

# Principles of Data Consolidation

Area 2 – Technologies  
Lesson 8 – Data Analysis  
Sequence ID – 27a

UPM





## DISCLAIMER

### A2.L8.T2 Principles of data consolidation

Pilar Barreiro, [pilar.barreiro@upm.es](mailto:pilar.barreiro@upm.es), Universidad Politécnica de Madrid, Spain , [0000-0003-4702-6059](tel:0000-0003-4702-6059)

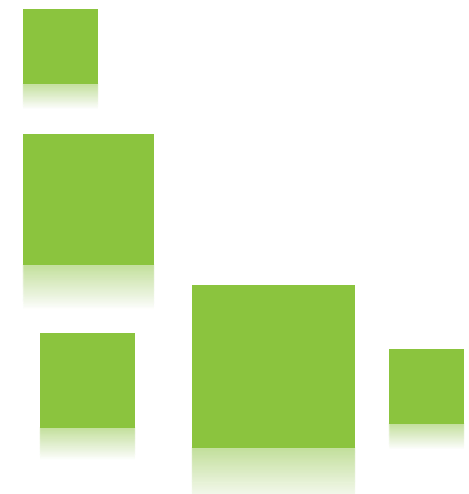
Eva Cristina Correa, Universidad Politécnica de Madrid, Spain

Belén Diezma Iglesias, [belen.diezma@upm.es](mailto:belen.diezma@upm.es), Universidad Politécnica de Madrid, Spain

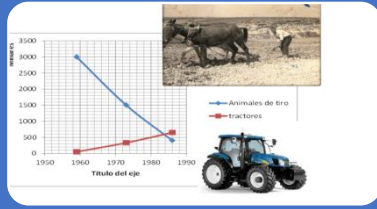
Pilar Barreiro, Eva Cristina Correa, Belén Diezma Iglesias, *Principles of data consolidation*, © 2020 Author(s), [CC BY-SA 4.0 International](#), [DOI 10.36253/978-88-5518-044-3.28](https://doi.org/10.36253/978-88-5518-044-3.28), in Marco Vieri (edited by), *SPARKLE - Entrepreneurship for Sustainable Precision Agriculture*, © 2020 Author(s), [content CC BY-SA 4.0 International](#), [metadata CCO 1.0 Universal](#), published by [Firenze University Press](#), ISSN 2704-6095 (online), eISBN 978-88-5518-042-9, [DOI 10.36253/978-88-5518-044-3](https://doi.org/10.36253/978-88-5518-044-3)



# The Decision Making Process



## AGRICULTURE 0.0 (before the Industrial Revolution)



- Before mechanization
- Before the Green Revolution

## AGRICULTURE 1.0 (Industrial Agriculture)



- Tractors, Combine harvesters, Machines in general
- Outsourcing inputs

## AGRICULTURE 2.0 (Advanced Mechanization, Precision Agriculture)



- On board technology (guidance, electronic control, VRT, ISOBUS)
- Remote sensing (satellites, aerial, UAV)
- Sensor Networks

## AGRICULTURE 3.0 (Digital Agriculture)



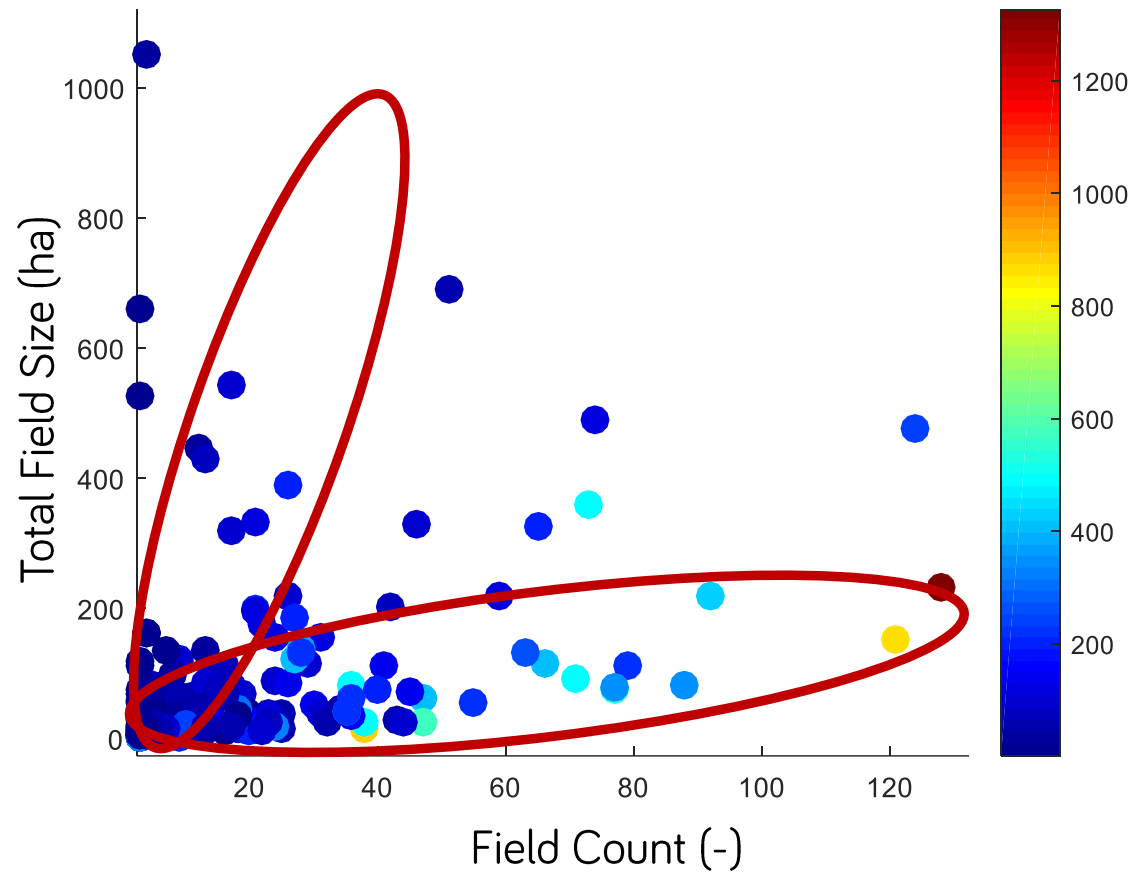
- GIS
- Digital notebook, Decision support systems
- Fleet management

## AGRICULTURE 4.0 (Permanent network of information exchange)



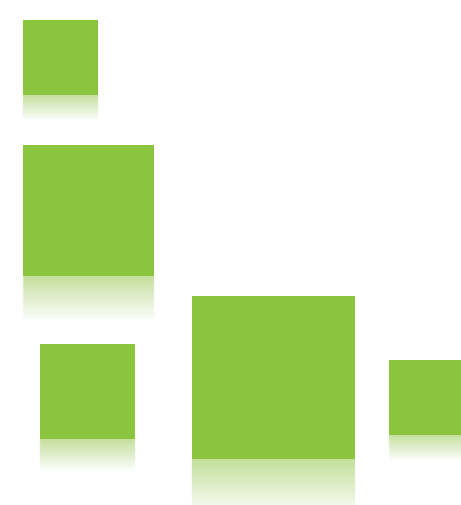
- Smart digital ecosystems (system of systems)
- Smart integration of bits and bytes
- Ubiquitous cyberspace (30Mbytes/s in all rural areas by 2020)

# Influence of Farm Structure on Decision Making

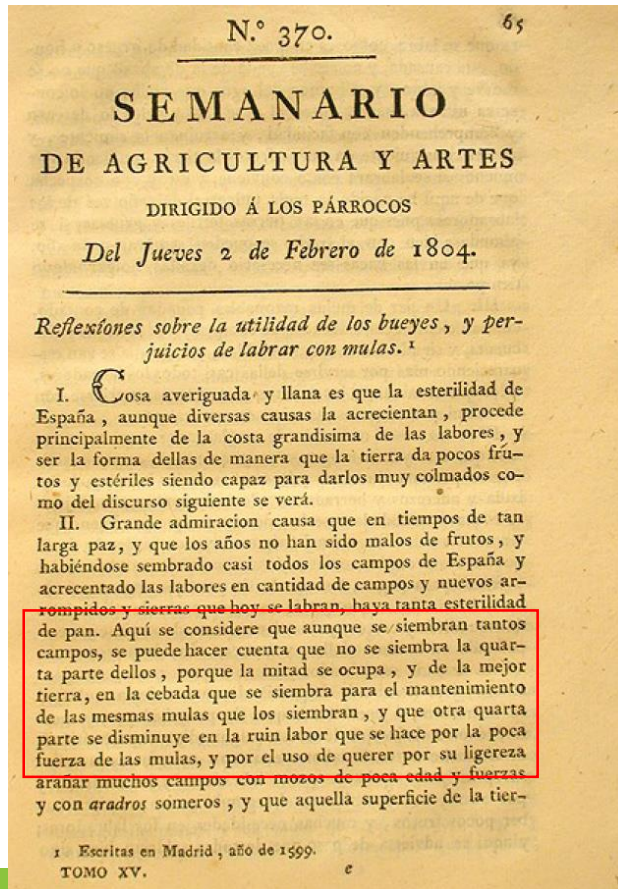




# The Beginnings of Farming Records



# Old and New Dilemmas

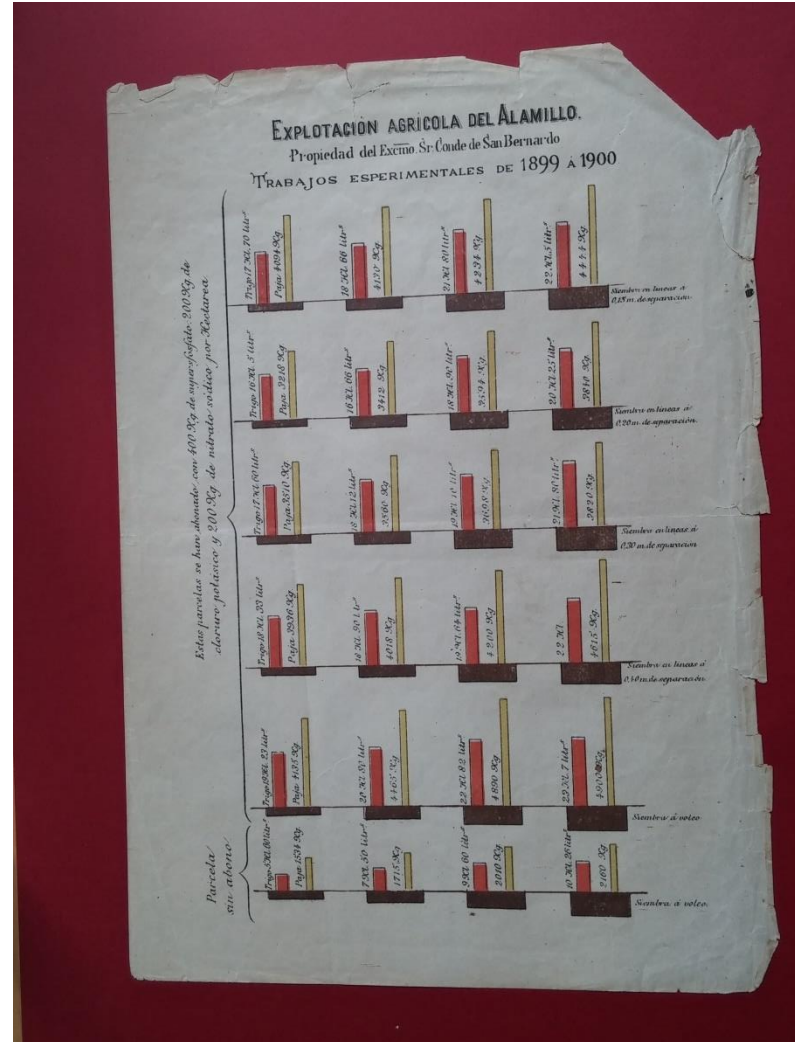


# Even the Simplest Records are Crucial



5 parcels fertilised with superphosphates

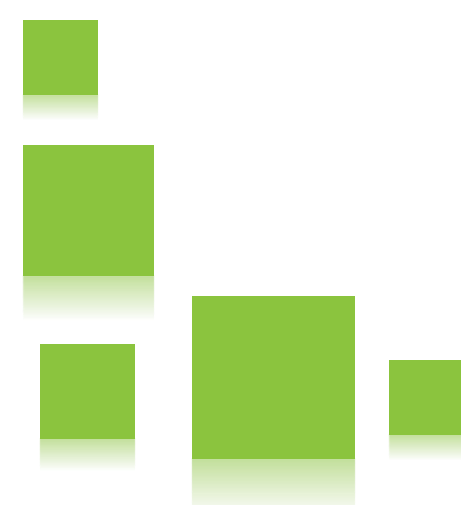
1 parcel without fertilisation





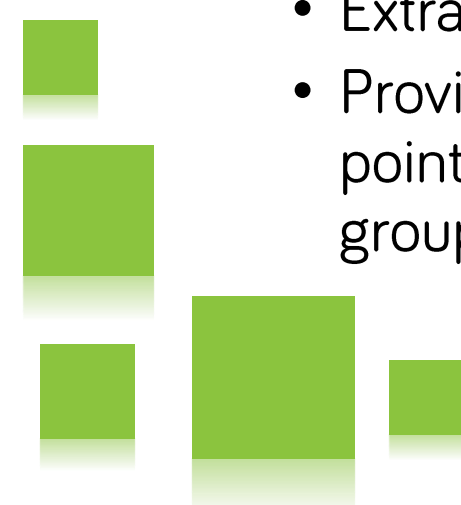


# A / B Comparisons in the Era of Digital Information

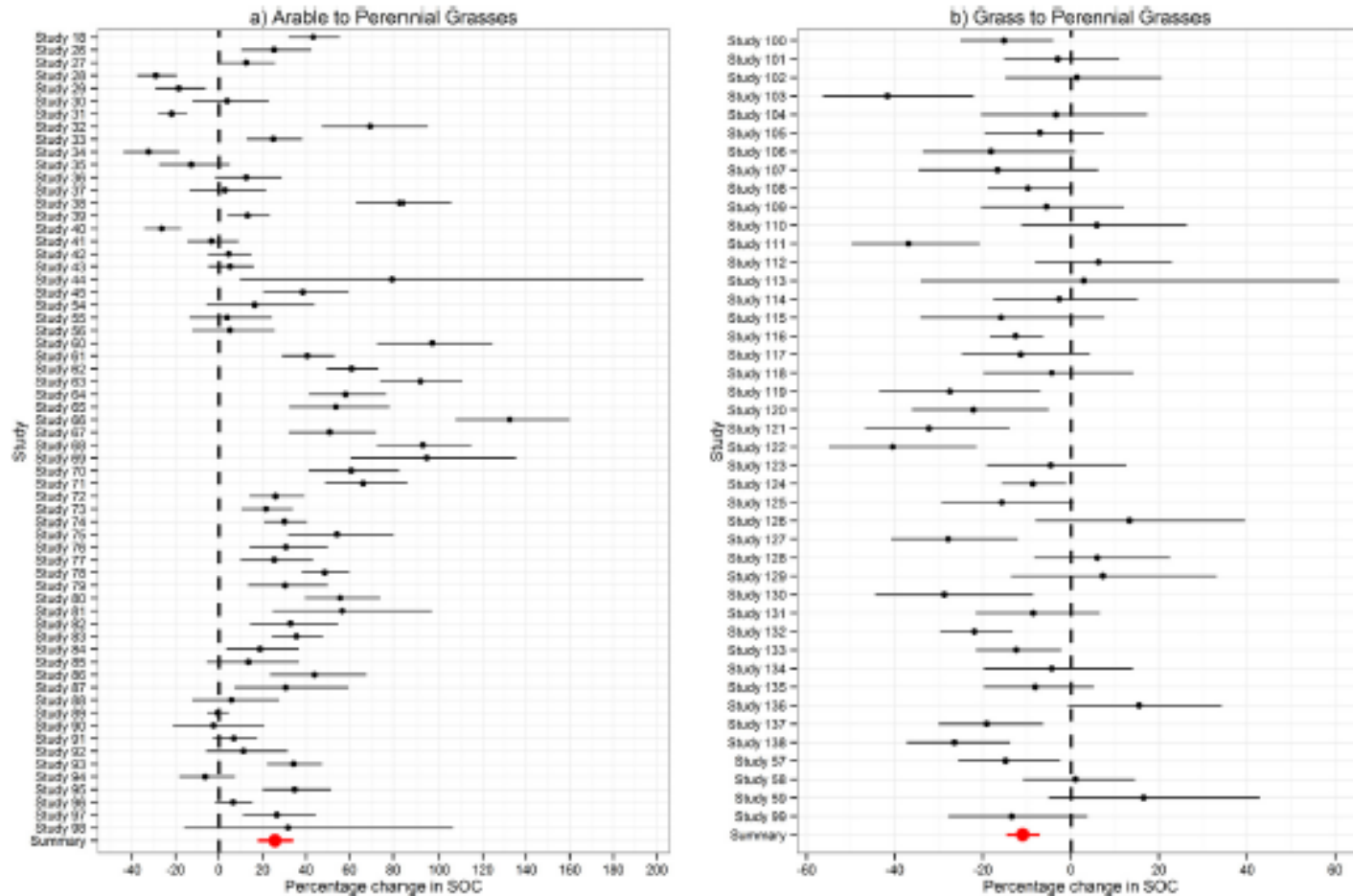


# The Concept of Meta-Analysis

- Perform a systematic document search
- Set the criteria to apply a study to your specific situation, they should:
  - Have adequate metrics that quantify a response
  - Take into account the transition period, quantifiably, relating the results to both the periods before and after the transition
  - Specify the location
  - Extract the data in predefined (standard) units
  - Provide three basic values: mean, standard deviation, and the number of data points, to allow comparison of the effectiveness of the treatment and the control group

