *Farm1-2017* and *Farm4-2017*, etc.

An example of this hierarchy is shown in Figure 19.

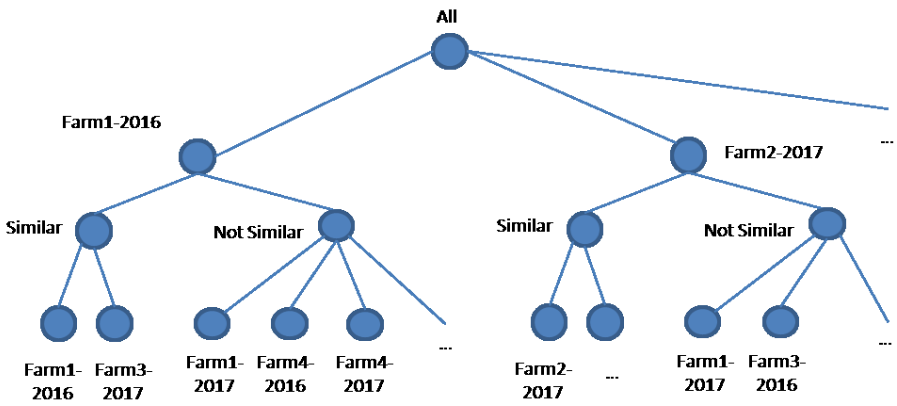


Figure 19: Example of the Similar *dimension's hierarchy*

Using this multidimensional complex model it is possible to answer to OLAP queries using at the same time GSI and IPM measures, and exploring farms that are similar according to the GSI value.

An example is shown in Figure 20. It shows the average of treatments at plot level (*AVG Treatments*) over 2017 and 2016 campaigns and the GSI values, for the farms not similar and similar to the Farm1 in the campaign 2016.

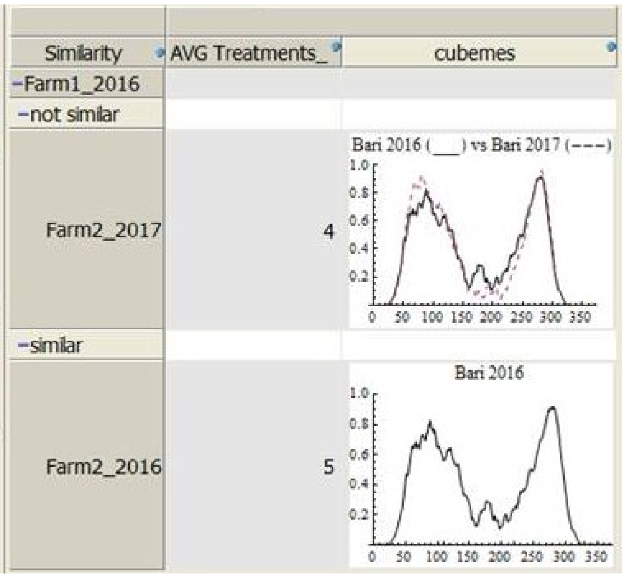


Figure 20: Example query performed using the advanced OLAP model

Finally, this new model solves the issues of the previous described naïve OLAP model.

7.5.4 Implementation and OLAP advanced analyses

The OLAP system is organized as described in the Section 7.4.2. The only differences regards the revised multidimensional model (Figure 21) and the JRubik client, which was modified in order to allow the visualization of the graphics representation of the GSI distribution, as showed in the previous section.

