Estimating the Environmental Impact of Agriculture by means of Geospatial and Big Data Analysis: The Case of Catalonia

Andreas Kamilaris[1], Anton Assumpcio, August Bonmati Blasi, Marta Torrellas and Francesc X. Prenafeta-Boldú

GIRO Joint Research Unit IRTA-UPC, Barcelona, Spain

Abstract: Intensive farming has been linked to significant degradation of land, water and air. A common body of knowledge is needed, to allow an effective monitoring of cropping systems, fertilization and water demands, and impacts of climate change, with a focus on sustainability and protection of the physical environment. In this paper, we describe AgriBigCAT, an online software platform that uses geophysical information from various diverse sources, employing geospatial and big data analysis, together with web technologies, in order to estimate the impact of the agricultural sector on the environment, considering land, water, biodiversity and natural areas requiring protection, such as forests and wetlands. This platform can assist both the farmers' decision-taking processes and the administration planning and policy making, with the ultimate objective of meeting the challenge of increasing food production at a lower environmental impact.

Keywords: Policy tool, Agriculture, Environmental Impact, Geospatial Analysis, Big Data Analysis.

1. Introduction

The central role of the agricultural sector is to provide adequate and good-quality food to an increasing human population and, because of its importance and relevance, it is on the focus of the global policy agendas.

---

[1] Corresponding Author. Email: andreas.kamilaris@irta.cat

Agriculture is considered an important contributor to the deterioration of soil, water contamination as well as air pollution *(Bruinsma, 2003)*. Intensive farming has been linked to excessive accumulation of soil contaminants *(Teira-Esmatges & Flotats, 2003)*, as well as to significant groundwater pollution with nitrate *(Stoate, et al., 2009)*. Almost 23% of global greenhouse gas emissions are attributed to agriculture. The negative environmental impact from livestock farming across Europe continue to make their mark *(Heinrich Böll Stiftung, 2014)*, resulting in new legislations, policies and large research programs. However, despite a huge amount of published material and many available techniques, doubts over the success of national and European initiatives remain, regarding the extent over which environmental targets are met *(Loyon, et al., 2016)*. Hence, a common body of knowledge is necessary to be developed, shared at local and regional levels of countries involved and affected, so as to allow an effective monitoring of cropping and animal production systems, fertilization and water demands, and impacts of climate change. This knowledge would assist policymakers to perform efficient regulatory enforcement considering sustainability and protection of the physical environment. To acquire this knowledge, web applications and mobile apps need to be combined with geospatial and big data analysis *(Mintert, et al., 2016)*, in order to develop online tools that allow policymakers to perceive, visualize and analyze the impact of agriculture, facilitating decision making towards mitigating or eliminating negative effects on the environment. Big data analysis is crucial for analyzing vast amounts of data (e.g. weather, air and water quality, pests and animal diseases etc.) coming from various sources in near real-time. Geospatial analysis is also important for large-scale planning while web/mobile apps can provide access to the users by means of graphical, user-friendly interfaces. In this paper, we describe how the aforementioned technologies (geospatial and big data analysis, web/mobile apps) can be combined to develop an online tool for policymakers (and possibly to farmers in the future as well), towards estimating the impact of agriculture on the environment. This is one of the first initiatives (in the agri domain) combining these technologies together in order to address such a complex multivariate problem. Hence, the contribution of this work is to describe this policy tool (Section 3), focusing on the area of Catalonia, Spain during our analysis efforts (Section 4).

2. Related Work

The majority of software in the agri-domain involves modeling software and simulations for soil, crops, water needs, adaptation to climatic change etc., targeting mostly the farm level *(Holzworth, et al., 2015)*. More related efforts include geospatial platforms and applications focusing on various problems of the agricultural sector, such as facilities management *(Lucas & Chhajed, 2004)*, spatial decision support systems *(Silva, et al., 2014)*, land use and land cover changes *(Embrapa, 2016)*, diseases control and epidemiology *(Cringoli, et al., 2007)* etc. Most relevant work involves geospatial platforms dealing particularly with the impact of agriculture on the environment. Examples include animal manure transportation *(Paudel, et al., 2009)*, selecting sites for safe application of animal waste as fertilizer to agricultural land *(Basnet, et al., 2001)* and environmentally sound management of livestock production *(Jain, et al., 1995)*. Finally, big data management and analysis platforms have recently appeared, focusing either on climate change *(Schnase, et al., 2017)* or earth observation *(Nativi, et al., 2015)*.

