Research Journal of
Forestry

Big Data Analysis (Business Analytics) in Agriculture and Forestry: A Bibliography Review[1]

[1]K. Kitikidou and [2]N. Arambatzis
[1]Department of Forestry and Management of the Environment and Natural Resources, Pandazidou 193, 68200, Orestiada, Greece
[2]Department of Crop and Animal Production, Prefecture of Chalkidiki, Dioikitirio, 63100, Polygyros, Greece

Corresponding Author: K. Kitikidou, Department of Forestry and Management of the Environment and Natural Resources, Pandazidou 193, 68200, Orestiada, Greece Tel: 0030-2552041179

ABSTRACT

Now-a-days we are getting flooded with information, largely deriving from the World Wide Web. This information is dispersed and no significant relationship may be obtained from it and neither can this great rate be managed by conventional computer storage. It is considered that organizations that are best capable to establish real-time business decisions applying Big Data solutions will flourish, whilst those that are ineffective to adopt and make the best of this change will progressively find themselves at a competitive disfavour in the market and face possible bankruptcy. Big data analysis assures adequate capacity in terms of computer storage and processing power to elastically deal with data of such magnitude and by the use of analytics to acquire value from it. This study focuses on agriculture and forestry data and the use of particular bid data analysis tools used to get some substantive information which can be utilized for strategic and successful agriculture and forestry. Preceding analogous studies are discussed and recommendations are given.

Key words: Big data analysis, business analytics, agriculture, forestry

INTRODUCTION

In a world where big data is encountered on a day to day basis, the storage should be such that it can handle the data and that it promotes heavy server processing needed in a cost effective way. With the elevating need to have agriculture and forestry inventory data being ecologically oriented, the best way to deliver it in a high quality form is through the implementation of fresh and relevant technologies. Liang (2013) revealed that the inventory data has evolved over time and there is need for a continued update of this data. Some of the common data needed for agricultural and forestry analysis are obtained from trends in soil, pest control and farm land ownership. Furthermore, other significant data is available from vegetation cover, trees crown conditions as well as crop diseases among others. When inventory information is dispersed and unsubstantiated, the management of farms and forests is left to use unreliable data and information. As a result, farmers and forest managers lack a clear picture on what strategies should be used to deal with challenges like deforestation and soil degradation and rely on information obtained from intuition developed from numerous years of experience. Additionally, proper reliable inventory data management facilitate the reduction of anthropogenic emissions like carbon dioxide resulting from deforestation (White *et al.*, 2013).

---

[1]This study is the extended version of a work entitled "Big data analysis (business analytics) in forest inventories; A bibliography review," presented at the 27th Pan-Hellenic statistical Conference, 23-26 April 2014, Thessaloniki, Greece

FACTORS COMPELLING INVENTORY BIG DATA STORAGE

Apart from providing relevant information for improved agricultural and forest management practices, the need for bid data analysis tools is emphasized through the requirement for green power, especially in data centers where real time forest inventory data is held and technical equipment like projectors and desktops are needed to display or output stored information for analysis (Munyua, 2000). Like organizations, farmers and forest managers require efficient, reliable and up to date information to support their vital decision intended for promoting sustainable agriculture and forestry (Maniatis *et al.*, 2011). Consequently, information technology penetration in agriculture and forest management through inventory data is promising, especially through e-science infrastructure and partnership on one side and the speedy development of digital gadgets and connectivity in various regions on the other hand. When coupled together, information technology devices and collaboration and e-science promote transformed creation, sharing and application of knowledge on agriculture and forest management. The storage of such information makes use of technologies like cloud computing in technologies like Amazon and Azure clouds coupled with faster data analysis technologies like SQL.

Food security is also another factor that promotes the adaptation of Information Communication Technology in agriculture. When applied to agriculture, ICT is perceived as the change instrument with the potential to offer to small scale farmers a huge transformation especially when it comes to make agricultural decisions that affect their livelihoods. Consequently, agriculture will become an information and knowledge-based process that is no longer the traditional input-intensive way of farming. Waga and Rabah (2014) highlighted that there are numerous tools that can be made and customized to offer farmers unique and distinct crop information and insurance. An example is Mansota's Climate Corporation start up in San Francisco (Waga and Rabah, 2014). Climate Corporation uses data from weather and agronomic to advise farmers on the best time for them to water spray pesticides and nutrients as well as reap their ready crop. Information Communication Technological tools like Eye on Earth that uses Microsoft Azure and Esri's cloud to allow researchers, policy makers and scientists to upload, share and analyze data related to the environment. Other ICT tools for agriculture are Google Earth engine which provides tools and parallel processing computing. With this tool, scientists can collect information that is time-sensitive on weather and water which when critically analyzed would result to information relevant to promote environmental sustainability (Ballantyne *et al.*, 2010).

