Data Management and Decision Making:
When Precision Agriculture puts us in Trouble

The use of sensors in Precision Agriculture together with the trend of increasing spatial and temporal sampling resolutions, lead to the generation of huge amounts of data. The appropriate management of such data sets to turn them into useful information to be used in the decision making process is quite a challenge. This is addressed in Stage 3 of the Precision Agriculture cycle we presented in the first Precision Ag Corner (New Ag International 64 in Nov-Dec 2016).

In this issue, we describe how to handle and use the gathered data to make more informed agronomic management decisions. Undoubtedly, this is the moment of truth: when PA puts us in trouble and we have to make an agronomic decision! New Ag International has partnered with the Research Group on AgroICT & Precision Agriculture (GRAP) of the University of Lleida-Agrotecnio Center in Catalonia, Spain. In every issue of the magazine Jaume Arnó, José A. Martínez-Casasnovas and Alexandre Escolà, with our Editorial Team, put together an editorial whose ambition is to help the various stakeholders bridge the gap between datanomics and commercial farming!

THE USE OF PROXIMAL AND REMOTE SENSORS allows a lot of data from the crop and its environment to be collected. Crop vigour, health, water status, canopy evolution, weeds, pest and disease, soil parameters and weather data are examples of what is possible to be measured by sensors today. Additionally, in Precision Agriculture all these data are georeferenced in order to assess the spatial variability of agronomic parameters affecting crop production and/or quality. Moreover, it is getting very easy and economical to have several measurements of either terrestrial and/or airborne sensors along the season. The improvement of the spatial and temporal resolutions of sensors implies increasingly bigger data sets that may accumulate up to several hundreds of Gigabytes of digital data per field in a season (e.g. when using terrestrial laser scanners the amount of data in acquired point clouds may easily reach some Gigabytes (109) per field, but the resulting map can be reduced down to some Megabytes (106)). To be classified in the Big Data domain, at least 2 “V” are required so we need to increase Volume, at least. Therefore, only when Precision Agriculture techniques are implemented in a region, one can think about Agro Big Data. The continuous monitoring of a large area allows for the estimation of yields (to deal with logistics, commercialization and food security issues), agricultural input needs (such as water, fertilizer or plant protection products) and weather forecasting (to deal with insurance related to risks and hazards, climate change, etc.). Open remote sensing data is currently being used in Agro Big Data since several hundreds of Terabytes (1012) or even Petabytes (1015) of Earth images are generated on a daily basis. However, big data in agriculture is not still as deployed as it is in the domain of social network analysis, for instance. One of the reasons may be that there is an important lack of proximal sensing data to combine with remote sensing which is either private or non-existent. Recently, big machinery and agricultural input companies are deploying either self-developed or acquired cloud platforms to gather as much data as possible. Internet of things (IoT) will certainly increase the databases. But they are currently not compatible making a general implementation more difficult. The use of these platforms arise some concerns: who’s the data owner? Who will use the data? With what purpose? This is also connected to ethical issues. Nevertheless, before focusing on large areas, in the Precision Ag Corner we want to address how to make more informed decisions related to in field management.

FIRST DECISION: UNIFORM OR VARIABLE RATES?
The first decision a farmer should consider is whether it is worth spending time and money doing Precision Agriculture. It is not a simple decision. Some researchers and agronomists have raised this in terms of opportunity. Is there enough spatial variability in a plot that justifies a differentiated management by zones within a field? Even more, is such variability conveniently structured by presenting a pattern that can be agriculturally manageable? In other words, will this pattern of variability allow the delineation of well-differentiated management zones...
compatible with the variable dosage requirements of agricultural machinery? To deal with these dilemmas, different technical opportunity indices (OIs) have been recently proposed. In this issue, we will not expand on making a list of the different OIs options. Rather, the interest lies in knowing what these indices are based on, the benefits of their application, and if there are still certain drawbacks pending to be solved and improved.