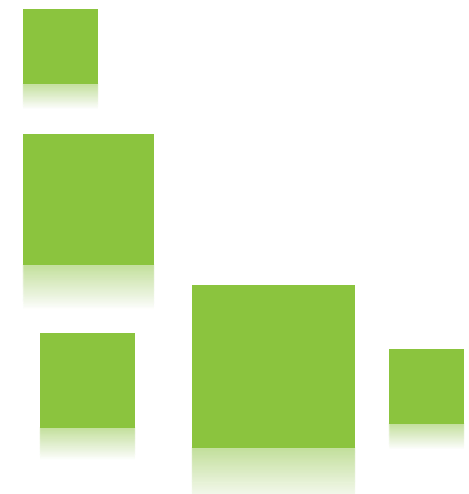


# Example of Meta-Analysis



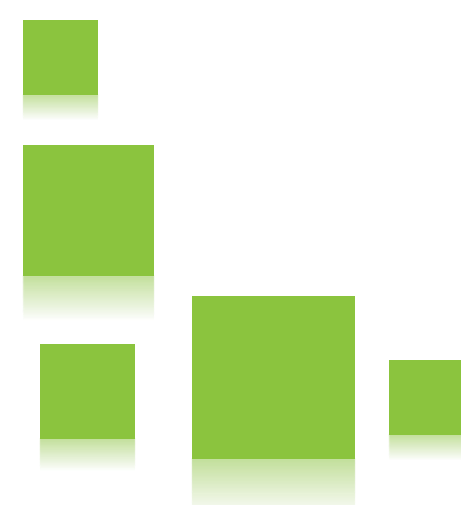
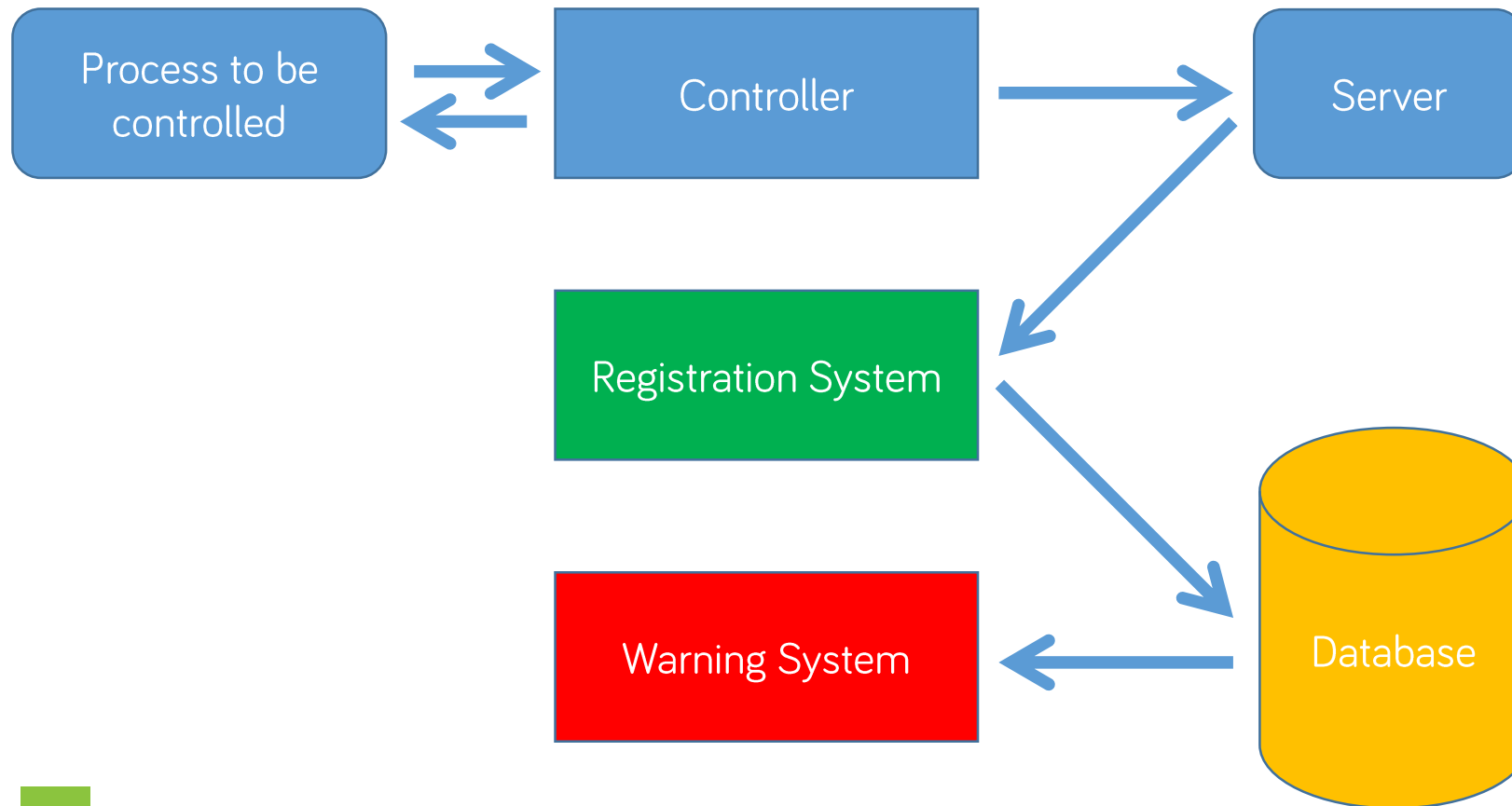


# SCADA Systems and the Data Boom



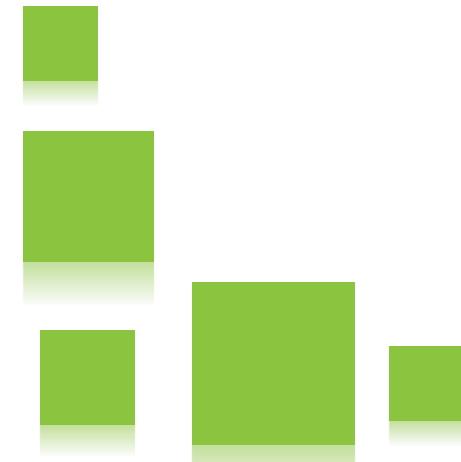
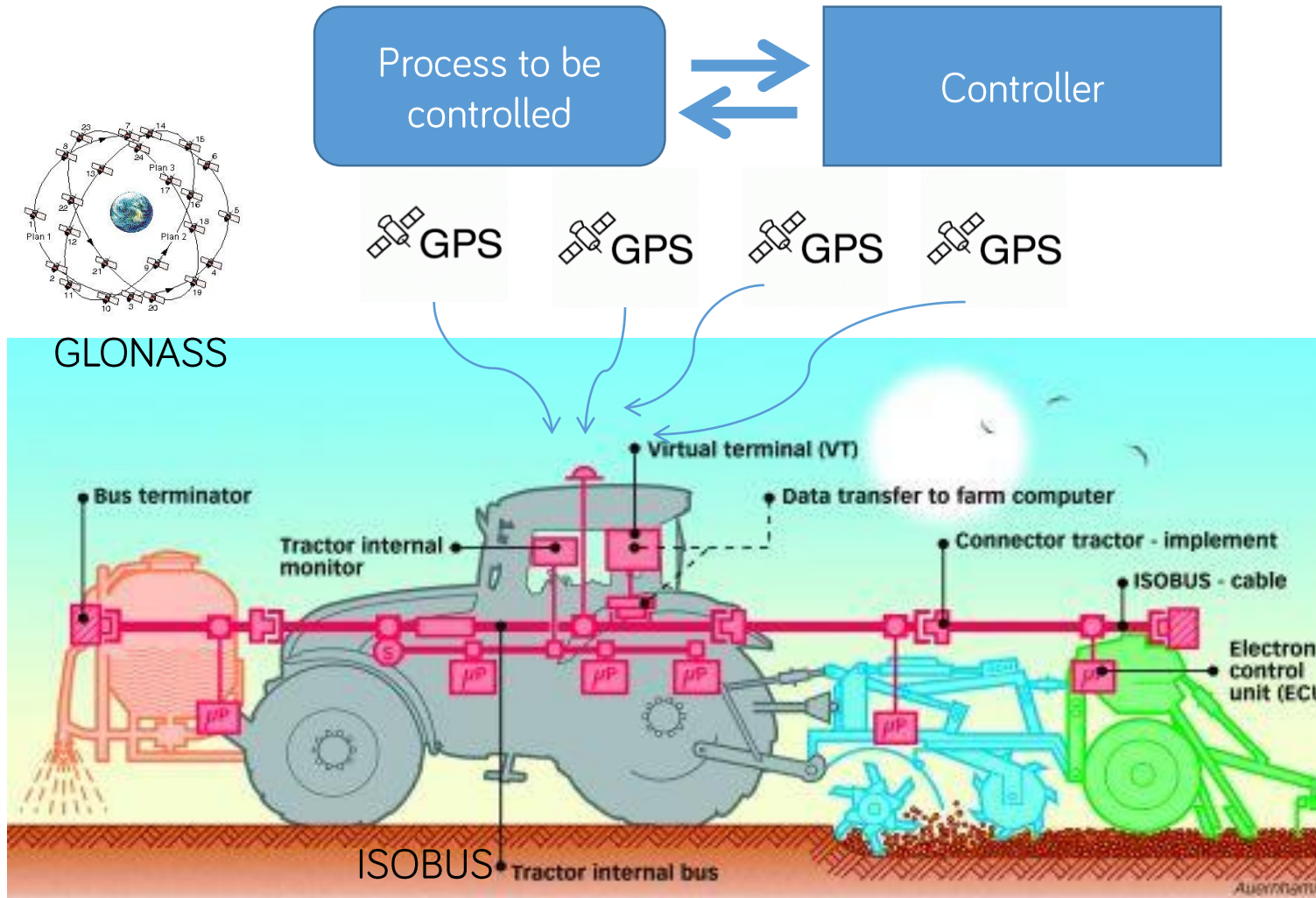
# What is SCADA?

(Supervisory Control And Data Acquisition)



# SCADA in On-Board Technologies

Incomplete SCADA System: only Supervisory Control



# Example: Distance Proportional Flow

Aguirre Experiment, 2015



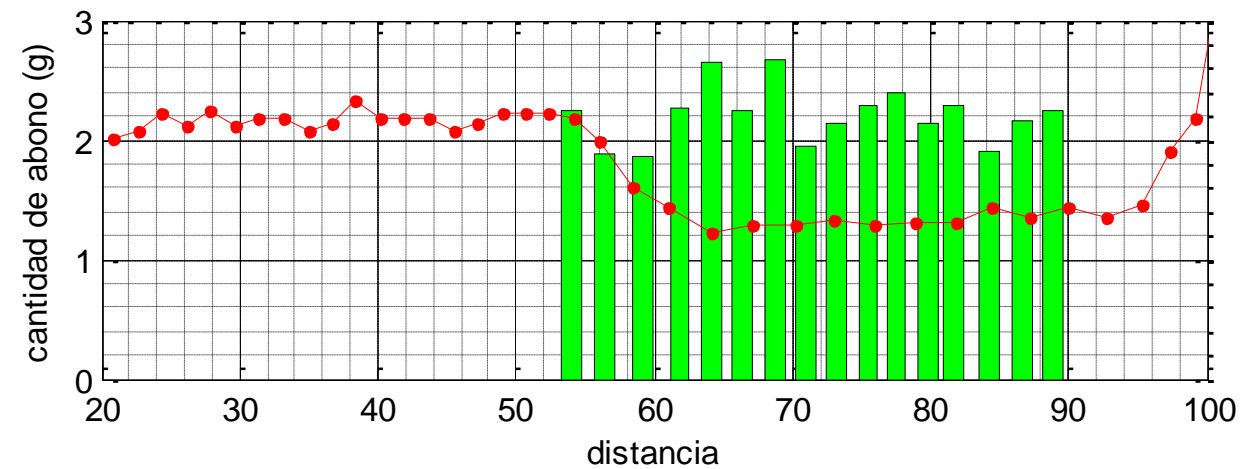
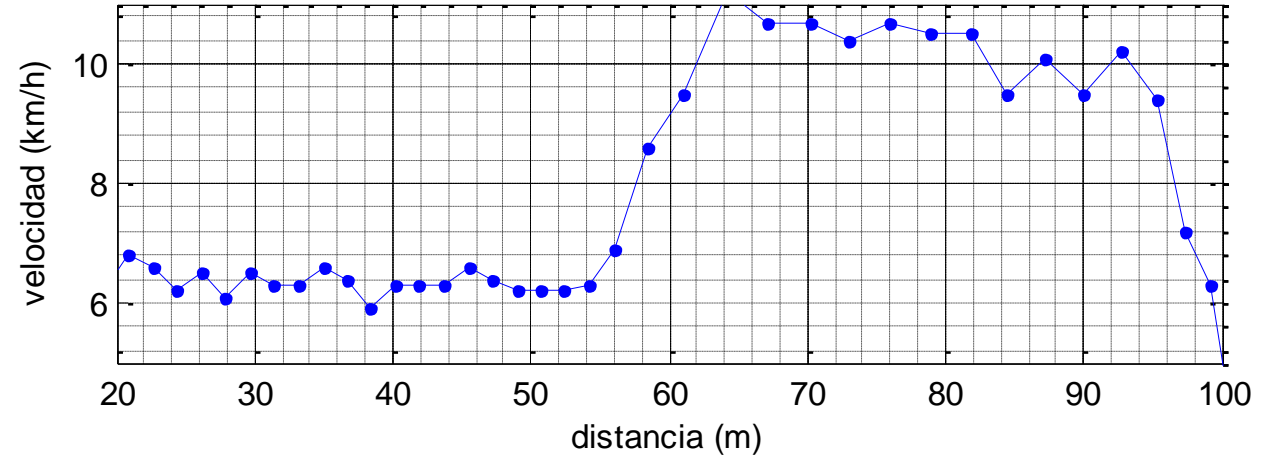
**PRUEBA DE CAMPO**

## Abonadora centrífuga DP44-4000 Aguirre, entre tradición y modernidad

Filix Barrón, Miguel Garrido, Wilma Da Costa Méto y Carlos García Fernández. UPMAGAZ, UPMAGAZ 15 de mayo

En esta prueba de campo realizada el 22 de octubre en Tafalla, pusimos a prueba las capacidades de regulación y control automático de la abonadora centrífuga de doble disco DP44-4000, suspendida y con una tolva de 4.000 l de capacidad, en un día con climatología adversa (velocidad del viento en el límite de lo admisible). Para ello trabajamos codo con codo con el dueño de la abonadora y explotación Pedro Ujue, y con el fabricante, Francisco Aranz (gerente de la empresa) y Mikel \*\*\*\* (técnico de campo), y establecimos un conjunto de ensayos que relatamos a continuación.

Las abonadoras centrífugas de Aguirre Agrícola se han desarrollado en estrecha colaboración con la estación de ensayo de Palencia. De acuerdo con su gerente, Francisco Aranz Jiménez, se han sido los retos en el desarrollo de una distribución transversal adecuada (figura 1), resolver mecánicamente el problema del ancho ajustable, e incorporar los sistemas de pesaje dinámico y control electrónico. El cuadro 1 recoge las características técnicas de los tres modelos disponibles. De los cuales se ha ensayado el DP44-4000 controlado por pesaje dinámico. De este modelo se venden cincuenta unidades al año en España, pero es en Francia donde se registra un mayor nivel de aceptación. La abonadora de doble disco Aguirre DP44 está preparada tanto para la fertilización como para la siembra de algunos cereales extensivos: trigo, cebada, avena, arroz, se sobrepone con posterior controlado con pase de rodillo. La máquina tiene un chasis que soporta los grupos de distribución (1 central y 2 laterales) todos girando a 1.000 rev/min, accionados por la toma de fuerza. En la parte inferior hay un cárter en el que está inserto el mecanismo de accionamiento del glugador de la tolva (juntas con los ejes

# Example: Dynamic Control System

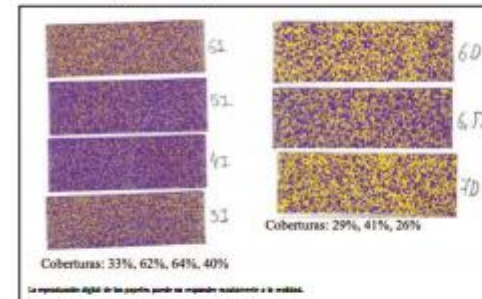
Gil Experiment, 2016



## PRUEBA DE CAMPO



FIGURA 3. Efecto del tipo de boquilla: normal (izda) y anti-deriva (derecha), ambas ISO azul (1,2 l/min a 3 bar).



en modo manual (caudal 1,2 l/min a 3 bar) de manera que la dosis real aplicada fue de 285 l/ha (la dosis inicialmente elegida fue 300 l/ha).

### Caudal proporcional a la velocidad de avance

El pulverizador Gil Teleno de 18 m dispone de un modo de funcionamiento manual (caudal constante) y otro automático (caudal proporcional a la velocidad real de avance). Para ello, dispone de un sistema de control que puede determinar la velocidad real (ground speed) bien mediante una antena GPS o bien mediante un sensor inductivo de rueda (opción más barata). En este ensayo empleamos la primera opción y verificamos posteriormente la calidad del segundo.