3. AgriBigCAT Platform Description

AgriBigCAT is an online software platform combining geophysical information from various diverse sources and web technologies, together with geospatial and big data analysis, in order to estimate the impact of the agricultural sector on the environment, considering land, water, air emissions (e.g. greenhouse gases), biodiversity and natural areas such as forests and wetlands. AgriBigCAT intends to promote more sustainable agriculture, constituting a knowledge-based platform, managing and analyzing a wide range of geospatial and sensory/multimedia information, accessible by standard communication technologies such as the internet/web and mobile apps. Its architecture is depicted in Figure 1.

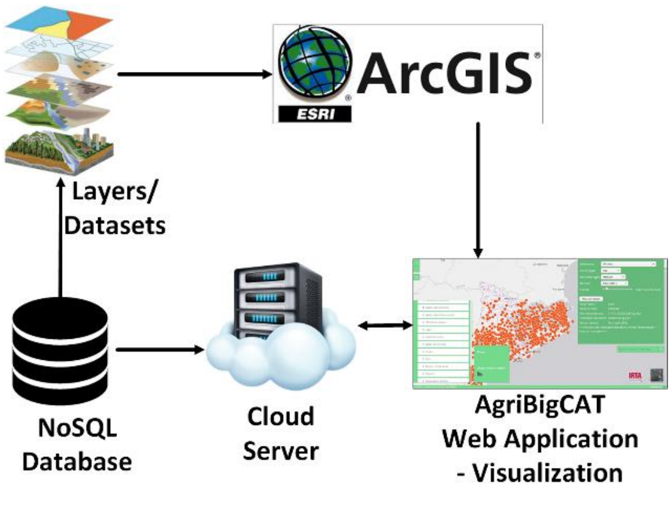


Figure 1: AgriBigCAT Architecture.

The platform allows large-scale data acquisition and analysis of relevant parameters in various agricultural systems in near real-time, facilitating geospatial management and use of inputs (e.g. energy, nutrients and water) and outputs (e.g. emissions, biomass yield, etc.). AgriBigCAT assists both farmers' decision-making and administration planning/policy making, with the objective of meeting the challenge of increasing food production at a lower environmental impact. Agriculture-related datasets are stored using Apache Hive, a (big) database software that allows management of large datasets residing in distributed storage. Visualizations of the datasets as well as geospatial analysis are performed by means of ArcGIS *(Esri, 2017)* and its API for JavaScript, which allows AgriBigCAT to be developed by using open web technologies (e.g. HTML, CSS, JavaScript, AJAX, PHP), but at the same time use the visualization and geospatial features of ArcGIS through its API.

Finally, specific, "heavier" scenarios of geospatial analysis are executed by means of ArcGIS and illustrated through AgriBigCAT in the form of *maps*. In this case, datasets imported into ArcGIS are abstracted as *layers*. Over time, multiple layers become part of a large-scale geo-database for spatiotemporal analysis. Table 1 lists the operations and technologies used by AgriBigCAT, according to big data characteristics as defined in *(Chi, et al., 2016)*. Data sources include geospatial datasets (~GB), sensor measurements (~MB), historical data in various file formats (~MB) and images (~GB-TB). Data veracity (V4) is assured by collaborating with reputable agencies for data collection, such as the Ministry of Agriculture of Catalonia (see Section 4).

Table 1: AgriBigCAT Operations, Technologies and Big Data Characteristics.

No.	Operation	Technology	Big Data Characteristic
1.	Data storage	Apache Hive	Volume (V1), Variety (V3)
2.	Data analysis	ArcGIS Cloud with Hadoop	Volume (V1), Velocity (V2), Variety (V3)
3.	Data visualization	AgriBigCAT	Velocity (V2), Valorization (V5)

4. Analysis

We have primarily focused on the territorial domain of Catalonia (Spain) motivated by the fact that it is one of the European regions with the highest livestock density, with farm concentrations of more than 6M pigs, 0.7M cows and 38M poultry. The high density of livestock in some areas, linked to the insufficient accessible arable land, has resulted in severe groundwater pollution with nitrates *(Nitrates Directive, 1991)*. Excessive soil accumulation of phosphorous and heavy metals from manure has also been reported in certain areas *(Teira-Esmatges & Flotats, 2003)*. The main purpose of this analysis is to optimize manure management by proper planning, aiming to prevent manure land overdosing eliminating negative effects on the environment. The analysis could also result in optimization on the application of agricultural inputs by means of more efficient policies and guidance. As a first step, large and diverse information datasets of the territorial domain of Catalonia were collected, summarized in Table 2. Difficulties in data collection are discussed in *(Kamilaris, 2017)*.