In general, the proposed OIs are based on two components (at least, the first published versions): field variability (or spatial variability) and, as it has already been mentioned above, the spatial distribution of this variability (or spatial structure). Under this dichotomous scheme, the idea behind an OI is simple: the plots that present opportunity for Precision Agriculture are only those that meet certain requirements (thresholds) in the two components. Obviously, how these components are formulated and, above all, what are the specific thresholds to be met in each plot are both key factors in getting useful and applicable results. Generalists OIs are those that, being able to be applied to different crops (arable crops or fruit trees), allow to adapt to zone management by providing indications of opportunity for a wide range of the most common mechanized operations in a farm. However, there are other options where evaluating the opportunity is limited to a single crop and a single mechanized operation. This is the case of the OI developed by the Research Group on AgroICT and Precision Agriculture of the University of Lleida (Catalonia, Spain) to assess the opportunity for selective wine grape harvesting. This OI is novel in that the index is only based on vigour maps derived from high-resolution remote sensing data. It is known in viticulture the presence in many plots of consistent spatial correlations between grape yield, vineyard vigour and some grape quality parameters. Then, given the price differentiation between grapes of different quality, selective harvesting based on prior classification of vigour maps may produce benefits by separating homogeneous blocks of varying grape quality within the plots. Moreover, winemakers using the OI can identify those areas within the farm that, predictably being of different quality, could be selectively harvested to fit the productive plan and winemaking needs of the winery.

To show an application example, Figures 1 and 2 were obtained for a big winery in Raimat (Lleida, Catalonia, Spain) to classify 36 vineyard plots according to their opportunity for selective harvestings. The goal was to establish a ranking of vineyards within the farm according to their index values. Since winemakers need to plan production according to grape varieties and quality, advanced information about fields that can be harvested evenly and non-homogeneous fields having appropriate characteristics for selective harvesting is of great importance. Assuming that vines with lower vigour (low NDVI values or values of the Normalized Difference Vegetation Index) produce higher quality grapes, only fields with a large magnitude of variation in the vigour of vines would present distinct grape quality. Then, applying a geostatistical analysis to the NDVI map obtained at veraison, both the spatial variability and the spatial structure components were evaluated to differentiate plots without and with opportunity. The index is aimed at identifying the latter, namely fields that can provide the required amounts of grapes and, by clearly defining zones of different grape qualities. They also optimize the operation and working time of grape harvesters. Successful application is expected. First, because the index allows the demands of other wineries and their particular management logistics to be adjusted. Second, because it makes use of software and technologies that are becoming increasingly known in the agricultural sector, such as remote sensing and geographic information systems (GIS) which can also be found as open source tools.

This is just an example. The reader can draw their own conclusions and imagine similar applications in fertilization, sowing or application of plant protection products. More information about this opportunity index can be found in the open access article published in Cogent Food & Agriculture by Arnó, Martínez-Casasnovas and Tejada-Moral in 2017. (https://www.tandfonline.com/doi/abs/10.1080/23311932.2017.1386438).

**MANAGEMENT ZONES: A MATTER OF DATA CLASSIFICATION**

How to turn data into usable agronomical maps? In the March/April 2018 issue of the magazine we introduced this key concept of data classification which is essential to achieve the benefits of Precision Agriculture according to the demand of the specific vineyard plot. Moreover, it is important to establish criteria that can be used to make decisions and quantify the decision making process, especially when the final application is grower driven and technology and information systems (GIS) can also be found as open source tools.

*demand of the specific vineyard plot*
An interview with Joel Wipperfurth, Director of Ag Technology, WinField United

“WinField United will continue to equip ag professionals with tools to access and leverage the power of data, now and in the future”.

WinField United is one of the largest U.S.-based providers of crop inputs, data, technology tools and agronomic expertise. Today’s agriculture technology marketplace resembles a train station where each traveller speaks a different language and is trying to reach a different destination. These travellers are perfectly capable people. But because they don’t understand each other, they can’t make informed decisions about which train is the right one for them. As a result, no one gets anywhere. What does a transit hub have to do with ag technology? Well, many precision ag tools can’t communicate with each other. And even when connectivity is possible, multiple steps may be necessary to achieve it. At WinField United, we work with ag retailers to help their farmer-customers choose and implement the right ag technology tools for their particular operations.