Cuando se trabaja a presión constante cualquier cambio de velocidad del vehicu-

desfavorables a la pulverización. Por ello, se sustituyeron tres boquillas del ramal derecho por boquillas anti-deriva (tipo venturi) del mismo caudal (ISO azul), realizándose todos los ensayos de campo con esta disposición.

### CUADRO I

		CAMBIO											
		tramo 1					tramo 2						
MANUAL	distancia (m)	6	9	12	15	18	21	24	27	30	33	36	
	velocidad (km/h)	7.30	7.30	7.00	5.90	5.00	4.70	4.50	5.00	5.00	5.00	4.70	
	cobertura (%)	57.6	45.9	65.2	81.9	93.8	90.4	69.8	95.2	94.9	100.0	93.2	
AUTOMATICO			tramo 1					tramo 2					
	distancia (m)	7	10	13	16	19	22	25	28	31	34	37	
	velocidad (km/h)	5.0	5.2	4.8	7.1	7.0	6.8	7.2	7.0	8.1	6.6	6.1	
		cobertura (%)	46.8	53.3	52.6	71.6	59.0	71.5	44.2	35.7	78.3	47.5	46.2



# WSN as a SCADA System

Wireless System Networks as Data Acquisition Systems



**SPECTRAL REFLECTANCE SENSOR (SRS)**  
Measurements: Leaf Area Index (LAI), Solar Radiation  
Applications: Biomass Production, Canopy Cover, Greenhouse Management, Plant Ecology, Plant Water Stress

OVERVIEW SPECIFICATIONS ACCESSORIES SUPPORT EDUCATION



2:13



# Complete SCADA Systems

JDLINK Fleet Management



The screenshot shows the JDLink web dashboard in a Windows Internet Explorer browser. The dashboard is titled "JDLink" and "JOHN DEERE". It displays various performance metrics for a selected machine (8335R).

**Utilización de máquina**

	Tiempo a ralentí	Tiempo de trabajo
Utilización de ...	30.7 hr	33.6 hr
Consumo de c...	157.6 l	810.1 l

**Horas de motor en período**

**Combustible consumido en período**

H	Combustible consumido en un ...
12.0	2.2 l
14.0	33 l
16.0	46.6 l
18.0	19.7 l
20.0	39.8 l
22.0	34.6 l

**Uso según estado de la máquina**

Fuel	Ralentí	Trabajando	Transporte
Combustib...	157.6 l	615.2 l	194.6 l
Utilización...	30.7 hr	28.1 hr	5.6 hr
Régimen d...	5.1 l/hr	22.0 l/hr	34.2 l/hr
Velocidad ...	0.0 km/hr	5.1 km/hr	34.5 km/hr
Velocidad ...	968.1 RPM	1440.4 RPM	1901.8 RPM
Factor de ...	14.5 prcnt	38.5 prcnt	52.5 prcnt

**Régimen de combustible promedio**

Promedio ponderado	l/hr
Consumo medio de combust	8.7

**Temperatura**

**Tiempo de marcha**

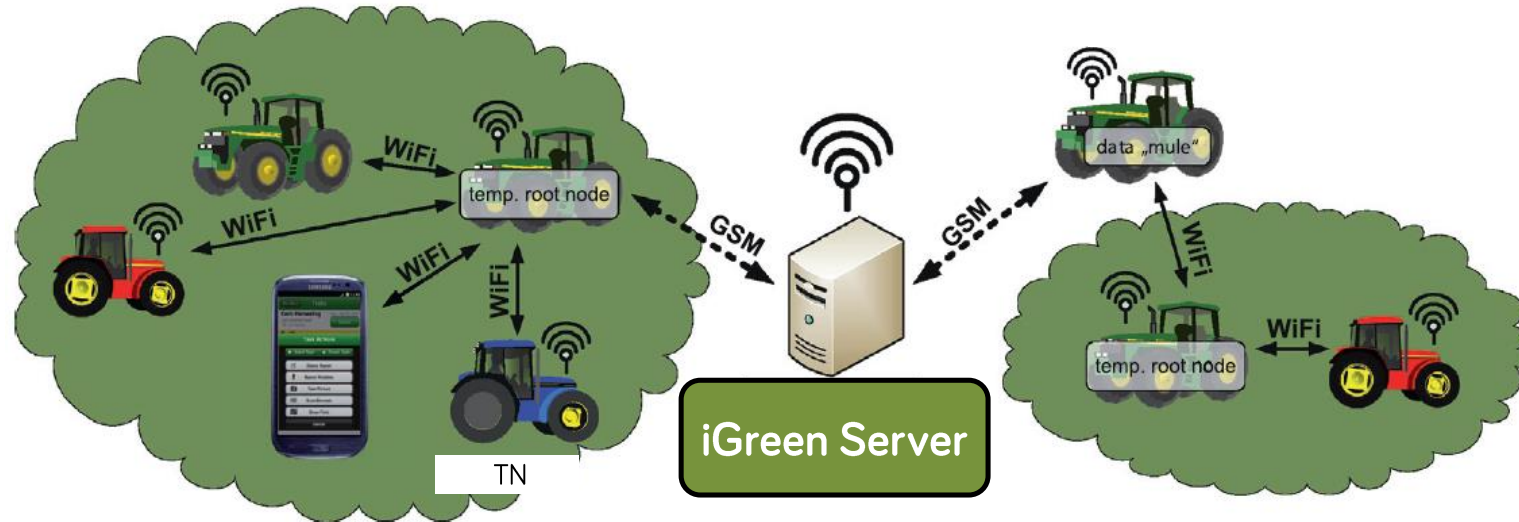
	Inactivo	Activo	Error	No disponible
Tiempo de marcha				

**Voltaje promedio de sistema**

Promedio ponderado	V
Voltaje promedio de sistema	

# Complete SCADA Systems

iGreen Conceptual Fleet Management



TN – Temporal Node  
Mule – Data Mule

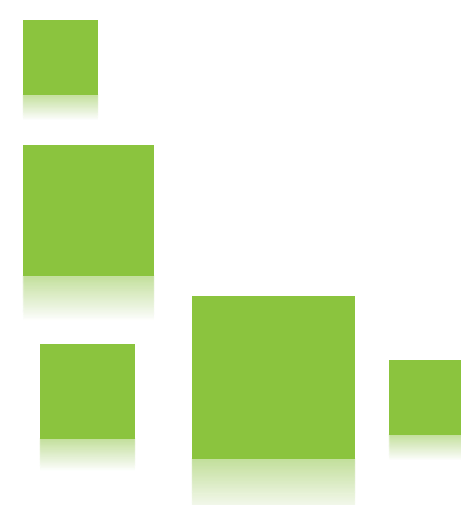
- The real-time information transmitted via
  - WiFi (IEEE 802.11g) has a delay of 0.1 to 3s and a success rate of >90%;
  - 3G has a delay of 1 to 10s and a success rate of >65%;
- When documentation information is transmitted (real time is not required) the delay is 5 to 85 minutes and the success rate is 100%.

(Blank et al., 2013)

# Advancements in the Digitalization of Information



Worldwide estimation	Year		
	2000	2007	2013
Analogical	75%	7%	2%
Digital	25%	93%	98%





- The data is starting to be included in company's balance sheets as corporate assets;
- In 1980, 40% of company assets were intangible;
- In 2000, this value is considered to have increased to 75% (e.g. brand, business strategy, talent, and data).



# Imperative Measure and Instrumental Blindness (Distortion)

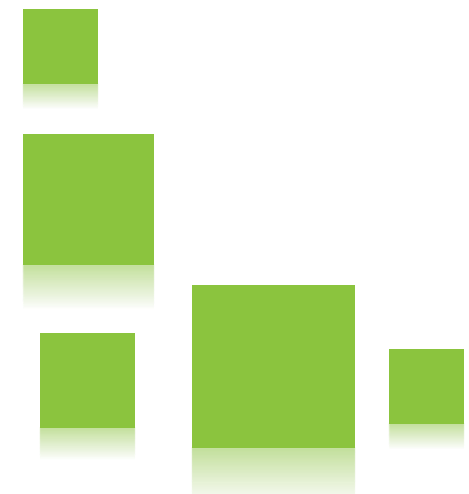


- Does it really measure the value of interest (the most relevant aspect) or simply what is easiest to measure?
- Not everything measured is always relevant, nor can everything that is relevant be measured...





# Geographic Information Systems

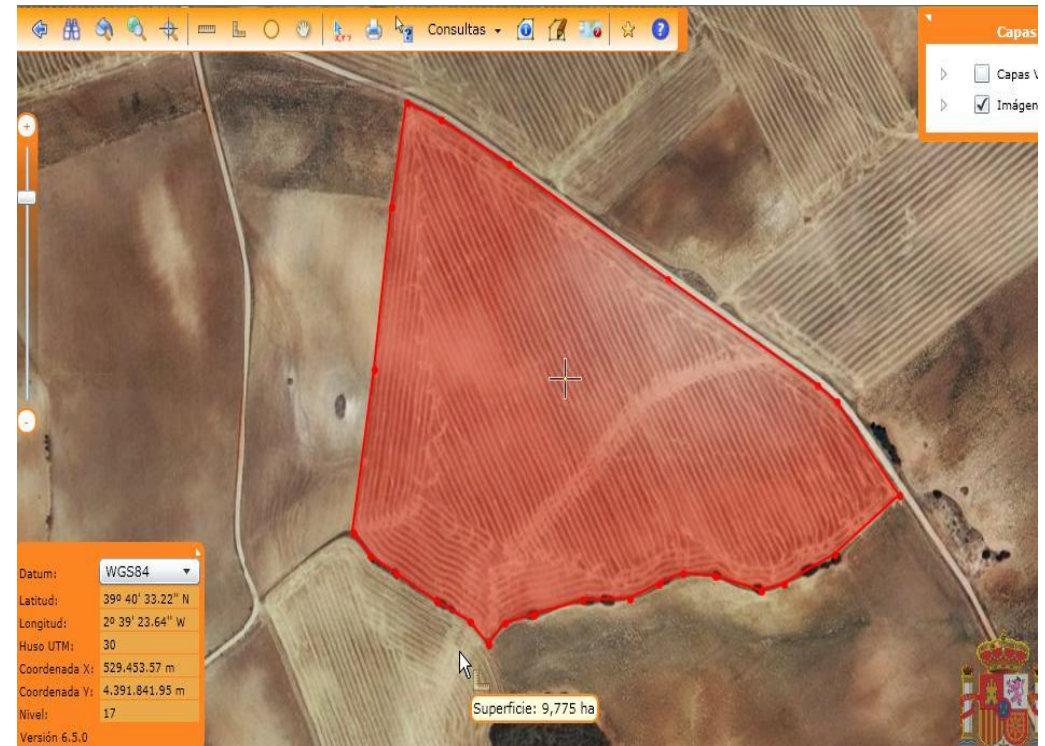
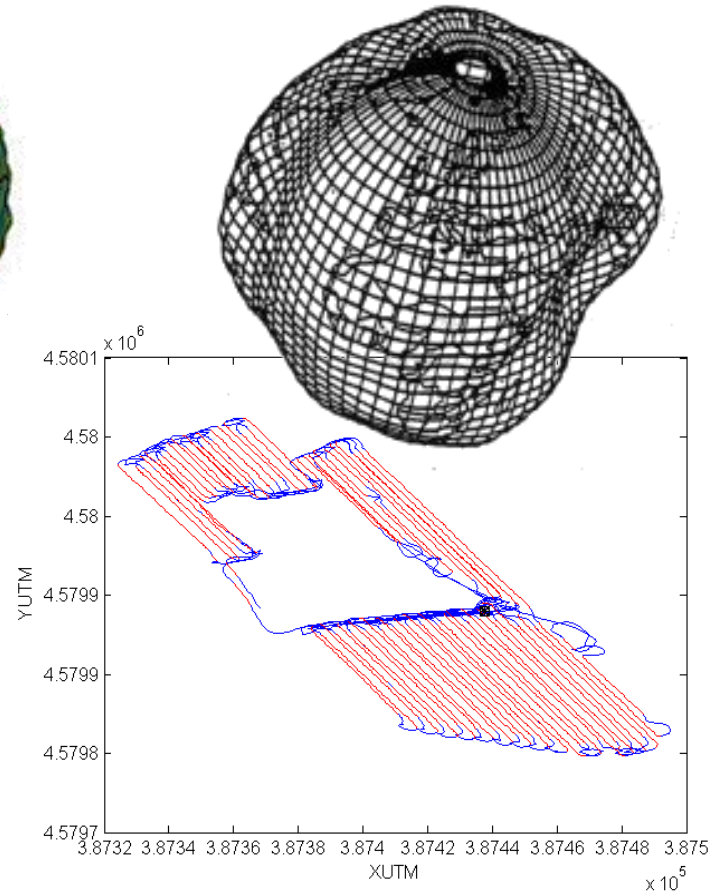
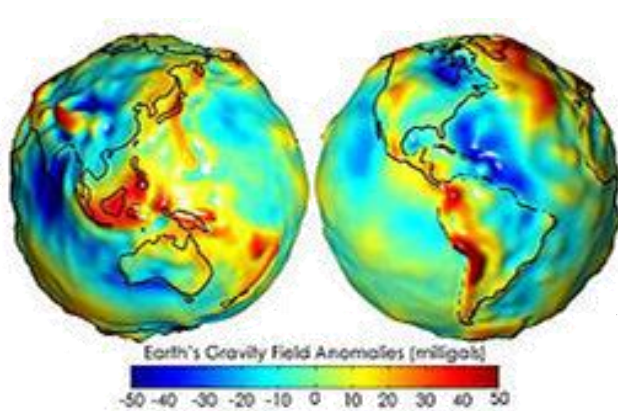


# Digital Mapping GISs

(e.g. Google maps, SIGPAC PNOA, etc.)

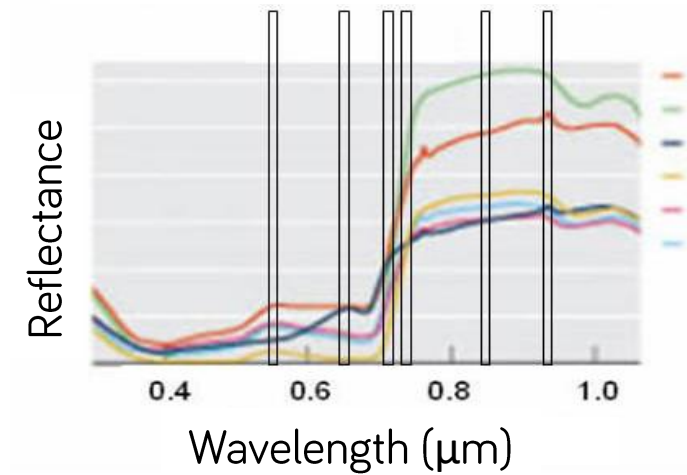
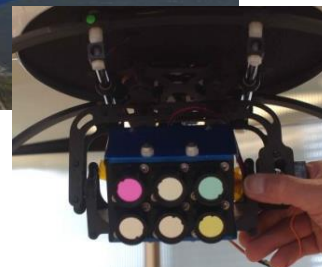
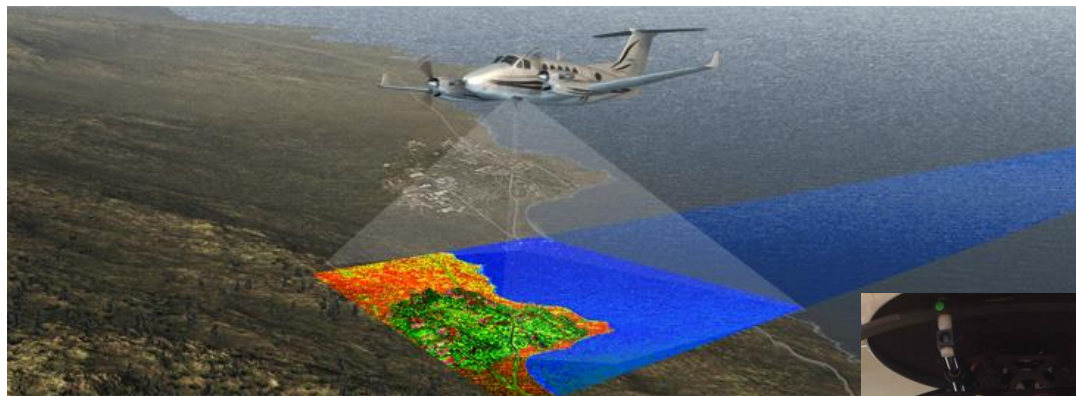