Table 2: Datasets collected for the application scenario.

No.	**Area/Group**	**Datasets**
1.	People, farms and animals	Manure management units, Farmers of Catalonia and numbers/types of animals they possess, Habitats of Catalonia.
2.	Areas and land	Fishing, Evapotranspiration/thermal regions, Nitrate vulnerable areas, Forests, Municipalities, Land parcels, Soils, Crops.
3.	Infrastructures	Wind parks, Road networks, Water/gas pipeline network.
4.	Biodiversity	Vertebrates, Lormophytes, Lichens, Bryophytes, Waterbirds.
5.	Climate and atmosphere	Thermal levels, Temperature, Rainfall, Noise maps, Climate type, Atmospheric emissions (C02).
6.	Water	Wetlands, Zones vulnerable in nitrates, Wastewater plants, Watersheds, Rivers, Reservoirs, Water network, Monitoring stations, Lakes, Water-deficit areas, Coasts, Bays, Aquifers.

As described in Section 3, these datasets were imported as layers in ArcGIS for geospatial analysis, as well as in our NoSQL database for real-time calculations/estimations. Combining these two aforementioned features/services with web technologies, we developed a demo application, available online *(P-Sphere Project, 2017)*, allowing for policymakers, technical advisory services, researchers and environmental scientists in Catalonia to estimate the impact of animal production on the physical environment. Specifically, users can select a particular area on Catalonia (or all areas), particular animal type (e.g. pigs, dairy cows, poultry, beef cattle or all animals), and emission type (e.g. carbon dioxide, methane, nitrous oxide, ammonia) or nitrogen/phosphorous excreted, and the platform calculates the farms and animals involved in the query, manure produced and estimated emissions/excretions in monthly or yearly basis, taking into account the existing weather and thermal conditions at the selected time period from historical data. Calculations are based on the *(IPCC, 2006)* guidelines (TIER1) while data about farms and animals have been provided by the Ministry of Agriculture of Catalonia (for the year 2016). Calculations also consider animal feeding practices widely used in the area, as well as popular manure management techniques employed in Catalonia. All assumptions involved are listed on the website, next to each visualization.

Figure 2 illustrates a snapshot of the demo application, where farm locations and rivers of Catalonia are visualized, in order to examine the impact of pigs' manure in relation to possible contamination of nearby rivers with nitrates.

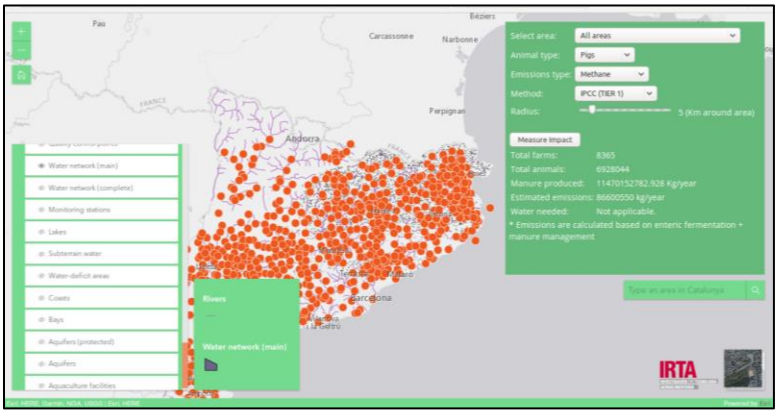
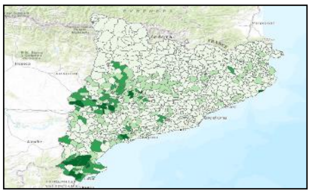
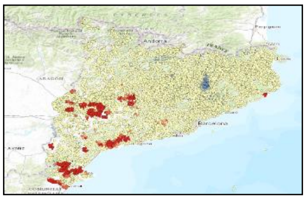
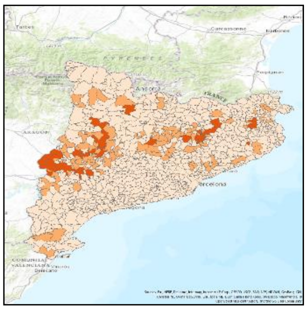
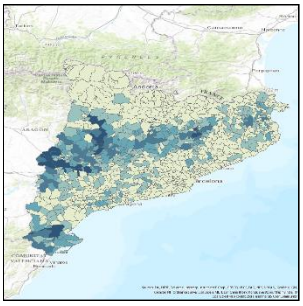
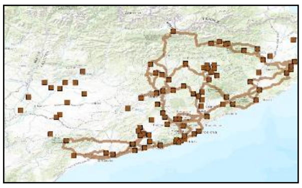