Our suite of ag technology tools facilitates smoother, more complete data management and more informed in-season decision-making — even if some of the tools we’re managing aren’t our own. For example, one of our solutions is the Answer Tech Data Silo™ — a cloud-based management system that helps retail agronomists store and synchronize farmer data. We recognize that a farmer may use several digital tools. If something like a field boundary needs to be changed, Data Silo enables this update to be made across all of that farmer’s systems, keeping everything in sync.

I believe that in the future, ag technology will include more sources of data from equipment sensors, in field observations, and the continued datafication of a farm’s cultural practices which will need to be conveniently parsed and supplied across multiple software systems farmers may use. These sources of data have the potential to help farmers not just plant one crop per year, but run thousands of scenarios to optimize outcomes while the crops value and potential undulate throughout the season. As the power of connecting data sources moves forward there is potential beyond growers sharing data in return for agronomic advice. Data monetization could be a near term opportunity as commodities move through the food supply chain, insurance companies look to underwrite risk, or traders looking for the edge on markets are allowed access to data with the growers choosing. There will no doubt be an evolving definition of privacy, but trust and transparency will be cornerstones of any business looking to access these rich and fertile data layers.

Winning growers and retailers will be defined by ag technology. WinField United will continue to equip ag professionals with tools to access and leverage the power of data, now and in the future.

process in the Precision Agriculture data flow. In addition to data acquisition, this process involves several steps, such as data cleaning, interpolation (in case of variable data acquired in discrete points) and classification of the spatial within-field variability into different potential management classes. The delineation of these potential classes can be done “manually” according to the expert knowledge of farmers and/or technicians, or automatically by applying classification algorithms. Nevertheless, this procedure may not assure that the created classes are statistically different. Other additional disadvantages are that the farmer/technician can hardly delineate the potential zones based on more than one variable at a time, since the delineation is by visual interpretation of maps; additionally,

Figure 2: Selective harvesting opportunity in the 36 vineyard plots in Raimat (Lleida, Catalonia, Spain). Green fields indicate opportunity for PA (Arnó et al., 2017).
the process is time consuming, particularly when the farm has a lot of fields to be classified. Because of that, and the new era of Big Data, automated classification procedures may be preferred.

The two most frequently used algorithms are k-means and ISO-DATA. Both are similar, but with the difference that the ISODATA algorithm allows for different number of clusters while the k-means assumes that the number of clusters is known a priori. These algorithms result in a “hard” map of classes, that is, each point of the field is classified into a single class. Nevertheless, sometimes there are points that are “in between” two classes and, in some situations, they could be classified as belonging to class “A” or to “class B”.

This indeterminate assignation of classes can be solved with the aid of the so called “fuzzy” algorithms, such as the fuzzy c-means algorithm. This method allows more than one class or cluster to be assigned to each individual point of the field.

**Figure 3:** Prescription map to plant maize with different seeding rates according to the productive potential of each zone. The delineation was made on the basis of the apparent soil electrical conductivity (A), the yield map of the precedent crop (winter wheat, B), and the expert knowledge of the farmer and agronomist of the farm about the soils (texture, salinity, drainage). The recommends doses are expressed in thousands of seeds per hectare (k seeds/ha).
point of the field, according to its characteristics and with different degrees of belonging. One example of software based on fuzzy classification is the Management Zone Analyst (US Department of Agriculture). It performs an unsupervised fuzzy classification procedure for a range of cluster numbers, and provides the user with two performance indices (fuzziness performance index and normalized classification entropy) to aid in deciding how many clusters are most appropriate for creating management zones. That is, based on the different input variables (e.g. apparent soil electrical conductivity, vegetation indices, yield maps, elevation, slope, etc.), the program indicates if a field should be divided into either two, three or four management zones. The application of this type of classification in Precision Agriculture is particularly interesting, but it does not solve one the key question in PA yet: (and perhaps the most difficult one) what to do in each zone?