## Geoide

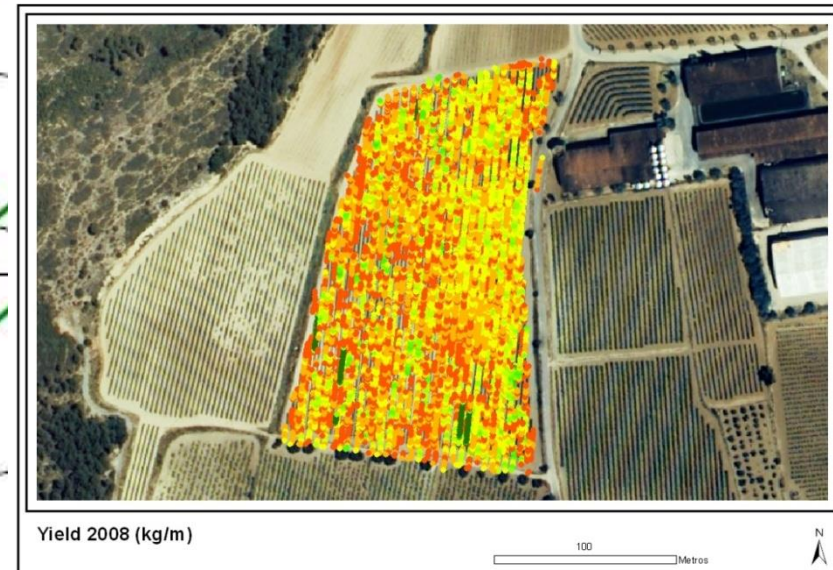
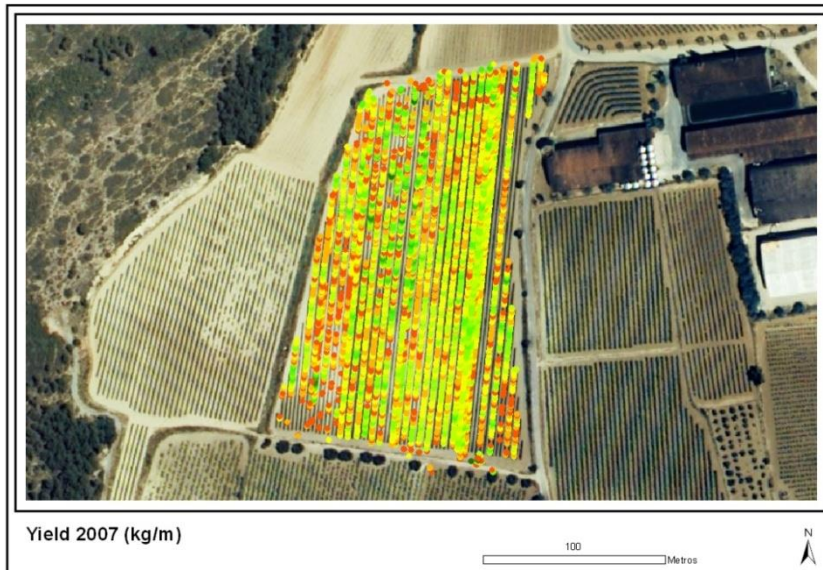
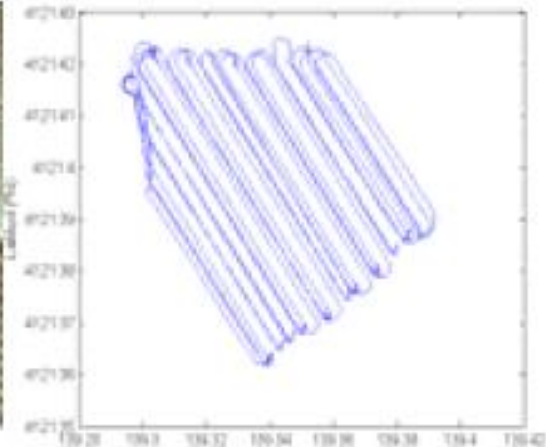
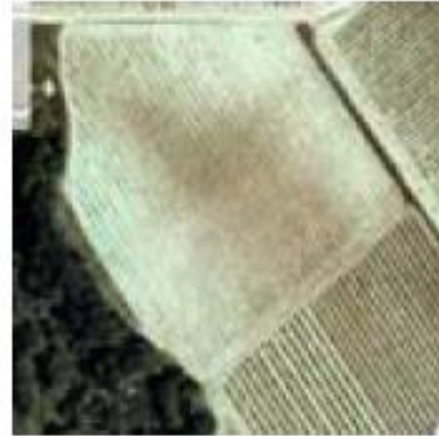




# Aerial Mapping with Remote Technologies

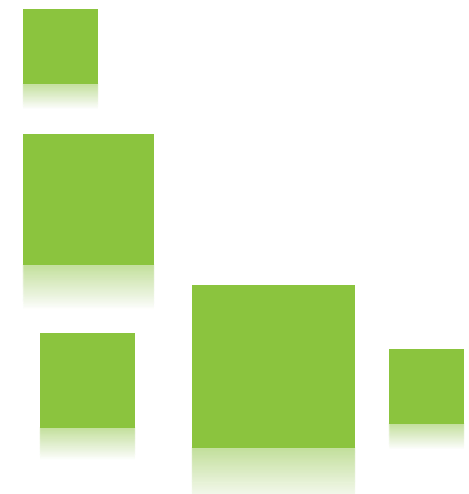


# On-Board Technology Based Mapping





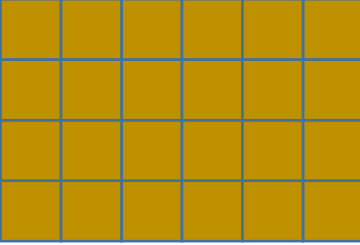
# Data aggregation



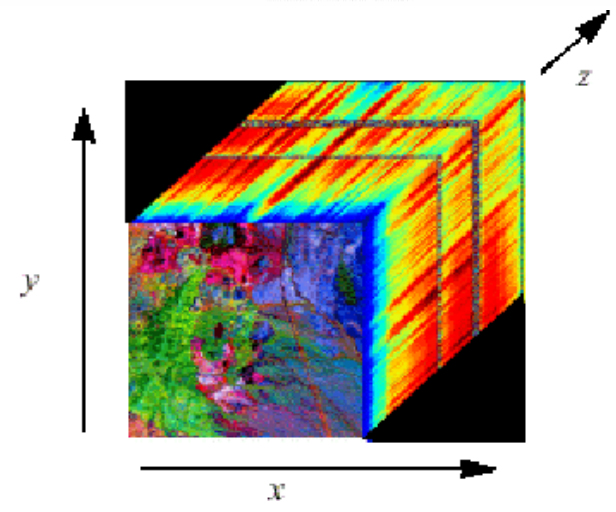
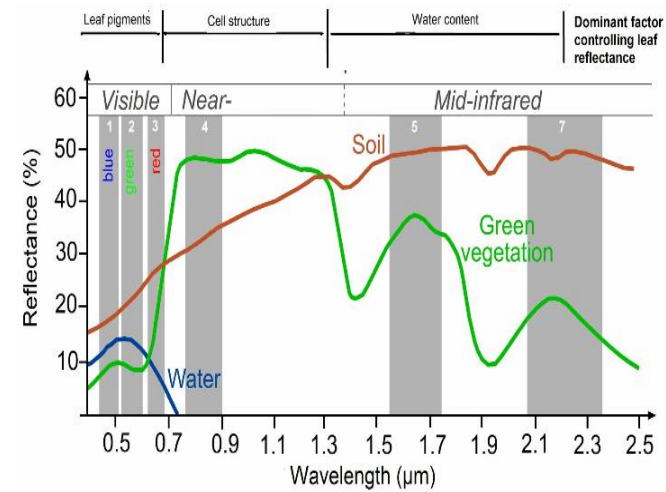
# Dimensions of the Data

- Scalar value: e.g.  $T^a$  

- 1D 

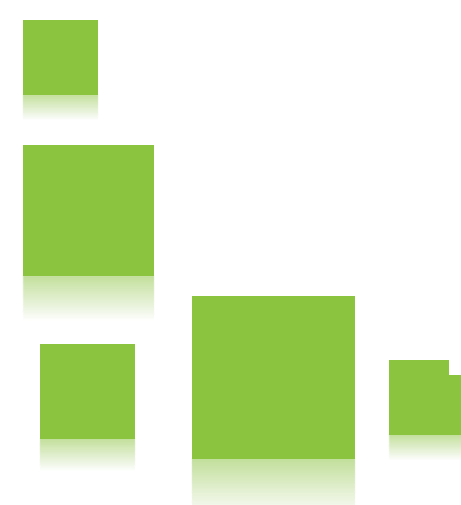
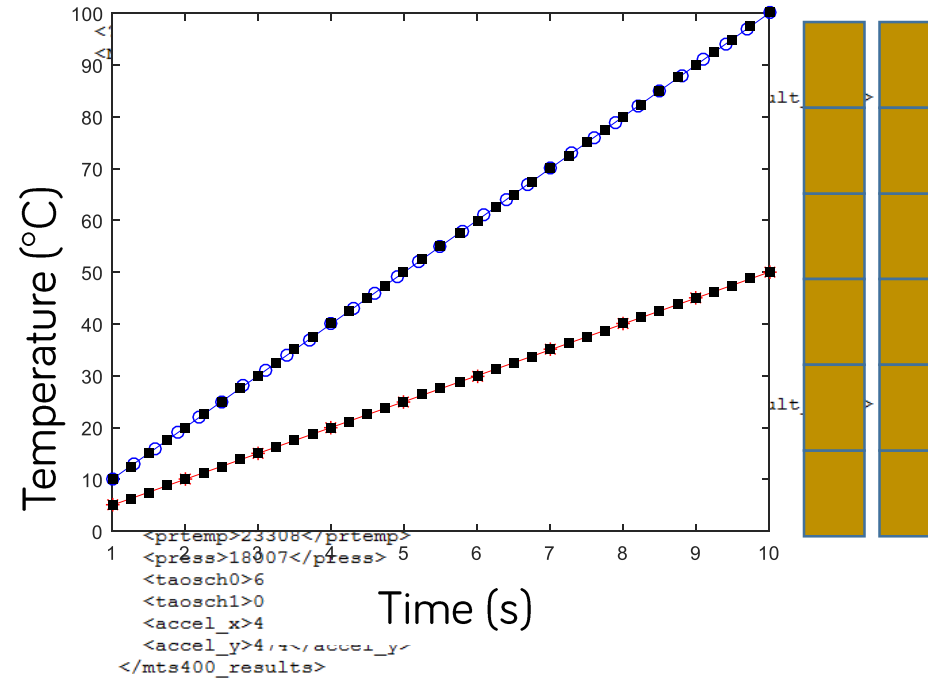
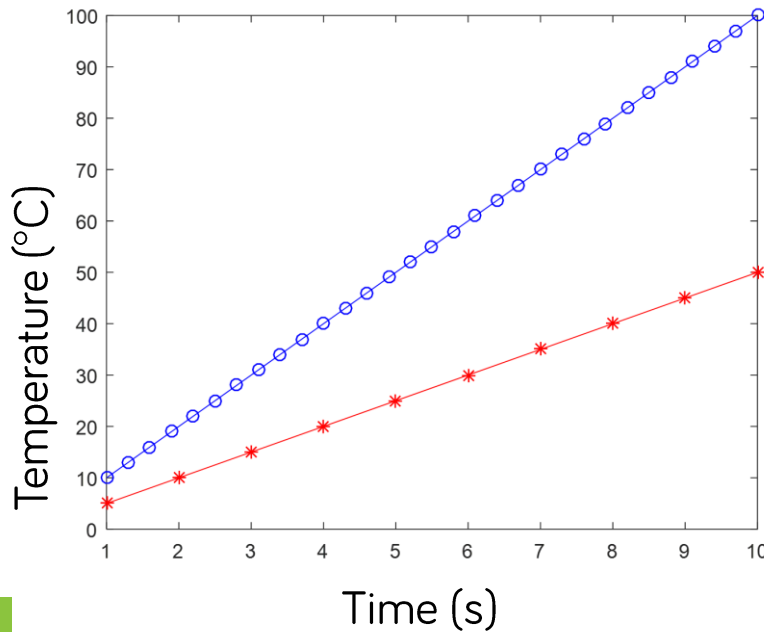
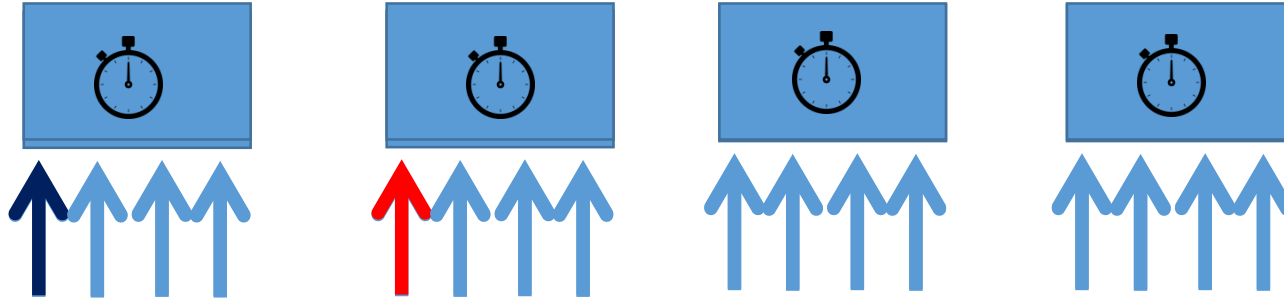
- 2D 

- 3D 



# Temporary Data Aggregation

The importance of consistent time units

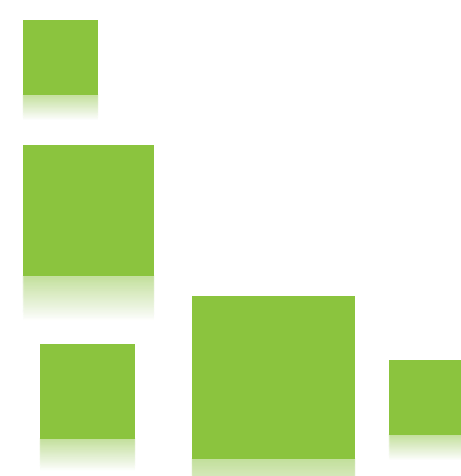
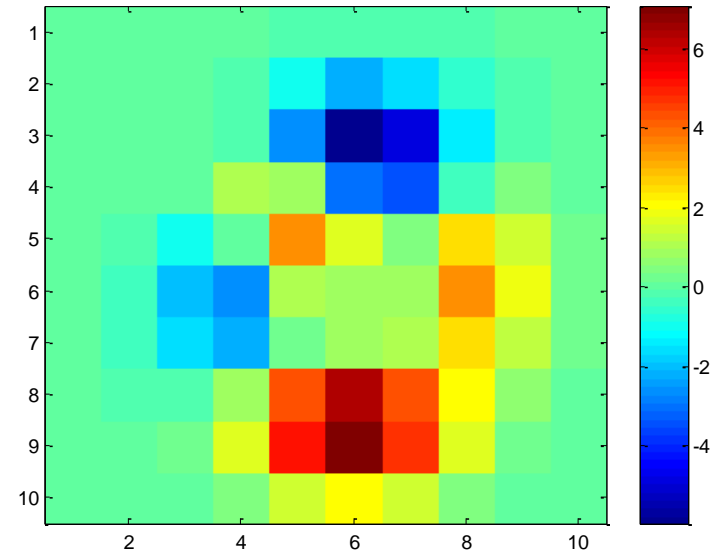


# Spatial Data Aggregation

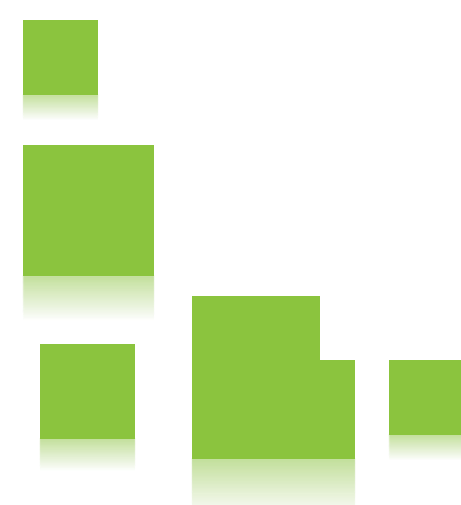
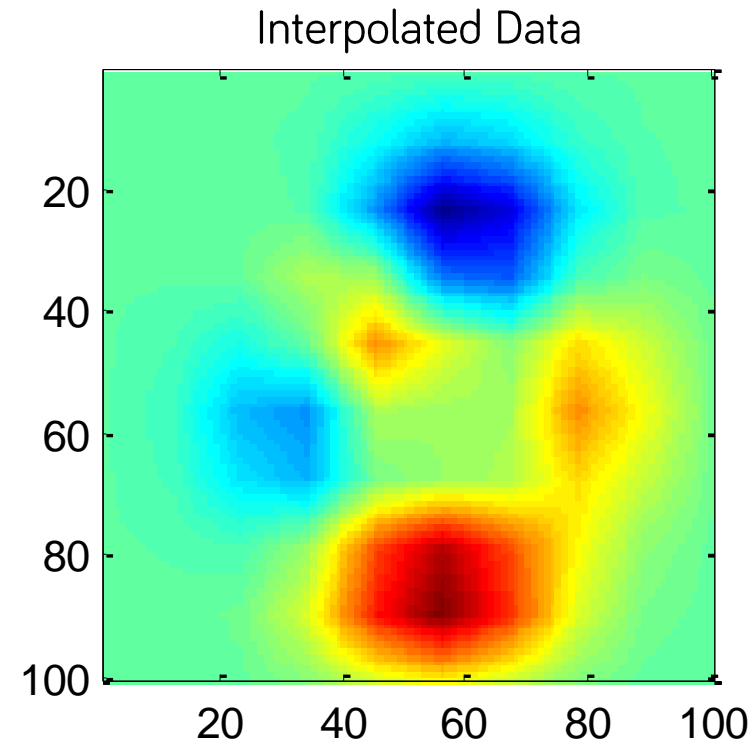
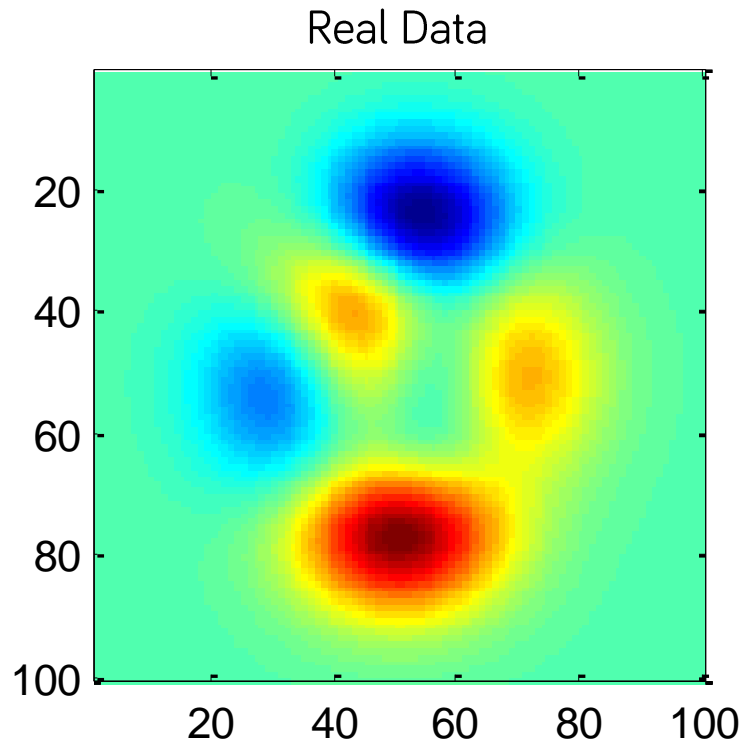
Spatial Interpolation Methods