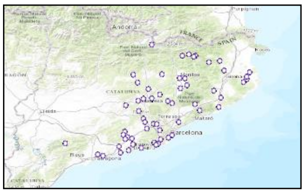
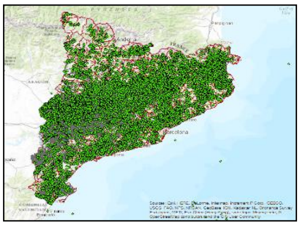
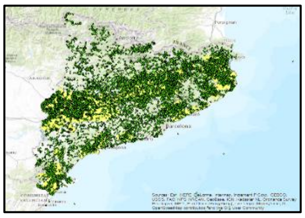
Figure 2: Snapshot of the demo application

Table 3 lists some visualizations and explanations of the geospatial problems addressed through our application. Our findings are being continuously reported to the Ministry of Agriculture of Catalonia for further actions.

Table 3: Geospatial scenarios in Catalonia and visualizations/solutions.

	Visualization of farms in Catalonia according to the type of animals they grow (i.e. pigs, dairy cows, cattle beef, poultry etc.). In this particular scenario, pig farms are displayed, which constitute the most popular farming industry in the region.

	Farms per municipality according to number of animals. Darker green means a larger concentration of farms/animals. Same maps have been created per animal type as well. The south-west part of the region has the highest concentrations.
	Hot spots of farms according to the total number of animals they have. It provides an indication of where high/low numbers of animals cluster spatially. Same maps have been created per particular animal type as well.
	Summary of yearly methane emissions (TIER2) at each municipality, based on existing farms/animals. Darker orange means a larger production of methane emissions, while the darkest one indicates that caution is required. Similar maps have been produced for other emissions such as nitrous oxide and ammonia. Same maps have been created per particular animal type as well. Calculations are based on the existing manure management treatment methods in the region.
	Summary of yearly nitrogen excreted from animals' manure at each municipality of Catalonia, based on existing farms and animal types. Similar maps have been produced for phosphorous too. Darker blue means a larger amount of nitrogen excreted. Same maps have been created per particular animal type as well. Calculations are based on IPCC TIER1. Similar maps will be created for the potential reduction of nitrogen/phosphorous by applying best practices (future work).
	Which are the best routes to connect farms together, e.g. for finding the best transportation solution to collect and carry manure by means of trucks? The best routes (i.e. minimizing total travel time/distance) are displayed in light brown color for the case of poultry farms. Similar maps have been created for other animal types and for all farms/animal types.

	According to the structure of the existing manure management plants, which are the closest plants that can serve the animal farms in Catalonia to dispose their manure? The best plans are displayed as yellow rectangles for the case of poultry farms. Similar maps have been created for other animal types and for all farms/animal types.
	If we selected specific municipalities/locations to build manure processing units, considering their proximity to farms, where would these be? The best options (i.e. minimizing total travel time/distance) for the case of poultry farms are displayed as blue circles on the map. Similar maps exist for other animal types and for all farms.
	Which are the possibilities of using manure as crops' fertilizers? Is there an actual need per municipality according to crops produced and animals' manure being produced at the same time? The picture shows a matching between livestock farms and nearby crop-based farms that could use manure as fertilizer. (Work in progress)
	In which degree are areas vulnerable of nitrates affected by livestock farms? In which degree are rivers and lakes (water vulnerable zones) affected? The picture shows livestock farms and their proximity/impact to nearby rivers/lakes and nitrate vulnerable areas. (Work in progress)

Specific details over each of the above geospatial scenarios (e.g. methods, analysis, inputs, parameters, results) are available on the project's website *(P-Sphere Project, 2017)*. Some "sensitive information" is currently available only to authorized users, such as employees of the Department of Agriculture.

5. Conclusion

This paper has described AgriBigCAT, an online software platform combining geospatial and big data analysis, together with web technologies, to estimate the impact of the agricultural sector on the environment. Serving as a knowledge-based platform, it constitutes a useful tool for administration planning and policy making, contributing to the challenge of increasing food production at a lower environmental impact. Moreover, an online application of AgriBigCAT, focusing on the local environmental issues of the agricultural sector of Catalonia, has been presented and described. As future work, we plan to examine best available techniques (BAT) available in agriculture for animal/manure management, and assess the potential environmental benefits (considering costs' trade-offs) of their application in Catalonia (or other regions considered in the future through our platform).