CHALLENGES WITH CLOUD COMPUTING

In computer networking, cloud computing is computing that involves a large number of computers connected through a communication network (Fig. 1).

Cloud computing is one of the newest technological developments to shake up the business world. As more companies migrate their data and applications to the cloud, business itself changes dramatically. With data and applications held in the cloud, processing times are much faster for everything. Companies can search for information, run an application or input new data in record time. This development speeds up all business-related operations, including buying, selling and data management, thus making businesses more efficient overall. However, the adoption of cloud computing in agriculture and forestry does not bring to an end the problem of relevant information availability for agriculture and forest management. On the contrary, cloud computing get slow, especially when large data is involved. As a result, the use of massive parallel computing with analytic algorithms is a crucial way of attaining, fast and efficient processing of large amounts of data. Parallelism is encouraged in data processing, given that each of the systems in parallel is designed to process certain data and create distinct solutions for all the segments of data
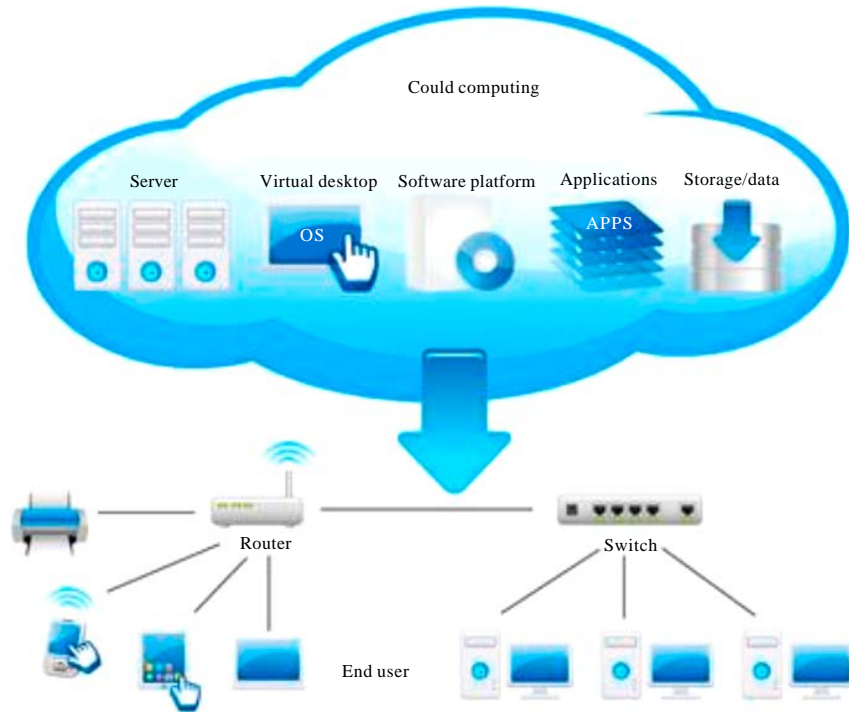
Fig. 1: Cloud computing. Source: http://www.rickscloud.com/5-ways-cloud-computing-is-changing-business/

(Supergo Technologies Inc., 2014; Ballantyne *et al.*, 2010). The result is the dire need for embracing data analytics such that organizations will manage to prepare and work on huge amounts of data to obtain findings that will efficiently benefit the farmer and forest manager. Different forms of data analytics are suitable for different organizational settings. Business analytics is an important tool, since it combines capabilities in databases, processing of natural language, data virtualization and artificial intelligent in the collection and analysis of information (Waga and Rabah, 2014).

AIRBORNE AND TERRESTRIAL LASER SCANNING

Despite the presence of the ever growing need to access relevant information for enhanced forest inventory data management, there is need to understand the data being collected. Additionally, the tools needed to collect data are also very important, especially in making sure that the data collection becomes a cost effective process. For this kind of data collection, the practice of agriculture and forest inventory should focus on Airborne Laser Scanning (ALS) data through an approach pertaining to each area (White *et al.*, 2013) (Fig. 2).