Example of varieties detected through rough sampling



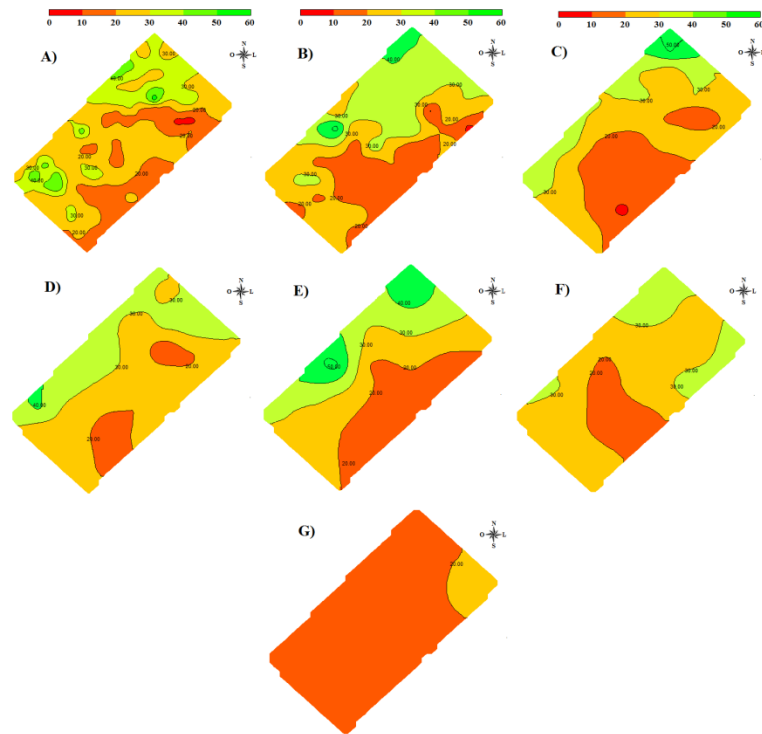
# Spatial Interpolation Results



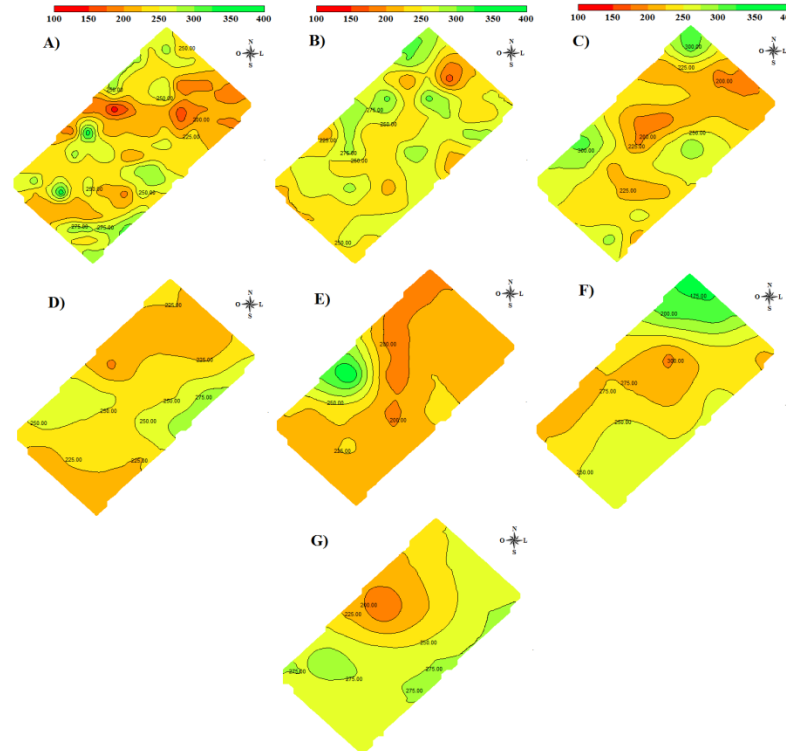
# Not All Ions (Variables) Require the Same Sampling Density



Phosphorus



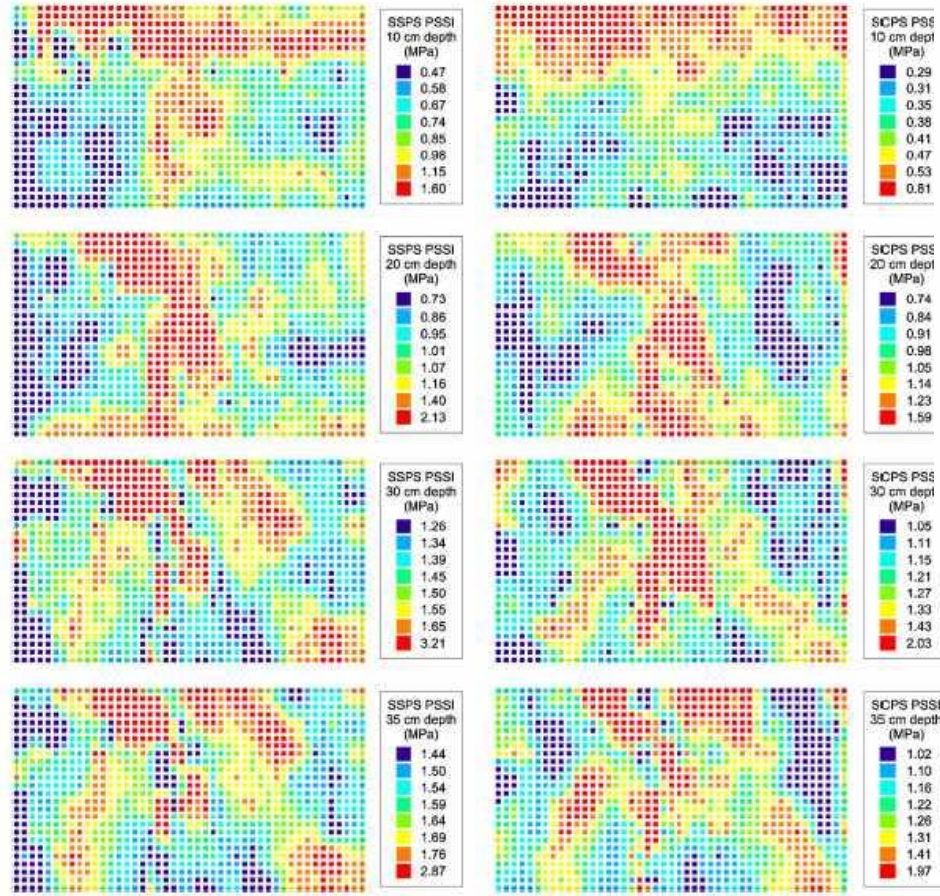
Potassium



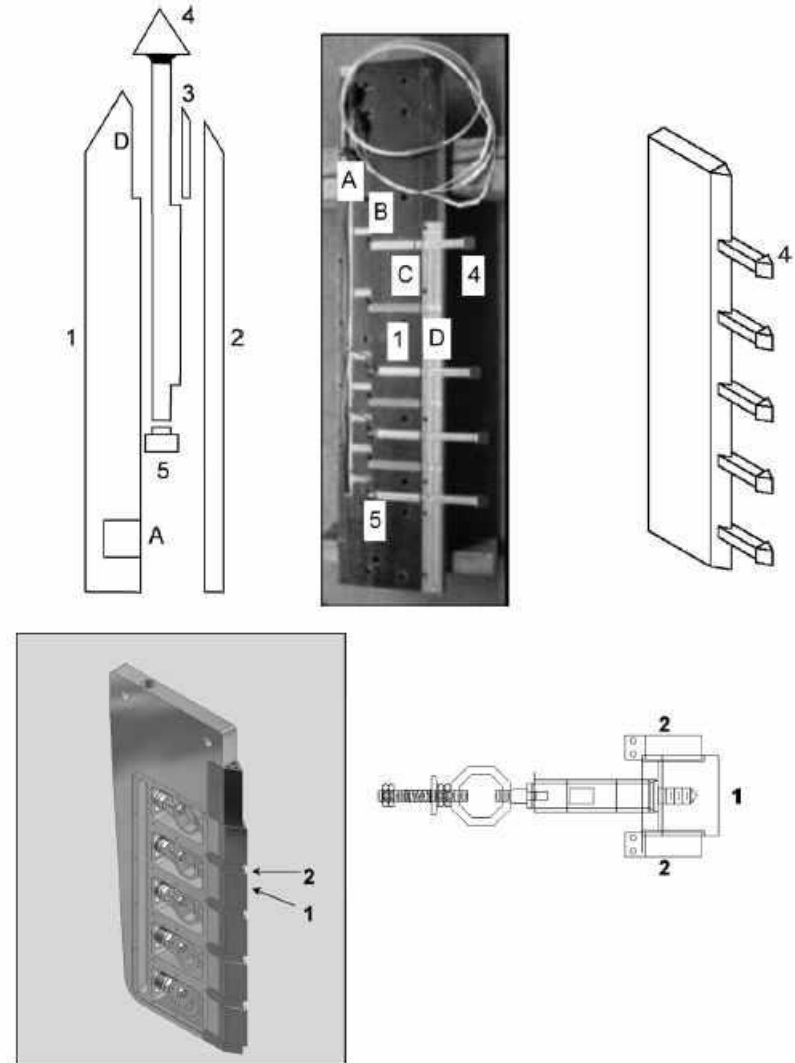


# Information Robustness

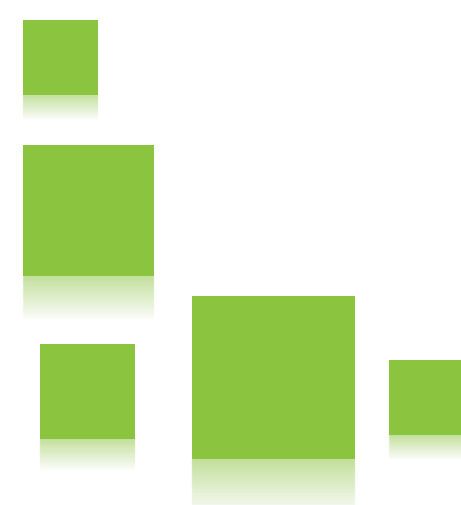
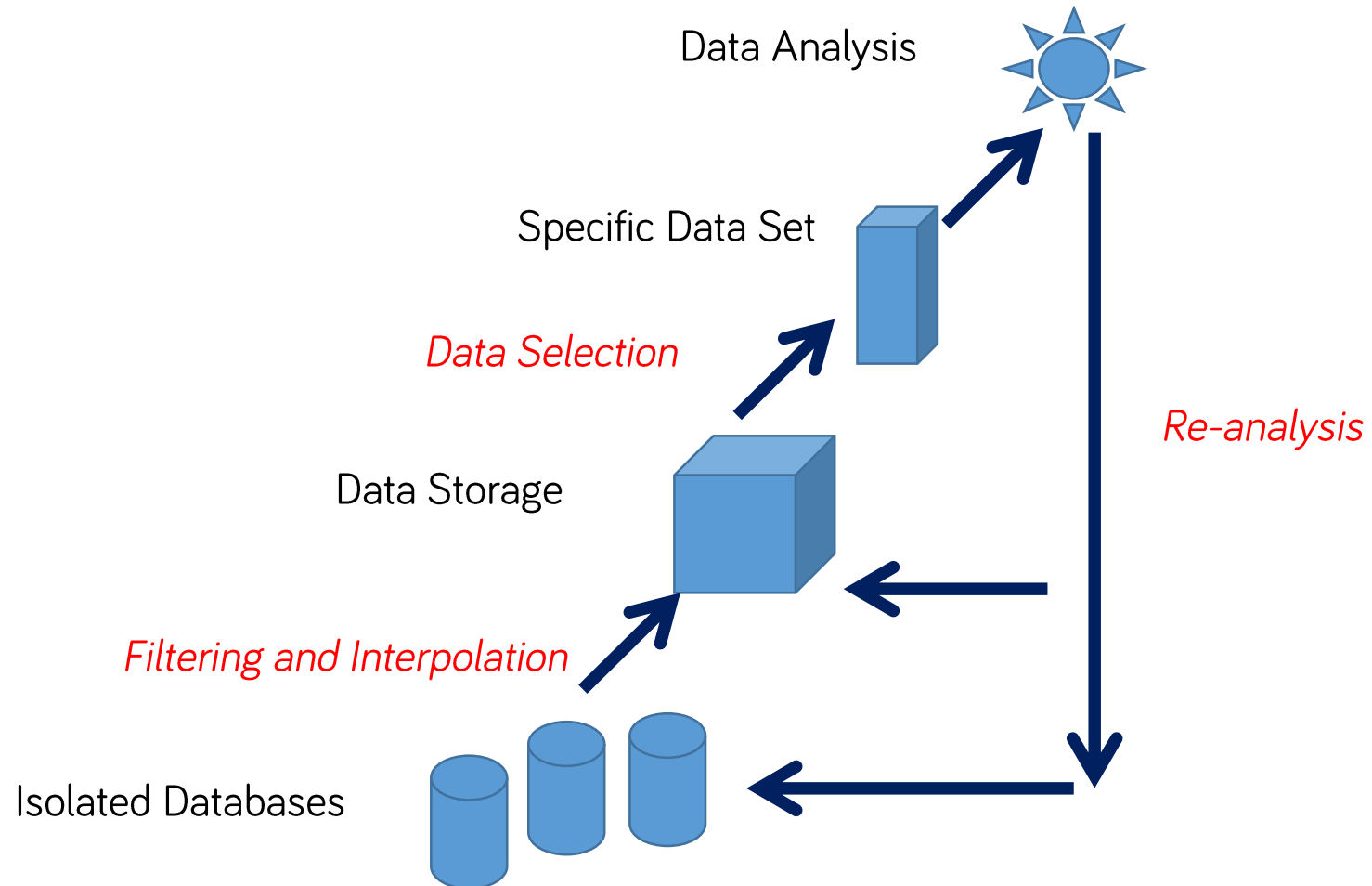
Electromechanical Sensors