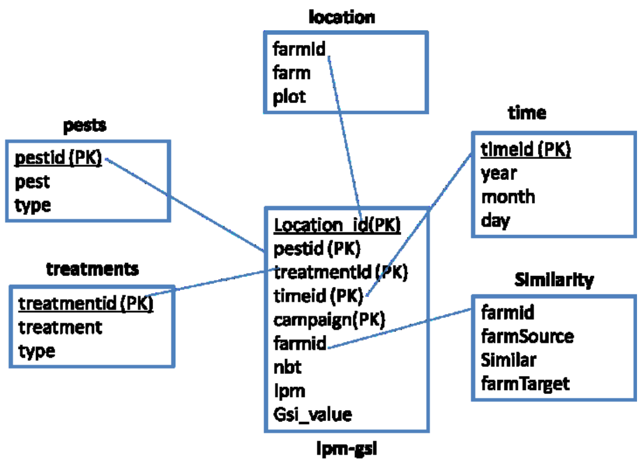


Figure 21: the revised multidimensional model

7.5.5 OLAP advanced analyses

The advanced OLAP model presented in the previous section allows carrying out a set of analyses. In particular, as shown in Figure 22, it allows to perform several inter-farms queries, based on the similarity obtained by the comparison of the GSI distributions. In the provided example (Figure 22), the user farm, *Farm1_2016*, is compared with similar, *Farm2_2016* and itself, and not similar farms, *Farm2_2017*, according to *Similar* dimension (Figure 19), previously described. The similarity or not, is underlined in the system by the graphic representation of the two GSI distributions investigated. In particular, for the similar farms the same GSI annual distribution is showed, while for the not similar farm, the system returns the comparison between the two distributions, highlighting the differences occurred. Moreover, the system permits to analyse the behaviour of the similar and not similar farms, according to the measures *Error type*, *Respect of IPM* and *Number of Treatments*.

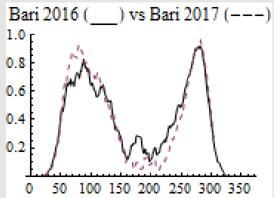
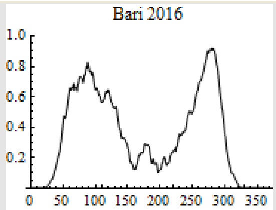
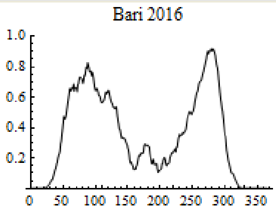
Similarity	cubemes	Error type_	Respect of IPM_	Number of Treatments_
-Farm1_2016				
-not similar				
Farm2_2017		Allowed	OK	4
-similar				
Farm2_2016		Errors	NO	5
Farm1_2016		A. S. not allowed in time	NO	5

Figure 22: example of inter-farms queries, based on the similarity.

It is interesting to argue, that the *Farm2*, is considered similar for the year 2016 and not similar for the year 2017. So in the Figure 23, is showed an intra-farm query, aimed at highlighting the differences occurred during the two analysed years. In this case, thanks to the indication provided by the *Error type* measure, is extremely easy to identify the wrong treatments performed in the 2016. At the same time, in the 2017 the whole crop protection strategy respects the IPM principles, underlined by the output "Allowed" showed at All Treatments level.

Similarity	Treatments	Error type_	Respect of IPM_	Number of Treatments_
Farm2_2017	+All Treatments	Allowed	OK	4
	Dimethoate	Allowed	OK	1
Farm2_2016	+All Treatments	Errors	NO	5
	Dimethoate	Allowed	OK	1
	Imidacloprid	Allowed	OK	1
	Deltamethrin	A. S. not allowed	NO	1

Figure 23: example of intra-farms queries, based on the similarity.

Bibliography

Boulil, K., Bimonte, S., & Pinet, F. (2015). Conceptual model for spatial data cubes: A UML profile and its automatic implementation. *Computer Standards & Interfaces*, *38*, 113-132.

Kimball, R., & Ross, M. (2013). The data warehouse toolkit: The definitive guide to dimensional modeling. John Wiley & Sons.

Orlandi, F., Garcia-Mozo, H., Dhiab, A. B., Galán, C., Msallem, M., Romano, B., Abichou, M., Dominguez-Vilches, E. & Fornaciari, M. (2013). Climatic indices in the interpretation of the phenological phases of the olive in mediterranean areas during its biological cycle. *Climatic change*, *116*(2), 263-284.

Sautot, L., Bimonte, S., Journaux, L., & Faivre, B. (2015, April). Dimension Enrichment with Factual Data During the Design of Multidimensional Models: Application to Bird Biodiversity. In *International Conference on Enterprise Information Systems* (pp. 280-299). Springer, Cham

8. Results on Survey analysis

The questionnaire shows answering from a sample within a Producer Organization (PO) in Lucania region. The respondents are 59.

In this section are exposed the results derived from the cluster analysis and the boxplot performed on the collected answers.

