

# MEMORIA PROYECTO FINAL: AGRO-TECH *CropIT*

**GRUPO: Tanya Moscoso, Germán Fernández**

## 1. DESCRIPCIÓN DEL PROYECTO

CropIT es una Agro-TECH especializada en el desarrollo de soluciones Big Data - ámbito de la Agricultura de Precisión

Enfocada en la toma de decisiones basada en el manejo y tratamiento de los datos permite maximizar el rendimiento de los cultivos mediante la optimización de los recursos empleados en la producción (agua, fertilizantes, energía, etc.)

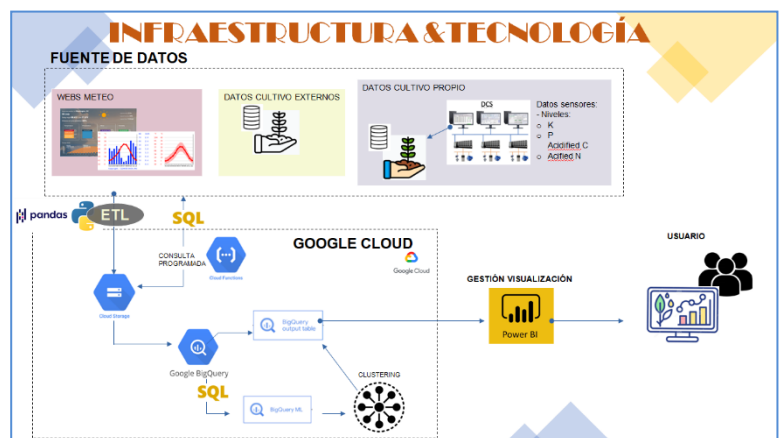
Destinada a:

- Pequeños y medianos agricultores que buscan mejorar las condiciones del sistema productivo a la vez que profundizar en el conocimiento de sus explotaciones.
- Técnicos de campo que realizan el manejo de una o varias explotaciones agrícolas.

## 2. INFRAESTRUCTURA Y TECNOLOGÍA

La infraestructura ha sido dividida en 3 bloques generales:

1. Fuente de datos. Especificada en el apartado 3.
2. Ecosistema de Google Cloud para la gestión de los datos.
  - a. Cloud Storage.
    - i. Almacenamiento de datos.
  - b. Cloud Functions.
    - i. Consultas programadas para la actualización periódica de la base de datos.
  - c. Google BigQuery.
    - i. Creación de tablas de salida para usar en PowerBI.
    - ii. Generación de modelos de ML para clusterizar los datos de suelo.
3. Gestión de visualización con PowerBI.



### 3. DATOS

#### 3.1. Base de Datos AfSIS (Africa Soil Information Service)

Se trata de una de las bases de tipo 'open data' más extensas de la región de África. Contiene tanto propiedades físicas como químicas de las distintas muestras de suelo que se han tomado y analizado en la región subsahariana a lo largo de 4 años (2009-2013). Todas las muestras están debidamente georreferenciadas.

Del archivo 'georeferences.csv' se han tomado como datos el id de la muestra (SSN) la latitud y longitud, así como el país y ciudad donde fue recogida.

De 'Wet\_Chemistry\_CROPNUTS.csv' se han tomado las medidas de conductividad eléctrica (EC), fósforo (M3 P), potasio (M3 K) y PH.

Por último, del archivo 'Wet\_Chemistry\_ICRAF.csv' se han tomado los datos de nitrógeno (Acidified nitrogen) y carbono orgánico (Acidified carbon); así como las composiciones texturales de los tres primeros horizontes ('Psa c1clay', 'Psa c1silt', 'Psa c1sand', 'Psa c2clay', 'Psa c2silt', 'Psa c2sand', 'Psa c3clay', 'Psa c3silt', 'Psa c3sand').

#### **¿Por qué se han seleccionado dichas variables?**

De los múltiples elementos que contiene el suelo, los macronutrientes primarios a destacar son nitrógeno, fósforo y potasio, dado que se necesitan en grandes cantidades en comparación con el resto de nutrientes. Por otra parte, el contenido en materia orgánica también es un factor muy importante a nivel del crecimiento y desarrollo de los cultivos. Adicionalmente, el nivel de pH influye en la disponibilidad de los nutrientes beneficiosos, siendo máxima cuando el nivel de pH tiende a la neutralidad (6.5-7). En suelos con pH ácido (<7) hay riesgo de toxicidad a causa del aluminio y del manganeso, además de una menor disponibilidad de algunos micronutrientes. En suelos básicos (>7) tanto los macro como los micronutrientes se encuentran menos disponibles. En cuanto a la conductividad eléctrica (EC), valores por encima de 1  $\mu\text{S}/\text{cm}$  podrían ser indicativos de un posible riesgo de salinidad por lo que habría que realizar un análisis más detallado para contrarrestar; no obstante, la mayor parte de las muestras presentaron valores por debajo de 1  $\mu\text{S}/\text{cm}$ , por lo que se trata de suelos normales – no salinos. Finalmente, en cuanto a la clase textural de los horizontes es importante tener en cuenta la proporción de las distintas partículas (arena, limo, arcilla) pues el diámetro de las mismas influye en la capacidad de retención del suelo, tanto de agua como de nutrientes, teniendo que maximizarse las aportaciones en texturas arenosas para cubrir el margen de pérdidas a causa del mayor tamaño de las partículas. En los suelos de texturas arcillosas el menor tamaño de las partículas aumenta la superficie activa total, favoreciendo la adsorción de los nutrientes por lo que se ve incrementada la fertilidad de los mismos.