For farmers and forest managers, information collected through Airborne Laser Scanning is in three-dimensional form of vertical vegetation which then translates to easily develop high quality digital elevation molds. Since the collection of data, categorization and surface generation activities needed when using Airborne Laser Scanning is tedious and expensive. In order to offset the high costs, farmers and forest managers have left the work of data collection to commercial agents who are contracted for this initial stage. In order to acquire consistent and transparent data gathering and dispensation, the inventory data standards are set and the commercial agents are observed.
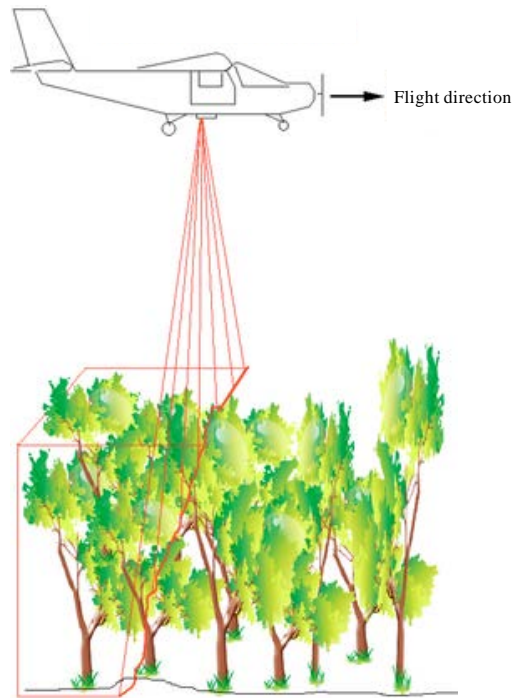
Fig. 2: Airborne Laser Scanning (ALS). Source: http://wiki.awf.forest.uni-goettingen.de/wiki/ index.php/3D-_Terrestrial_laser_scanning
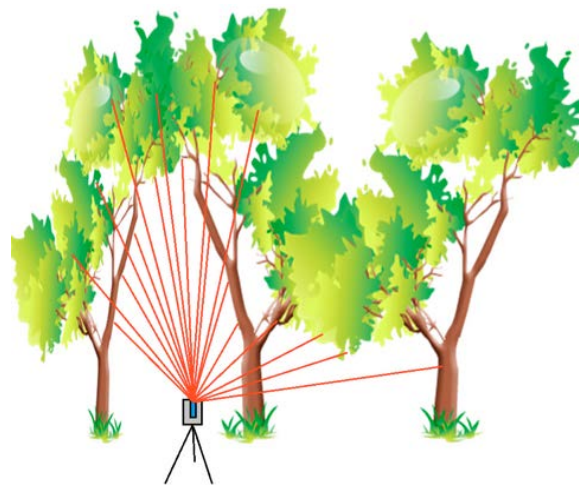

Fig. 3: Terrestrial Laser Scanning (TLS). Source: http://wiki.awf.forset.uni-goettingen.de/wiki/ index.php/3D-_Terrestrial_laserscanning

Apart from ALS forest inventory data collection, Terrestrial Laser Scanning (TSL) has also proved to be promising, especially in the correlation of numerous tree attributes (Liang, 2013; Papalampros, 2013) (Fig. 3).

However, this technology has not managed to embrace data processing automation as well as its cost. Conversely, embracing TLS means tapping into the technology's ability to document details of sample plots and its use of automated mapping technique that results to precise measurements of numerous vegetation attributes. Farmers and forest managers intending to use the most accurate data should focus on utilizing ALS and TLS technologies to enhance the tree attributes needed to promote proper agriculture and forest management (Liang, 2013).

REFERENCES

Ballantyne, P., A. Maru and E.M. Porcari, 2010. Information and communication technologies-opportunities to mobilize agricultural science for development. Crop Sci., 50: S63-S69.

Liang, X., 2013. Feasibility of terestrial laser scanning of plotwise forest inventories. Ph.D. Thesis, Aalto University, Espoo, Finland.

Maniatis, D., Y. Malhi, L.S. Andre, D. Mollicone and N. Barbier *et al.*, 2011. Evaluating the potential of commercial forest inventory data to report on forest carbon stock and forest carbon stock changes for REDD. Int. J. For. Res., Vol. 2011. 10.1155/2011/134526

Munyua, H., 2000. The role of information and communication technologies in rural development and food security. Sustainable Development Department (SD), Food and Agriculture Organization of the United Nations, Rome, Italy, November 2000.

Papalampros, L., 2013. Evaluation of a terrestrial laser scanner methodology for the determination of forest inventory parameters. M.Sc. Thesis, Aristotle University of Thessaloniki, Greece.

Supergo Technologies Inc., 2014. Implementing a cloud-based GIS system for agricultural resource management. Supergo Technologies Inc., Taipei, Taiwan, February 18, 2014.

Waga, D. and K. Rabah, 2014. Environmental conditions' big data management and cloud computing analytics for sustainable agriculture. World J. Comput. Applic. Technol., 2: 73-81.

White, J., M.A. Wulder, A. Varhola, M. Vastaranta and N.C. Coops *et al.*, 2013. A best practices guide for generating forest inventory attributes from airborne laser scanning data using an area-based approach. Information Report FI-X-010, Canadian Forest Service, Canadian Wood Fibre Centre, Victoria, BC, Canada.