Acknowledgments

This research has been supported by the P-SPHERE project, which has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skodowska-Curie grant agreement No 665919.

References

Basnet, B. B., Apan, A. A. & Raine, S. R., 2001. Selecting suitable sites for animal waste application using a raster GIS. *Environmental Management,* 28(4), pp. 519-531.

Bruinsma, J., 2003. *World agriculture: towards 2015/2030: an FAO perspective,* s.l.: Earthscan.

Chi, M., Plaza, A. & Benediktsson, J. A., 2016. Big data for remote sensing: challenges and opportunities. *Proceedings of the IEEE,* 104(11), pp. 2207-2219.

Cringoli, G. et al., 2007. Geo-referencing livestock farms as tool for studying cystic echinococcosis epidemiology in cattle and water buffaloes from southern Italy. *Geospatial Health,* 2(1), pp. 105-111.

Embrapa, 2016. *SOMABrasil.* [Online]
Available at: http://mapas.cnpm.embrapa.br/somabrasil/webgis.html

Esri, 2017. *https://developers.arcgis.com/javascript/.* [Online]
Available at: http://www.esri.com/arcgis/about-arcgis

Heinrich Böll Stiftung, 2014. Meat Atlas: Facts and figures about the animals we eat. *Heinrich Böll Stiftung. Friends of the Earth Europe.*

Holzworth, D. et al., 2015. Agricultural production systems modelling and software: current status and future prospects. *Environmental Modelling & Software,* Volume 72, pp. 276-286.

IPCC, 2006 . *Chapter 10: Emissions from livestock and manure management,* s.l.: IPCC Guidelines for National Greenhouse Gas Inventories.

Jain, D. K., Tim, U. S. & Jolly, R. W., 1995. A spatial decision support system for livestock production planning and environmental management. *Applied engineering in Agriculture,* 11(5), pp. 711-719.

Kamilaris, A., 2017. *Big data analysis and Integration of Geophysical information from the Catalan Agri-Technological sector,* Barcelona: Interest Group on Agricultural Data (IGAD), Available at: https://www.slideshare.net/kamiWeb/big-data-analysis-and-integration-of-geophysical-information-from-the-catalan-agritechnological-sector.

Loyon, L. et al., 2016. Best available technology for European livestock farms: availability, effectiveness and uptake. *Journal of Environmental Management,* Volume 166, pp. 1-11.

Lucas, M. T. & Chhajed, D., 2004. Applications of location analysis in agriculture: a survey. *Journal of the Operational Research Society,* 55(6), pp. 561-578.

Mintert, J. et al., 2016. *The challenges of precision agriculture: is big data the answer?,* San Antonio, Texas: Southern Agricultural Economics Association Annual Meeting (No. 230057).

Nativi, S. et al., 2015. Big data challenges in building the global earth observation system of systems. *Environmental Modelling & Software,* Volume 68, pp. 1-26.

Nitrates Directive, 1991. Council Directive 91/676/EEC of 12 December 1991 concerning the protection of waters against pollution caused by nitrates from agricultural sources. *EUR-Lex,* 375(31), pp. 1-8.

Paudel, K., Bhattarai, K., Gauthier, W. & Hall, L., 2009. Paudel, Krishna P., et al. Geographic information systems (GIS) based model of dairy manure transportation and application with environmental quality consideration. *Waste Management* , 29(5), pp. 1634-1643.

P-Sphere Project, 2017. *AgriBigCAT.* [Online]
Available at: http://www.psphere-project.com/irta/

Schnase, J. et al., 2017. MERRA analytic services: Meeting the big data challenges of climate science through cloud-enabled climate analytics-as-a-service. *Computers, Environment and Urban Systems,* Volume 61, pp. 198-211.

Silva, S., Alçada-Almeida, L. & and Dias, L. C., 2014. Development of a Web-based Multi-criteria Spatial Decision Support System for the assessment of environmental sustainability of dairy farms. *Computers and Electronics in Agriculture,* Volume 108 , pp. 46-57.

Stoate, C. et al., 2009. Ecological impacts of early 21st century agricultural change in Europe–a review. *Journal of environmental management,* 91(1), pp. 22-46.

Teira-Esmatges, M. & Flotats, X., 2003. A method for livestock waste management planning in NE Spain. *Waste Management,* 23(10), pp. 917-932.