**THE MOST DIFFICULT: WHAT TO DO IN EACH ZONE?**

To begin with, let’s take a plot of an arable crop with opportunity for nitrogen fertilization according to potential management zones previously delineated by applying an appropriate classification criterion. Once arrived at this point, what N dose should be applied in each zone? Undoubtedly, this is the moment of truth: when PA puts us in trouble and we have to make an agronomic decision. The conversion from data to agronomic decisions is a difficult task. The use of agronomic models can help us, and this explains why researchers and agronomists are devoting so much effort in experimenting with sensors and models of crop response. An example is the development of algorithms to estimate N fertilizer requirements using terrestrial optical sensors working in the red and infrared bands of the electromagnetic spectrum. Under this approach, it is worth mentioning the web-based application developed by Oklahoma State University, in order to take advantage of using the GreenSeeker™ handheld sensor through the subsequent implementation of a free, online sensor-based nitrogen recommendation calculator (Sensor-Based Nitrogen Rate Calculator, SBNRC). The method is based on mid-season measurements with the sensor (NDVI readings) and comparing the reflectance from unfertilized plants to those in a non-limiting N strip (N-rich strip) placed in the same field. To run the algorithm properly, it is also important to know the number of days from planting (or days with positive values of GDD, growing degree days). Thus, dividing the non-fertilized NDVI reading by the number of days, an estimate of the potential yield (YP 0) can be obtained, as this ratio provides the biomass produced per day on the specific date that the readings are collected. Yield potential (YP N) attainable when N fertilizer is applied is then calculated multiplying YP 0 by RI (response index in yield to additional N fertilizer). The RI is essentially the NDVI from the N-rich strip divided by the NDVI from the rest of the field (or farmer practice). Assuming that YP N does not exceed the maximum yield that can be obtained for a given area or region, the N rate recommendation can be finally obtained as:

\[
\text{Napplied} = P_a \times (\text{YP}_N \cdot \text{YP}_0)
\]

NUE

where Napplied is the N application rate (kg/ha); \(P_a\) is the percentage N in the grain; and NUE is the N use efficiency. Indeed, the farmer only has to enter a few data into the SBNRC to obtain as outputs the two yield potential estimates (YP 0 and YP N) as well as a field-specific N rate recommendation. This approach can be applied both for uniform management and for adjusting prescription maps based on areas of differentiated management of nitrogen fertilization. Drawbacks are also present. The algorithm is made region or site-specific, so the calibration coefficients used in each case should be checked. In addition, the farmer must be willing to invest time in the experimental design that is required in his/her own farm for NDVI readings. This ‘new’ approach of experimentation carried out by the farmer him/herself should be considered as a good tool to help improving soil and crop management. The value of ‘on-farm experimentation’ is to gain information that farmers can trust. Many times what works well in one farm does not necessarily work in another. At this point in the decision-making process, the use of the results of ‘on-farm experimentation’ (as referenced above) can be the key that, together with the farmers’ experience, contributes to reducing uncertainties to make informed decisions adapted to uncontrolled site-specific factors that occur in the agricultural environment.

**Glossary of terms**

**Uniform or variable rate:** Uniform or constant rates are used in conventional agriculture, where the management unit is the field. Precision Agriculture works on a scale basis smaller than the field. Delineating different management zones within the field with differential rates on a site-specific crop management basis implies using a variable rate approach.

**Potential management zones:** Areas in which a field is divided to carry out a differentiated agronomic operation with different doses in each zone. In general, a field is classified into potential management classes, each class including one or more zones within the field.

**Opportunity index:** Numerical value that allows fields to be ranked by the opportunity they present to do Precision Agriculture. The index usually evaluates the technical opportunity, that is, if fields have enough spatial variability and allow zones to be well delimited and managed according to the available variable rate technology.

**On-farm experimentation:** Using their own farm to carry out experiments related to fertilizing, seeding, irrigation, etc. with the proper statistical support in order to obtain specific coefficients to adjust general models to the farm local conditions.

**Agro Big Data:** The use of Big Data techniques to obtain, analyse, visualize and valorise agricultural data meeting the 3 "V" condition: big volume, high variety and high velocity. Generally, the data should cover a large area to achieve enough volume of data. The results could be used in yield estimation, agricultural input needs determination and risk assessment.