The queries are aimed at evaluate the number of farms that adopt ICTs tools, associating evidences from the farm structure and the age of the farmers. The cluster analysis shows two main clusters characterize as follows (Table 6):

Table 6: Two emerging clusters after processing data. The result concerns a selected number of variables from the survey.

	CLUSTER	
	1	2
Legal status	1	1
Crop type	2	3
Age	2	2
UAA	2	2
incomes	2	4
ICT	0	1

The resulting clusters present different features, only concerning three variables: crop type, incomes and ICT. In general, both clusters appear to consist of farms established as partnership, with farmers being medium age ranking old (35-50) and farm area between 10 and 50 hectares. Differences come from:

- Crop type: the cluster 1 is featured by herbaceous crop, while the cluster 2 is featured by mixed crop (herbaceous and tree crop);
- Incomes: the cluster 1 gains to the extent between 50.000,00 and 120.000,00 euros, instead the cluster 2 lines up between 250.000,00 and 500.000,00 euros;
- ICT: the cluster 1 is represented by farms not adopting ICT tools, the cluster 2 is featured by farms adopting ICT.

Table 7 shows the frequencies within each cluster. There appear the 38 clustered in the first (1) group and the 21 grouped in the second (2) cluster.

Table 7: Number of cases for each cluster.

Cluster	
1	38,000
2	21,000
Valid	59,000
Missing	,000

The major number of farms is concentrated in the first cluster. Although the last evidence, the cluster analysis points out that, in accordance with Dimara and Skuras, (2003), Sauer and Zilberman (2012), Stefanides and Tauer (1999) asserted, come up two groups: innovators and remaining groups. The cluster 2, populated by innovators, registers revenues much more relevant than the 1. The measure corresponds to the range between 250.000,00 and 500.000,00 euros against the cluster 1 featuring incomes between 50.000,00 and 120.000,00 euros. Furthermore, it results that the cluster adopting ICT is characterized by mixed crop. In this regards, the crop diversification is associated to higher incomes (Di Falco and Zoupanidou, 2017). Taking into consideration that the ones encompassed in the cluster 2 adopt ICT tools, the result in terms of incomes is significant. When farms adopt ICTs seem to improve the performance. On the other hand, there is a dependency between those two variables, though it is not defined the direction: at this stage is not clear what kind of factors push farms in innovating with ICT. In fact, it can depend on the achievement of excessive dimension and, due to the increasingly complexity, farms need to improve the data collection and management phases; otherwise, it can depend on the need to improve the revenue performance and so, the adoption of ICT tools cause the incomes increasing. For making clearer the explained point, it has been done another analysis. Assuming that the variable type of crop is excluded due to it has been chosen to conduct the analysis not considering such qualitative agronomic variable. The focus remains on the economic aspects, exploring relations with economic parameters.

Figure 24 shows the boxplot where are intersected four different variables, looking for stratum where farms adopt ICT. Outputs put in evidence that there are three main groups classified by UAA and age. Firstly, results seem to be in comply with the conclusions of Diedern et al., (2003). Indeed, the picture features three different relevant groups that can be summarized as follows:

- Innovators, characterized by age between 18 and 35 and UAA no more 10 hectares;
- Early adopter, mainly featured by age between 35 and 50 and UAA between 11 and 50 hectares;
- Laggards, principally classified by age over 50 and UAA between 35 and 50 hectares.

It has been assumed that if the age and the UAA present low value, hence the ICT adoption positively affects the productivity and, in consequence, the incomes. The picture displays that the major concentration of the ICT matches with the square corresponding to age level 1 (18-35) and UAA level 1 (0-10 hectares). This consideration confirm that younger farmers are much more incentivized and motivated to resort to ICT tools for managing farming activities (Plechowski, 2015). The innovators label is due to the age of the farmers, who, even though the low profile in terms of utilized lands, got medium-high level of incomes. The data is also confirmed by the legal status. In fact, the corporations are concentrated within the innovators group. In this regard, this type of legal status costs more than the one for partnership, and for sustaining the effort farms need to account sufficient resources in terms of revenues. This consideration allow answering the question coming from the previous analysis: in this innovators group ICT seem to push the incomes. Early adopters are characterized farmers aged between 35 and 50 and farming lands between 11 and 50 hectares. In that case, the growing of the farm dimension seems to pull ICT tools for managing the growing data complexity. Conversely, the group where age corresponds to the level 2 and the UAA to the level 3, even though is represented by a niche of respondents, it is another cluster of innovators. Therefore, the boxplot clustering analysis goes through the data catching more details than the first one. As a result, the innovators' cluster is morphologically more various than the one emerged from Table

6. Finally, laggards (or not innovators) are featured by the oldest classified farmers, not interested in introducing ICT devices in farm processes.