UC Davis vs. USDA

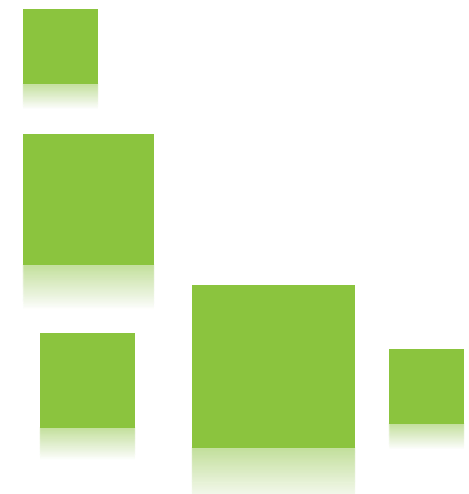


# Summary of Data Aggregation





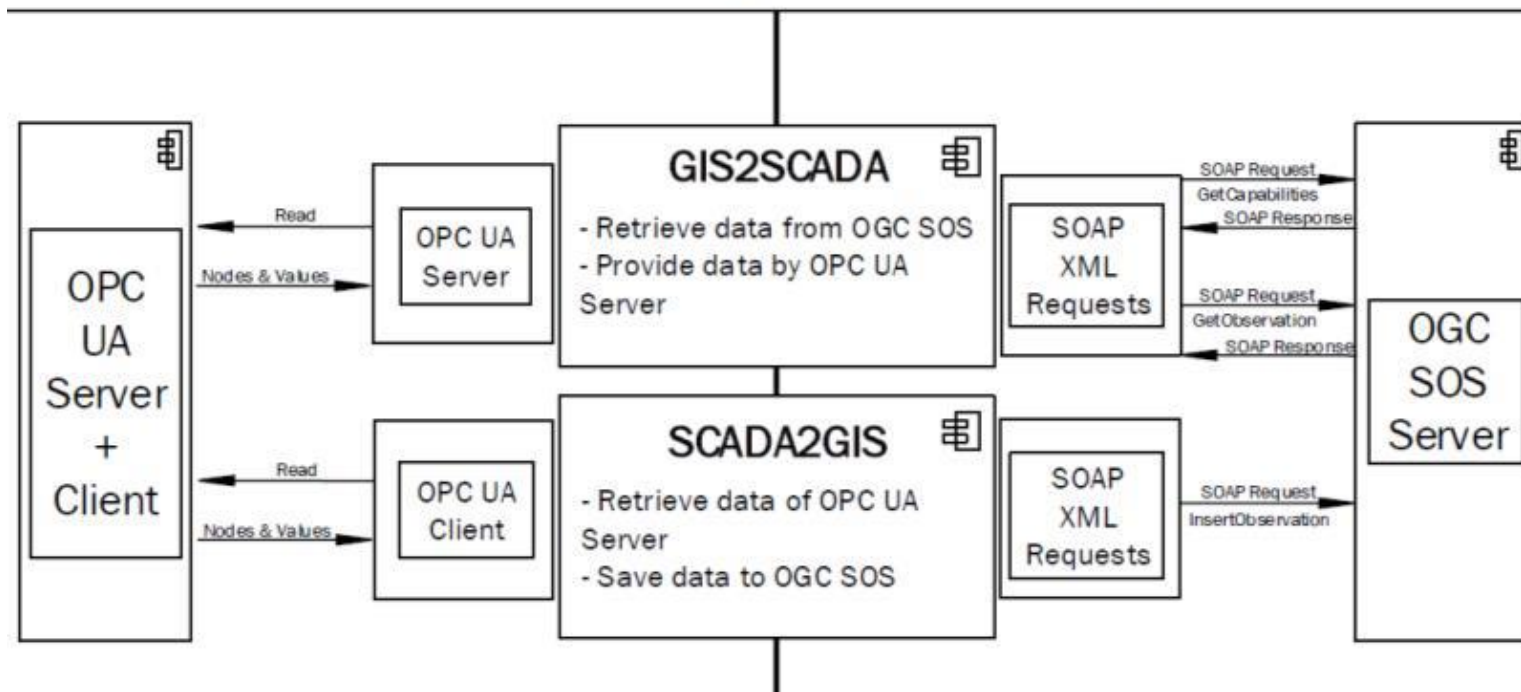
# The Interoperability Concept



# The Data Gateways

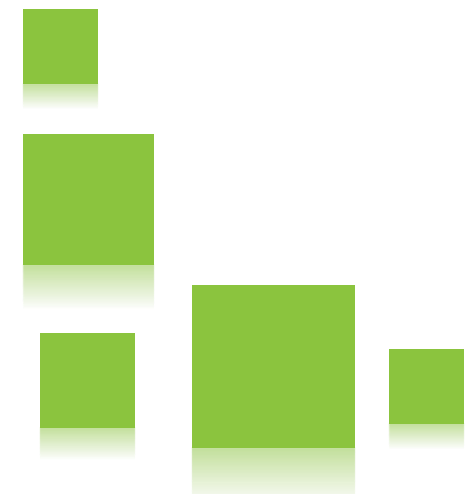


## Software Adapter Conception



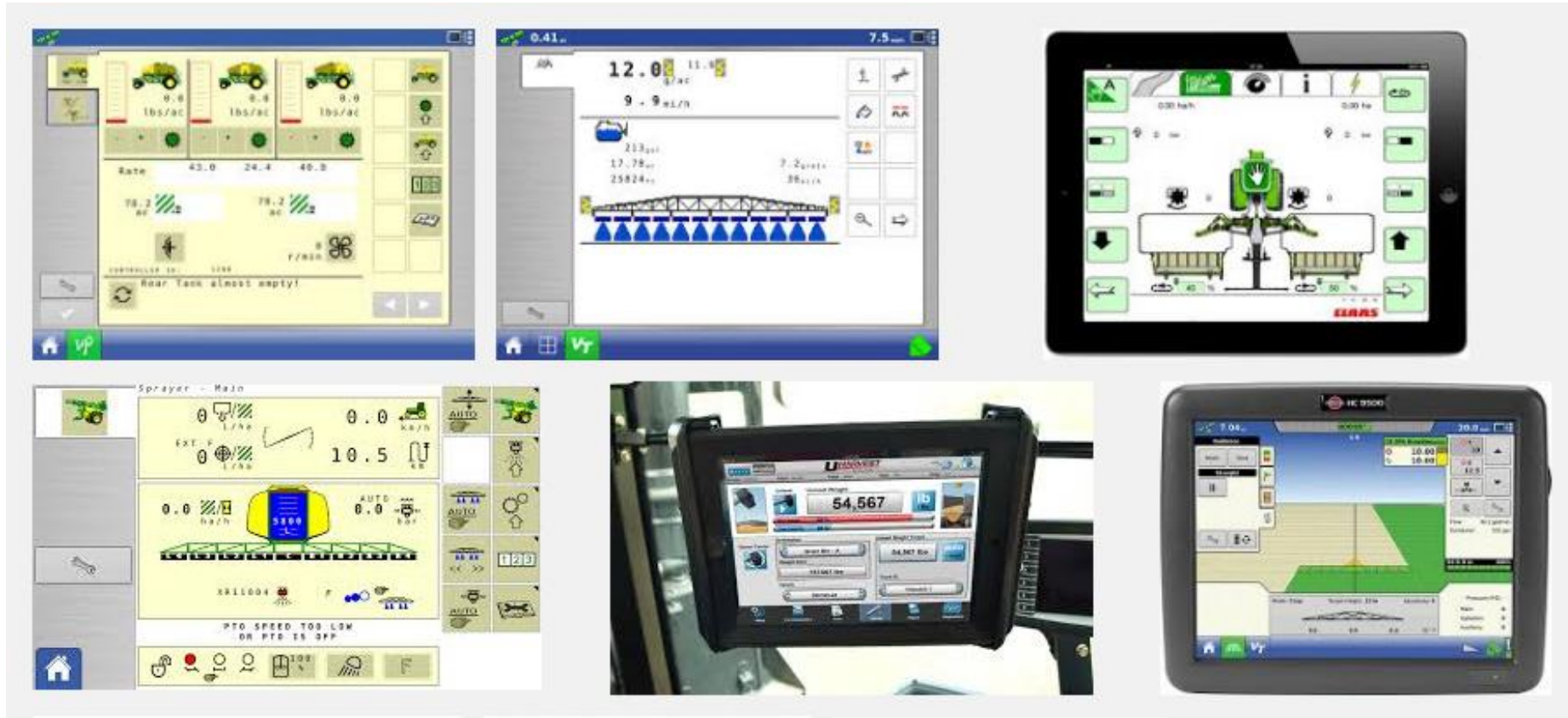


# Data Visualisation

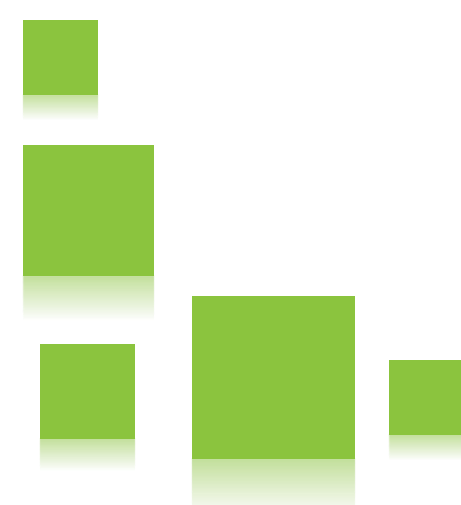
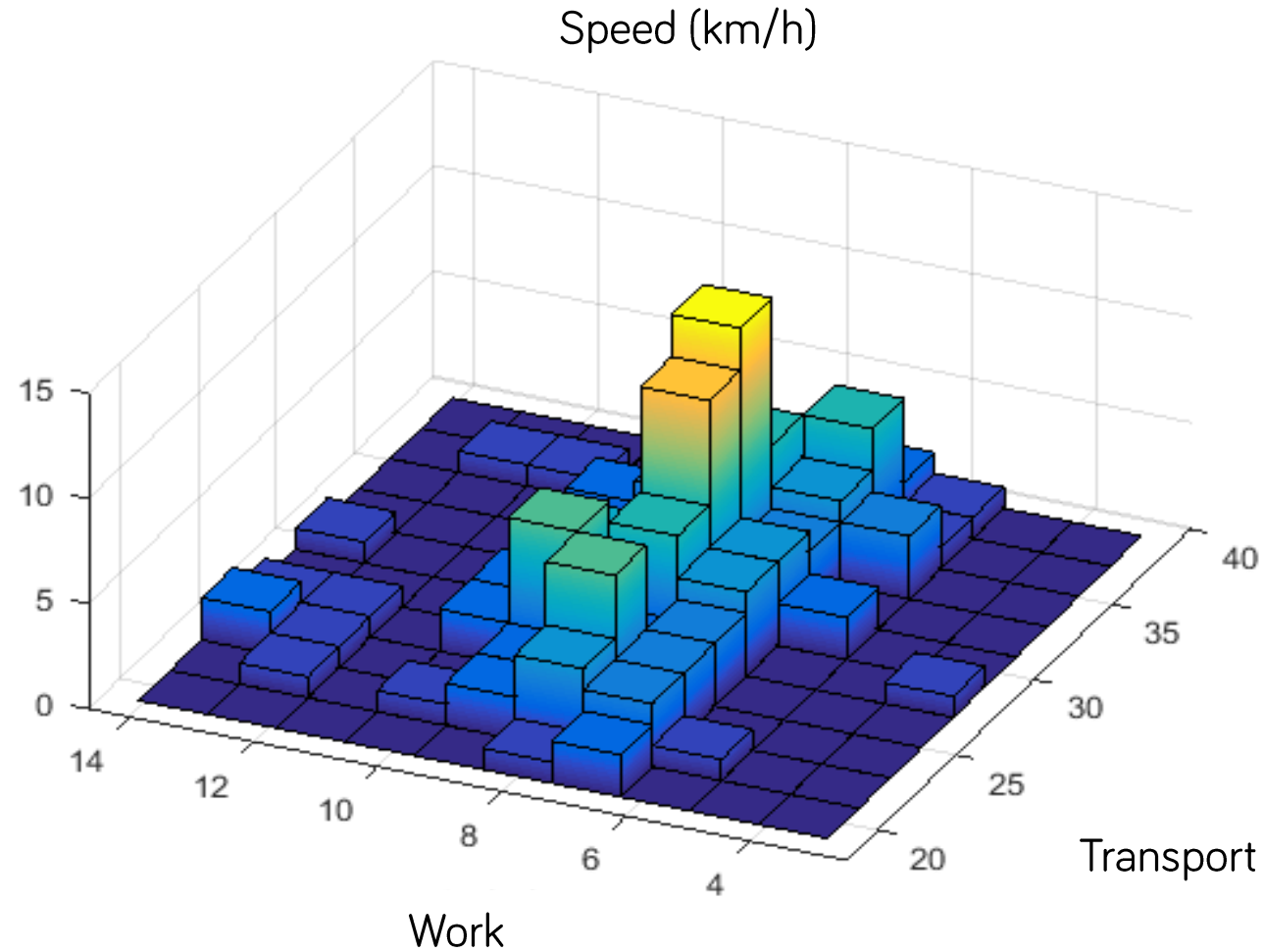


# Visualisation

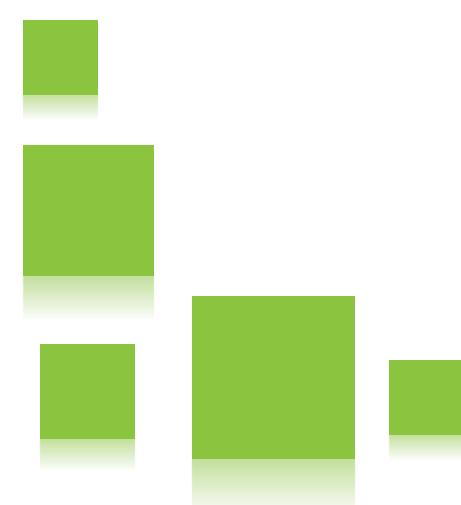
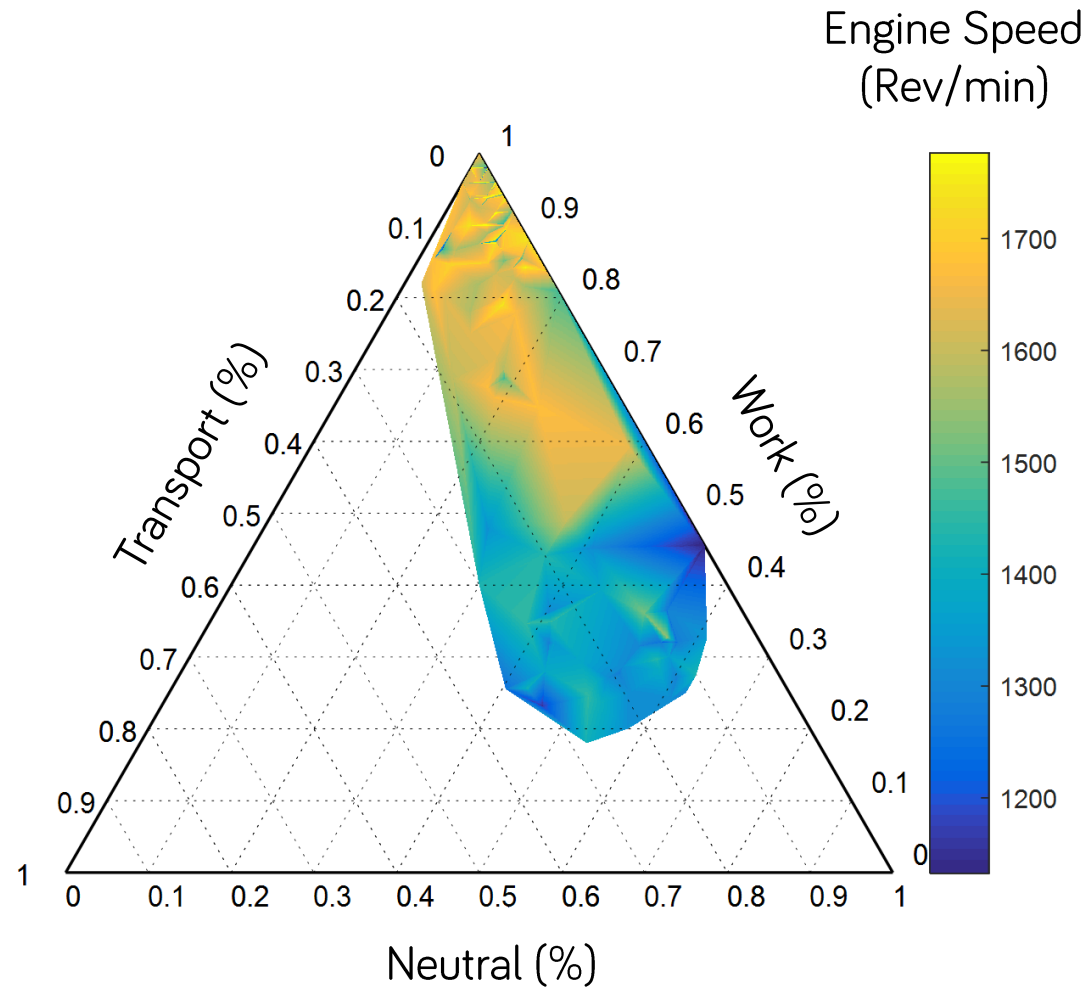
Translating a Mental Image to Something on the Screen