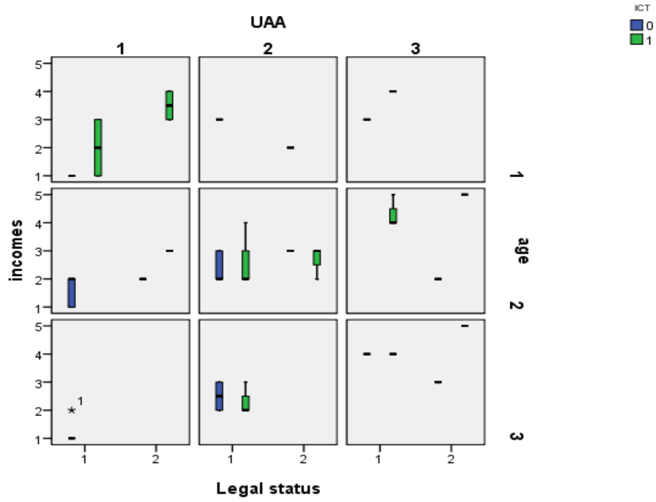


Figure 24: Boxplot grouping clusters according to age and UAA.

Bibliography

Di Falco, S., & Zoupanidou, E. (2017). Soil fertility, crop biodiversity, and farmers' revenues: Evidence from Italy. *Ambio*, *46*(2), 162-172.

Diederen, P., Van Meijl, H., Wolters, A., & Bijak, K. (2003). Innovation adoption in agriculture: innovators, early adopters and laggards. *Cahiers d'Economie et de Sociologie Rurales*, *67*, 29-50.

Dimara, E., & Skuras, D. (2003). Adoption of agricultural innovations as a two-stage partial observability process. *Agricultural Economics*, *28*(3), 187-196.

Plechowski, K. (2014, October). Youth, ICTs and agriculture-exploring how digital tools and skills influence the motivation of young farmers. In *eChallenges e-2014, 2014 Conference* (pp. 1-8). IEEE.

Sauer, J., & Zilberman, D. (2012). Sequential technology implementation, network externalities, and risk: the case of automatic milking systems. *Agricultural Economics*, *43*(3), 233-252.

Stefanides, Z., & Tauer, L. W. (1999). The empirical impact of bovine somatotropin on a group of New York dairy farms. *American Journal of Agricultural Economics*, *81*(1), 95-102.

9. Conclusion and future work

This section closes the work, providing a set of relevant conclusions derived from the critical analysis of the presented results. Moreover, the section 9.2 is dedicated to the future works related to the open research questions.

9.1 Conclusion

The ability to create a model that deals with traceability and certification of agri-food product based on RFID technology, would mean improve the quality and the logistics of the chain and offer therefore, environmental friendly and cost-effective solutions to optimize production flows, by networking the existing Italian hubs. The advantages offered by RFID technology compared to the barcode, currently the most widely used identification systems, is enclosed in a series of characteristics listed below:

1. Efficiency: more RFID tags can be read simultaneously from a single remote reader, even without visibility (presence of intermediate materials). The tags show their best performances especially in environments with dust, moisture, corrosive agents.
2. Memory: RFID tags can store more information;
3. Durability: theoretically, there is not limit of read/write cycles for RFID tags;
4. Flexibility: RFID tags are reusable, in fact, theoretically, they can be written;
5. Speed: RFID tags work in real time and automatically;
6. Tracking: RFID tags allow obtaining information about the products during all the steps of the supply chain;
7. Automation: the RFID tag can be read automatically, while, usually, the barcode reading required a human intervention. This ensures an errors reduction and increases the operations efficiency.

In conclusion, the two solutions developed in the CESAR project show clear benefits. Compared to random checks system, normally used in the agri-food chain, the RFID technologies proposed allow a constant monitoring of each individual production unit and the possibility to take prompt action in case of anomalous values. This implies, therefore, to increase the quality process and ensures the reduction of cost due to less use of resources and time for the non-compliance management. Another tangible benefit regards the improvement of the traceability system accuracy, given by the possibility to track the product by a unique identification code associated to the label that is generated by the chip manufacturer. Finally, there is the opportunity to improve business processes through analysis of the huge amount of data that derives from the use of RFID technologies.

However, have been identified a set of critical issues concerning the adoption of the proposed solutions, listed below:

- Damage to the RFID system: in case of damage to the label, produced in different form and size, that preclude the reading of the data, does not exist an alarm system able to inform the operator. Therefore, it is necessary to provide an action mode to predict when readers have not completed the acquisition of data;
- Changing standards: the standards change may result in the occurrence of false negatives and positives. The first case occurs when a valid tag cross the prescribed range of an RFID reader, but that does not extract the information. This can happen if the tag is completely hidden, or the signals are absorbed or blocked by the materials that surround the label. The second case occurs when a tag accidentally comes in an RFID reader range. The solution is the interoperability of standards of existing and future systems;
- Improper functioning of the RFID transponder: to guarantee the correct reading it is important to ensure the absolute minimum distance between the tag and the reader, so it can also be read inside other containers;
- Depletion of the active tag battery (label): the proposal battery must be used at least, because it runs out quickly;
- Configuration and management of reading devices: sometimes reading and interpretation of the data leads to difficulties for the operator; in this case, the process should be simplified by using automated processes for the set-up, configuration and data management.

Regarding the data management and analysis, in this work, was presented BI4IPM, a business intelligence framework that combines on-line transaction processing with on-line analytical processing to verify adherence to the Olive Crop IPM provided by the Apulia Region.

The OLTP approach is used to verify the compliance of daily farm-management operations, collected through a dedicated FMIS called AFC-WDM, with the requirements included in the general part of the IPM. The data model presented allows the identification of all the information useful to compare the farmer's behavior with the requirements necessary to access the economic incentive.

The OLAP model allows verification that farmers adhere to the Olive Crop IPM defense rules. As shown in this work, the user can use the system to obtain detailed spatial and temporal information related to his/her farm thanks to the different hierarchies implemented. Moreover, two groups of queries have been developed and tested. The first group concerns the technical specification compliance, investigating the adherence to the defense rules at the inter- and intra-farm levels. The second group includes the management queries, allowing the user to analyse the crop-protection strategies of his/her farm from different points of view, such as the number of treatment for a plot or parcel, the type of interventions for each pests, etc.

The two subsystems that comprise BI4IPM are strictly connected. In particular, a part of the information recorded in AFC-WDM is loaded in the OLAP system. Moreover, thanks to the feedback information provided by the OLAP analyses, the user can zoom in on the additional information present in the FMIS at level of a single instance.

The proposed models have been developed based on real needs expressed by farmers. In fact, the IPM is not only mandatory in European Countries but could represent an opportunity for the farmers to improve their incomes, as in the case of the Apulia region that stimulates the diffusion of certified integrated production through an economic incentive. BI4IPM could be satisfactorily used by farmers to verify that all the requirements included in the crop-specific disciplinary provided by the Region have been fulfilled, first as a condition to obtain the abovementioned incentive. Moreover, producers are also checked by an independent control authority, who verifies that all the farm's operations are compliant with the IPM. For this stage, BI4IPM could be used conveniently not only by the farmers but also by the control authority. In this work, these hypotheses were validated at a very small scale (using one agronomist). For example, using the OLAP system, the agronomist has easily understood why the IPM was not adhered to (Figure 10) simply by analysing data at a plot level (Figure 11). In detail, for the Bimonte Farm, five treatments were with Dimethoate, and the error type reported for the aggregation is "*Exceeded nb inter*" (Figure 7). At the plot scale, using a drill-down operation (Figure 11), the treatments are separated, and only one does not adhere to the IPM requirements: the third in Plot 2. The reported queries could also be used by the producer organizations (POs), which aggregate groups of farmers. In fact, in this case, the PO responsible for a single account could check all the farms in his or her organization. Moreover, it is possible with a single technical specification query to verify that all the farmers have adhered to the IPM requirement at the farm scale. At the same time, this analysis allows to identify the "best farms" in a PO, such as the farms that use fewer chemical products, to suggest to the other farmers the same protection strategy adopted by the best farm.