# The Scientific Approach

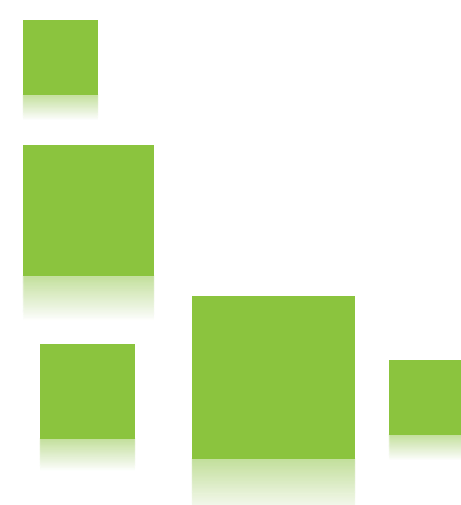
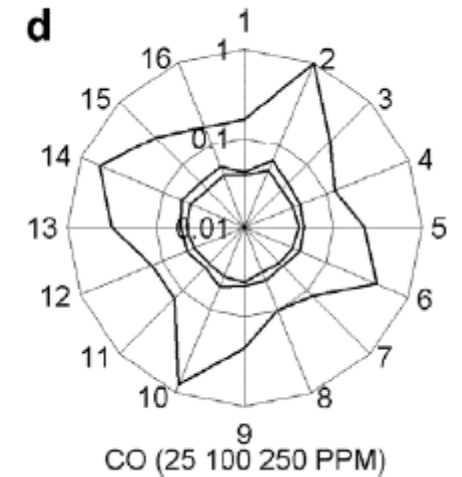
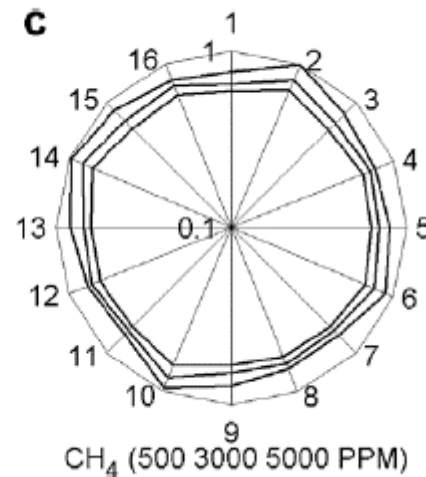
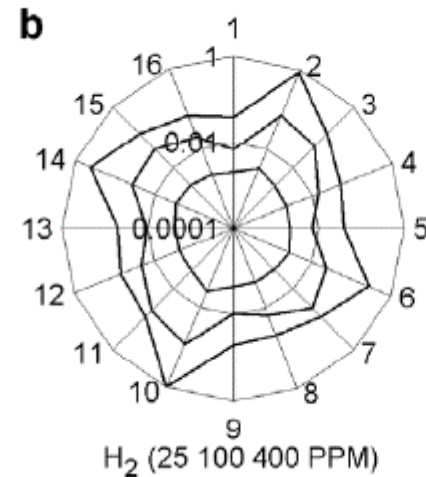
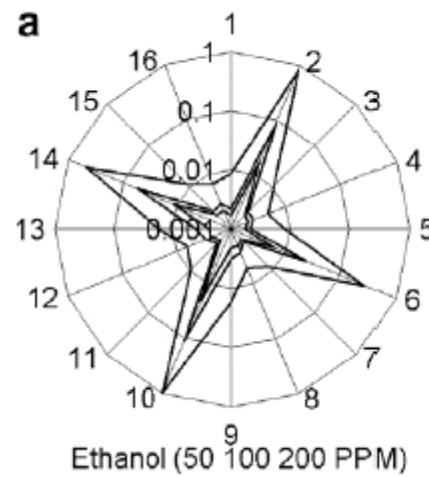
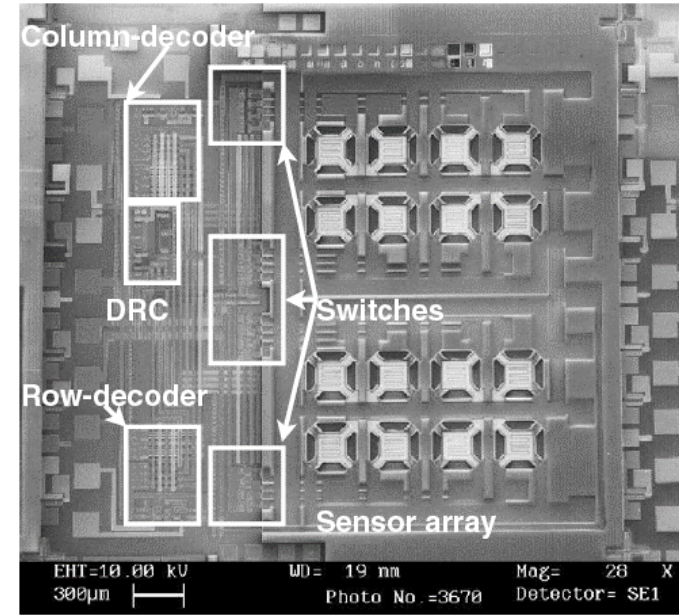


# Visualising Multiple Parameters





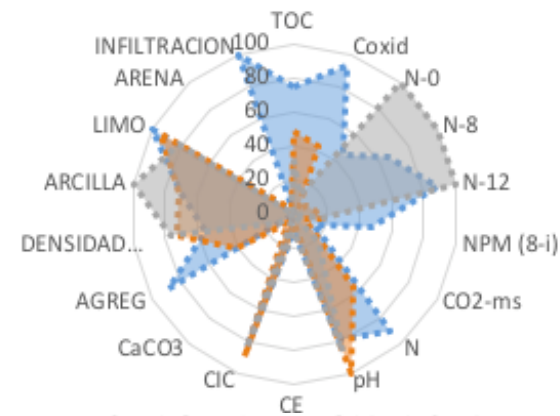
# Visualising Behaviour Patterns



# Example: Soil Characteristics

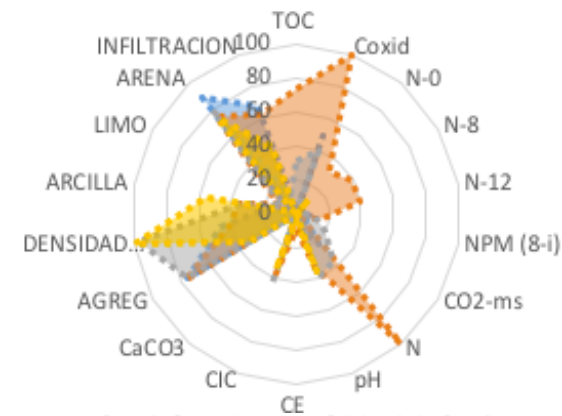
Soil Origin 5

5.2 5.3 5.5



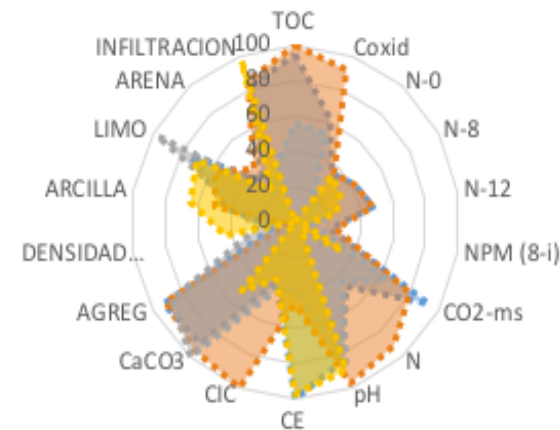
Soil Origin 3

3.1 3.2 3.3 3.5



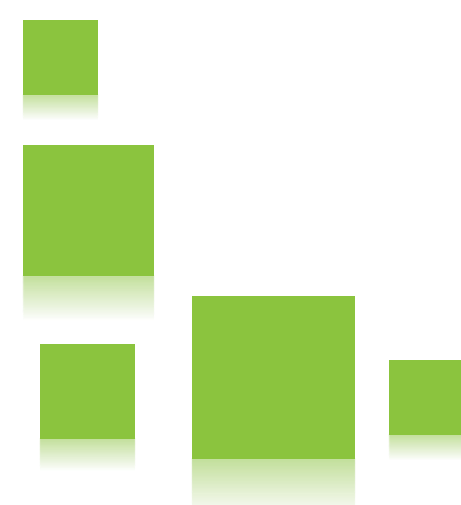
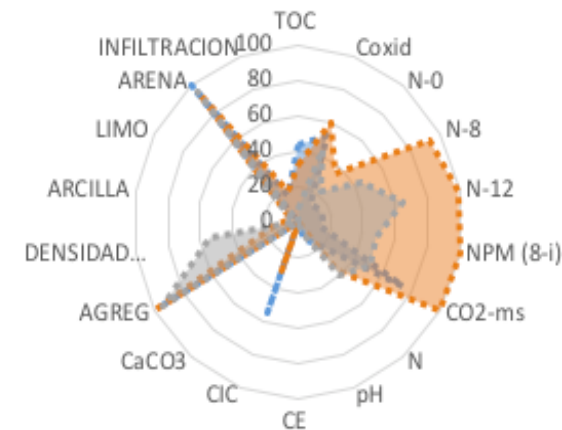
Soil Origin 4

4.1 4.2 4.3 4.5



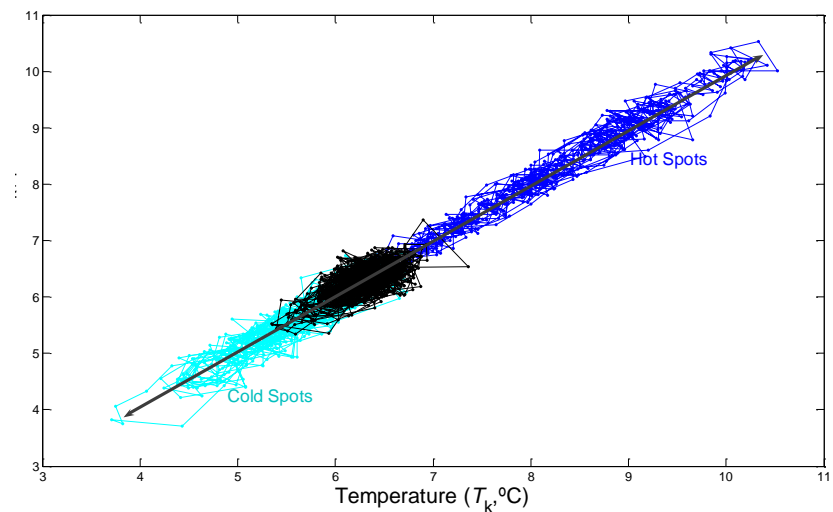
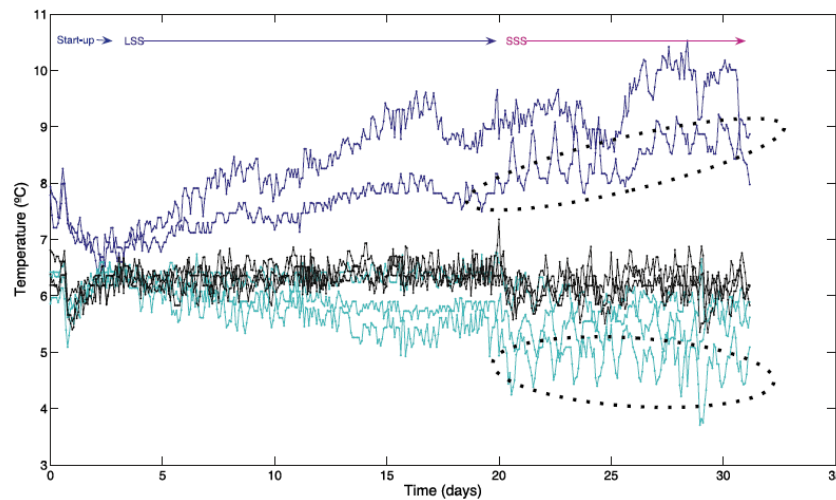
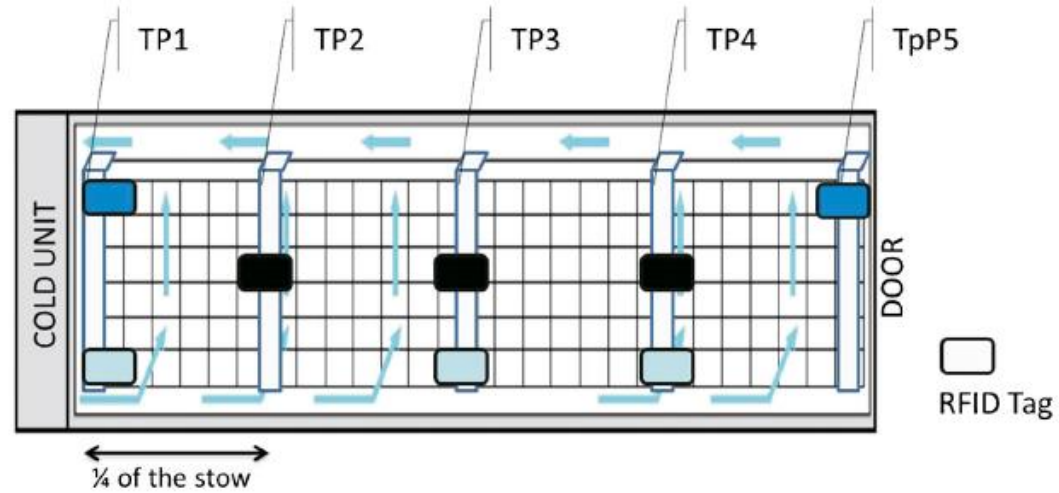
Soil Origin 1

1.1 1.2 1.3



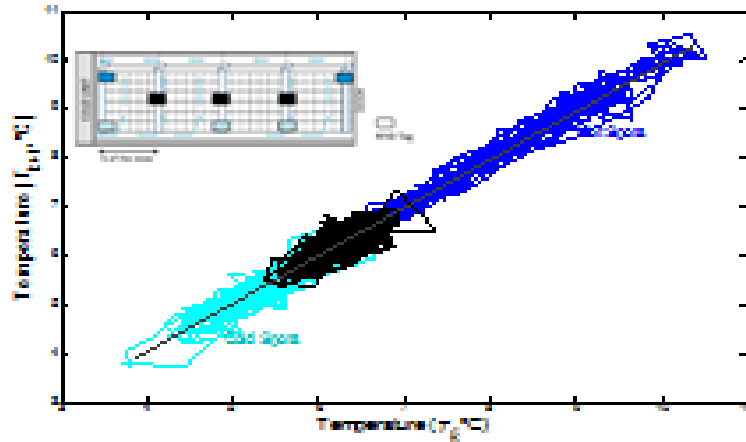
# Phase Diagrams

To Simplify Temporal Information

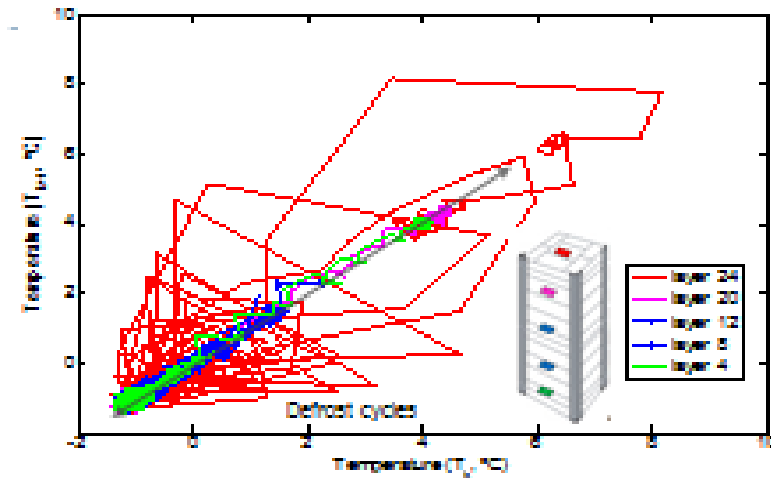
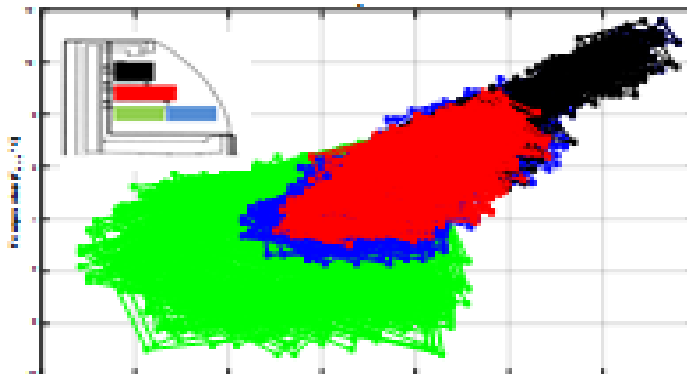


# Phase Diagrams

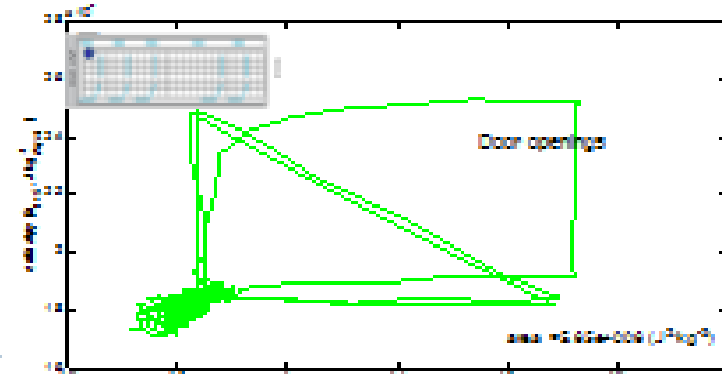
For Event Identification



• Identification of hot and cold spots



• Identification of events



# Management Zones

To Simplify Spatial Information

Fig. 1 EM38 map of soil variability for the Nickels site along with soil-sampling sites for ground truthing designated by the circles