Furthermore, the integration of the GSI model in the BI4IPM represents an opportunity for the farmer to analyse the crop protection strategy taking into account the climatic conditions. In fact, the GSI represents the phenological status of the Olive tree, in response to the most important climatic and environmental factors, such as the temperature, the solar radiation, and so on.

The performed survey analysis permits to map the profile of farms in Lucania region, putting in the spotlight good practices in terms of smart devices used for improving the efficacy and efficiency of the farms decisions and daily actions. Practitioners can optimize processes rising to be considered good practices when they reduce the inputs and improve the outputs, becoming increasingly revenues and more sustainable. In addition to previous reasons inspiring this study, the survey was also aimed to show good practices adopted by Lucania farmers. The goal fits with the objectives of the H2020 (EC, 2015) project Short Supply Chain Knowledge Innovation

Network (SKIN)[2], granted by European Commission and started on the 1st November 2016. The project consists in collecting good practices operating in short food supply chain and involving them in European building network in order to boost and facilitate knowledge transfer and real innovation uptake. The metrics indicators referenced within the project activities for collecting good practices, points out that farms raise to be good practice also if adopt smart tools, such as ICT tools, for improving the economic, environmental and/or social sustainability.

Grouping collected data in clusters allows identifying the most significant features qualifying smart organizations. The innovators and early innovators are ready to get into the network providing their experiences and gaining from other farms experiences. On the other hand, laggards can benefits after network will be built and synergies will be engaged. They can align their profile to the smarter ones. Innovators in terms of ICT adoption are mentioned like the ones able to promptly fit farm' activities to the environmental complexity and thence the ICT tools play an important role in moving to that category due to they return efficiency and efficacy (if they are rightly implemented). By contrast, laggards, even though the growing environmental complexity (external factors deflecting the right activities implementation if not correctly managed) do not adopt solution to make simple processes. However, innovations come up to be needed in rural and agro-food transition to allow farms becoming economically sustainable. Such necessity is implied in the farm size that is mainly medium-small and it reduces the competitiveness in findings profitable markets. The agricultural shocks are going to increasingly be frequent due to the market uncertain. ICTs facilitate the information management and the shock control.

Finally, this study considered a small sample of farms from Metaponto area where are mostly concentrated agricultural activities, in the Lucania region. It appears, obviously, as weakness. Nevertheless, it has been tried to look at the composition of

[2] The results related to the ICT survey are part of the SKIN project (www.shortfoodchain.eu). This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under grant agreement N. 728055.

the sample interviewing three different types of farms according to the Crop Type (tree crops, herbaceous and mixed).

9.2 Future works

The two RFID presented solutions need to be tested and evaluated on reals agri-food companies. In fact, the strategy to be adopted will therefore help to bridge the gap that still exists between research and market in the innovation process. This strategy ensures that the solutions will be targeted to specific requests or fields, providing a direct application of the research findings and benefits in terms of quality and marketing for the agri-food productions.

The design and implementation of BI4IPM should be considered as a starting point. Indeed, as on-going work, full-scale trials in experimental farms across the Apulia region will be prepared. Moreover, the integration of more real datasets to evaluate the tool's performance in different situations are needed. In particular, applying and testing BI4IPM with other key crops in the Apulia region, such as grapevine and wheat, generalizing the proposed model and taking into account the requirements included in the technical specification of these two cultivations. Finally, has been planned to analyse the use of the proposed OLAP approach for all farm-management activities to understand hidden possible correlations among agronomical and protection activities.

Bibliography

European Commission, EC (2015). Calls for proposals and related activities under the 2016-17 work programmes under Horizon 2020 — the Framework Programme for Research and Innovation (2014-20) and under the Research and Training Programme of the European Atomic Energy Community (2014-18) complementing Horizon 2020. OfficialJournal of the European Union (2015/C399/02).