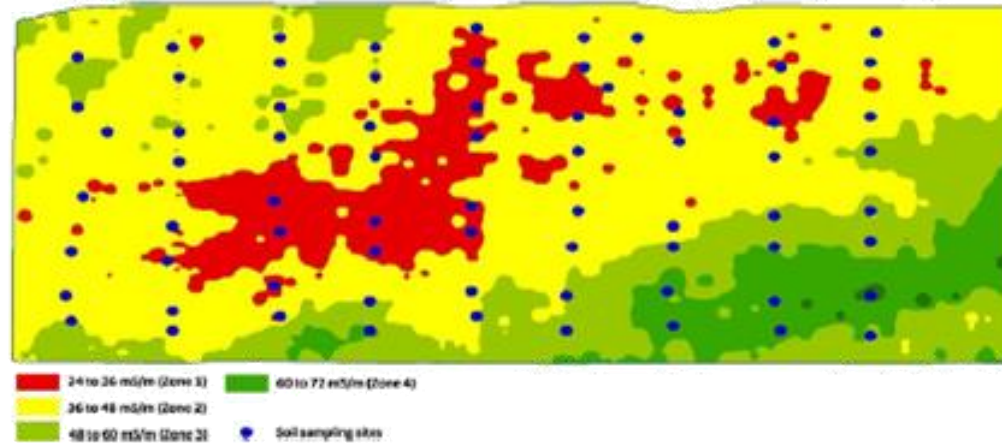
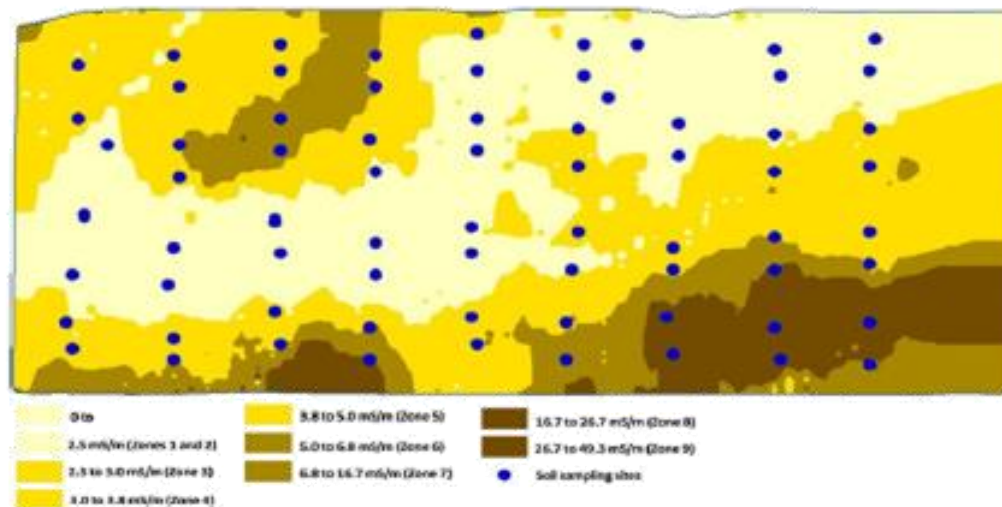
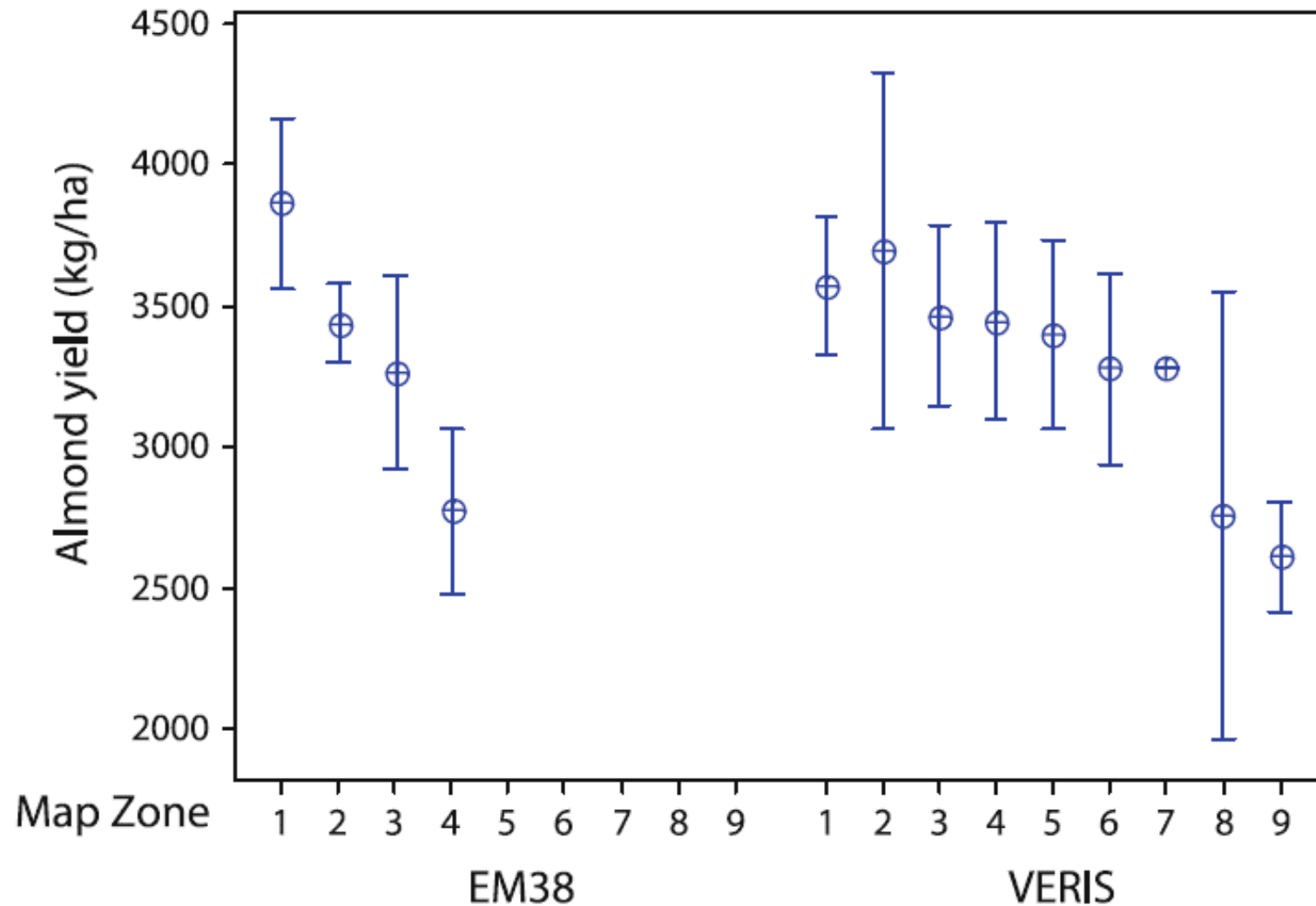


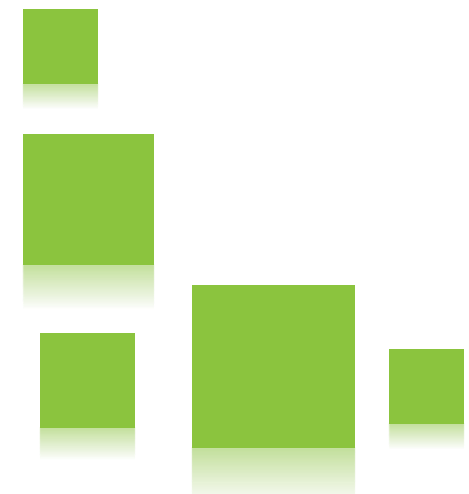
Fig. 2 VERIS map of soil variability for the Nickels site along with soil-sampling sites for ground truthing designated by the circles



# The Importance of Defining Management Zones



# Multivariate Analysis



# Common Multivariate Procedures for Pattern Recognition



PCA

Non  
Supervised

Dimension  
Reduction

Cluster

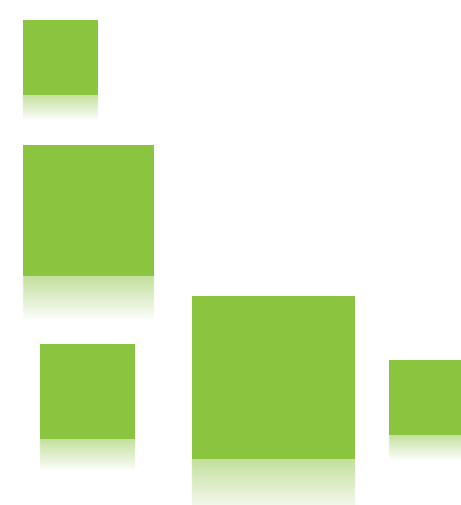
Non  
Supervised

Data Grouping  
(Pattern  
Recognition)

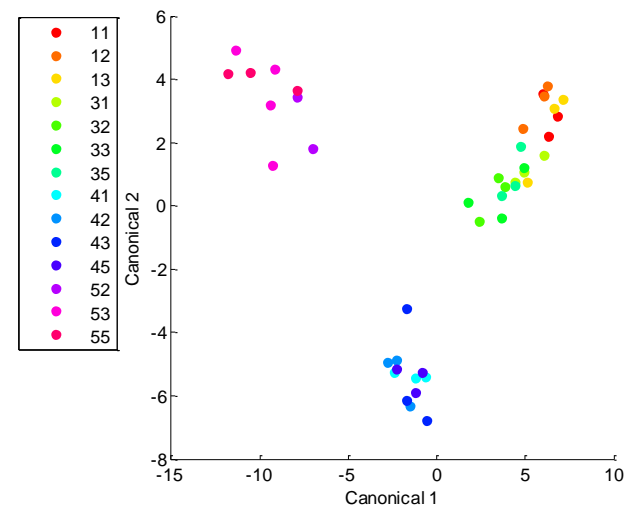
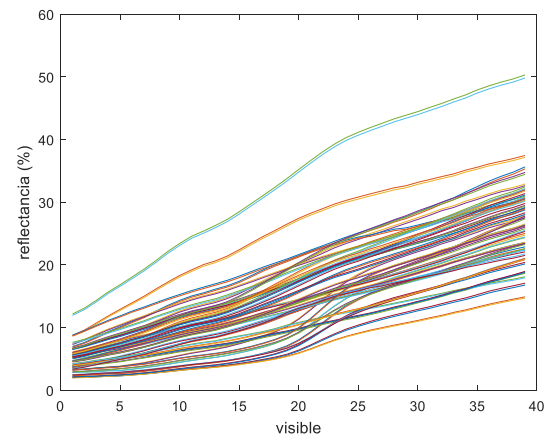
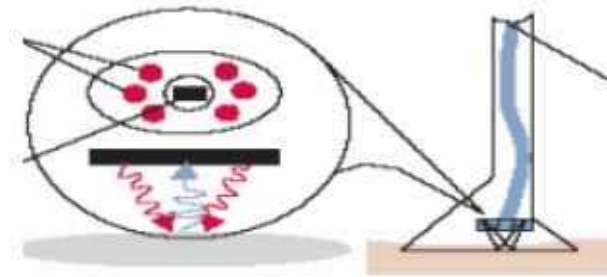
Discriminant

Supervised

Projection that  
maximises  
differences





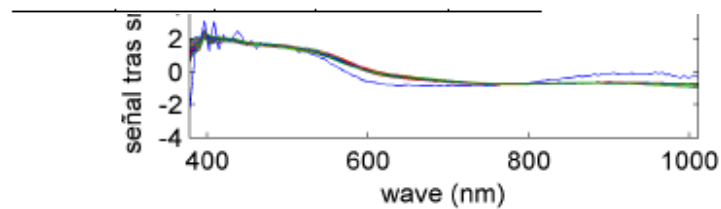


# Example: Summarised Spectral Information

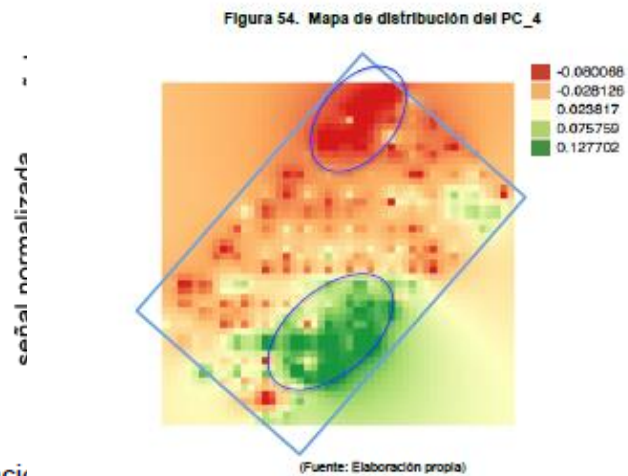
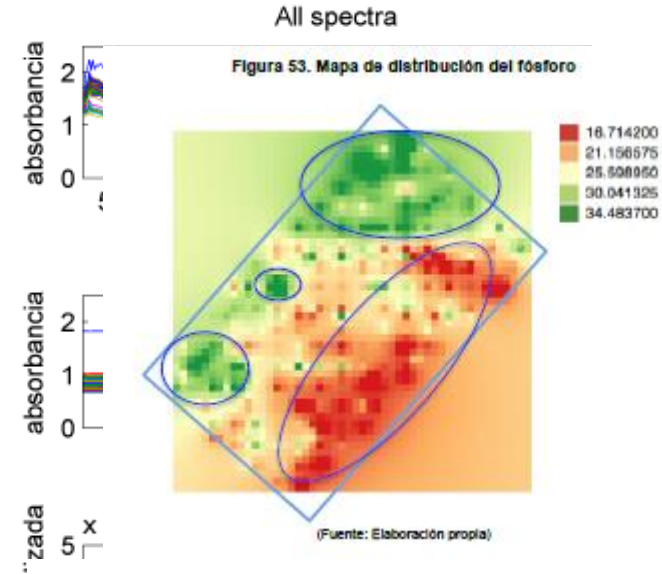


Tabla 8. Longitudes de onda (nm) importantes en cada uno de los 5 PC

PC_1	PC_2	PC_3	PC_4	PC_5
414	414	414	420	506
483	720	570	650-700	857
715	1916	624	1413	1610
	2145	965	1924	1805
		1103	2153	2005
		1394		2153
		1413		
		1912		
		2153		

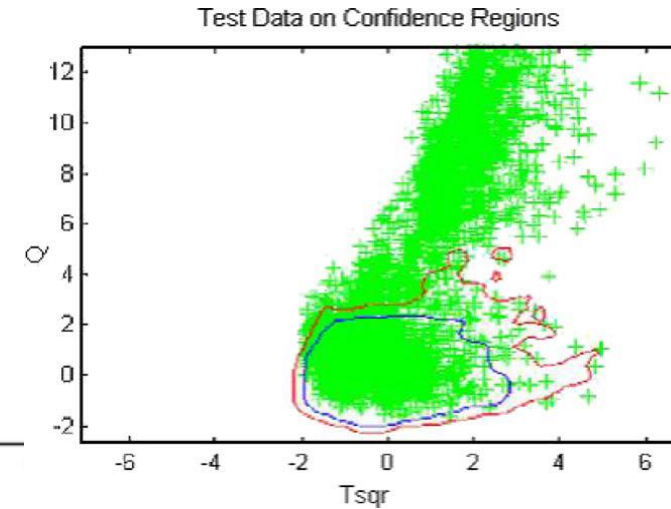
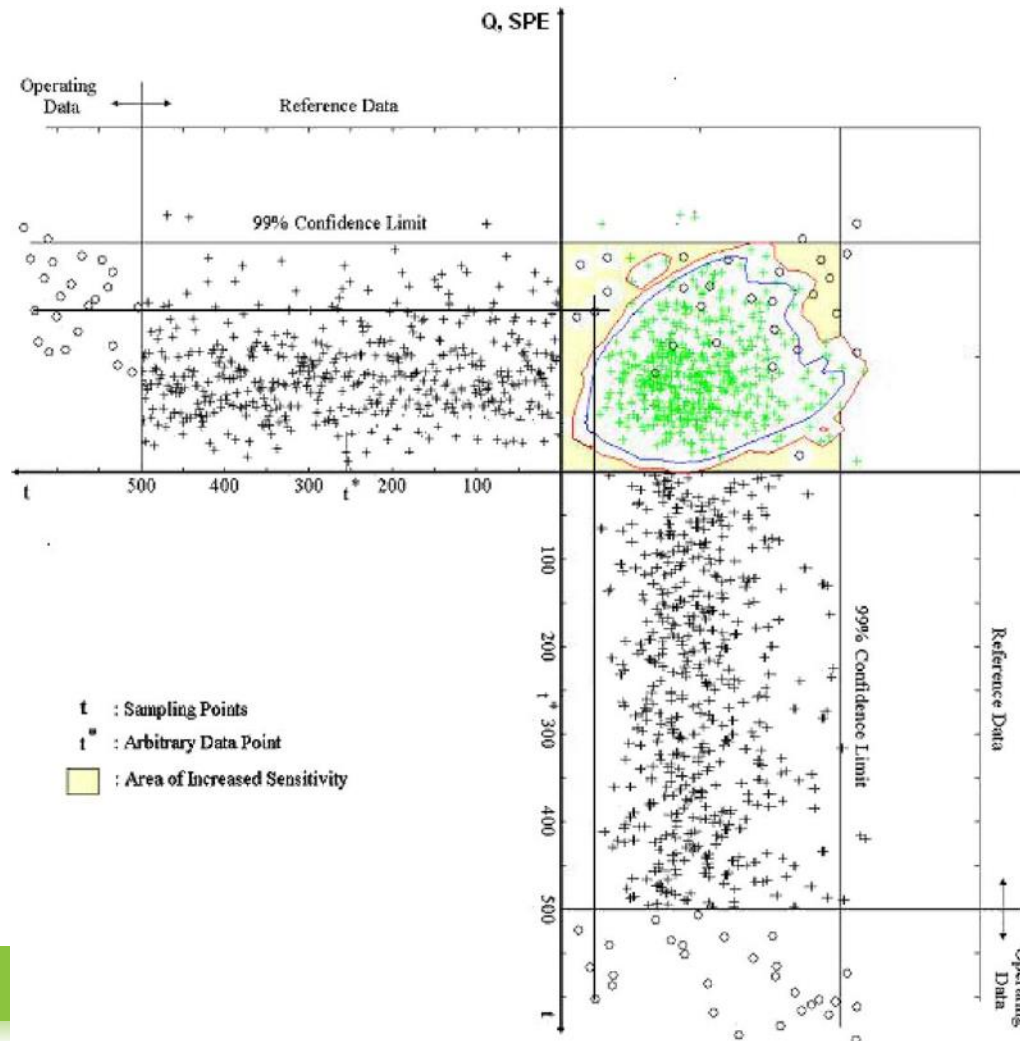


(Fuente: Elaboración propia)



# Multivariate process control

Use of T2-Q Parameters





*“Repensar lo pensado, des-saber lo sabido y dudar de nuestra propia duda, que es el único modo de empezar a creer en algo”*

(“Rethinking what is thought, un-thinking the known, and doubting your own doubt, is the only way to start believing in something”)

JUAN DE MAIRENA