### 3.2. Datos climatología

Se trata de un dataset de precipitaciones mensuales por cada una de las ciudades donde se dispone de datos de suelos, un total de 58. En algunas localizaciones se ha tomado el registro anual completo del año más reciente disponible en internet, y para aquellas donde no se ha encontrado registro o el registro anual no estaba completo, los datos se han generado de manera aleatoria.

La variable precipitación es de gran importancia dado que el agua es un elemento fundamental en el desarrollo de los cultivos. Tanto es así que en aquellas regiones donde se presenta un déficit en alguna época del año (precipitaciones inferiores a las necesidades del cultivo), éste debe ser corregido mediante riego artificial.

Así mismo, se dispone de un listado con los distintos cultivos que pueden darse en el territorio, las necesidades hídricas totales de cada uno de ellos ( $\text{m}^3/\text{ha}$ ) y el periodo (días) en que deben distribuirse las aportaciones.

### 3.3. Paneles fotovoltaicos

Se ha seleccionado un proveedor de paneles fotovoltaicos (Eco Green Energy) para extraer los datos de las especificaciones técnicas de cada panel y crear un dataset con los valores de potencia máxima entregada por panel, dimensiones y precio, lo que nos ha permitido crear la parte de optimización de los recursos energéticos del proyecto.

## 4. VISUALIZACIÓN

### ○ Principal

En la primera sección se puede encontrar un mapa con la localización geográfica de los puntos de muestreo acompañado de dos gráficos de barras, en los cuales se puede consultar el número total de muestras por país y, dentro de estas, el número de muestras por ciudad. Por decisión del propio equipo que ha elaborado la base de datos, los clúster se han efectuado en la mayoría de casos tomando una muestra superficial y otra subterránea por cada punto físico, por lo que en cada clúster se observarán 16 puntos a pesar de tratarse de un total de 32 muestras a nivel de datos. El contenido de nitrógeno, fósforo, potasio, carbono y pH se muestran en su etiqueta correspondiente, con una clasificación numérica y categórica del nivel que contiene la muestra, según los rangos que se muestran a continuación:

	Muy bajo 1	Bajo 2	Medio 3	Alto 4	Muy alto 5
CE en suspensión (mS/cm)	if > 1 posible riesgo de salinidad, analizar extracto de saturación				
C. orgánico (%)	<0.5	0.5-1.9	1.9-2.8	2.8-5	5-7.9
Nitrógeno (%)	<0.05	0.05-0.2	0.2-0.3	0.3-0.5	0.5-0.8
P (Olsen) (ppm) texturas gruesas	0-9	9-18	18-37	37-108	>108
K (ppm)	<30	30-100	100-160	160-350	>350
PH (adimensional)	<5 muy ácido	5-6.6 ácido	6.6-7.5 neutro	7.5-9 básico	>9 muy básico

Al detener el selector en cada punto ('tooltip') se muestra el contenido de los nutrientes de manera gráfica, así como el detalle de la composición textural de los 3 primeros horizontes.

- Necesidades hídricas

En el gráfico de barras se puede consultar los niveles de precipitación media mensual de la ciudad deseada, y su valor en mm en la parte inferior del mismo. Estos niveles de precipitación se han agrupado en las cuatro estaciones del año (invierno, primavera, verano, otoño) con el fin de orientarlos al ciclo del cultivo y facilitar la interpretación de los resultados. A la izquierda del panel se puede seleccionar tanto la ciudad como el cultivo de interés de la lista correspondiente, obteniéndose el total de precipitaciones en cada estación y el nivel de déficit hídrico (necesidades del cultivo – precipitaciones) que deberá aportarse mediante riego cuando el valor sea positivo (caudal necesario en  $\text{m}^3/\text{h}$ ).

Con el caudal necesario obtenido se deberá seleccionar el sistema de riego adecuado del que obtendremos las necesidades de potencia para nuestra instalación fotovoltaica.

- Paneles PV

En esta parte el usuario puede seleccionar el tipo de panel que más le convengan teniendo en cuenta sus necesidades energéticas (sacadas de la parte de 'Necesidades hídricas'). La herramienta le permite 'jugar' con varios datos (potencia, dimensiones, precio) para valorar, según sus necesidades de potencia y el espacio que tenga disponible, cuál es la instalación fotovoltaica óptima para sus requerimientos.

- Estimaciones nutricionales

En este apartado se calcula el contenido específico de nitrógeno, fósforo y potasio (N-P-K) de cada muestra en función de la profundidad del horizonte, misma que puede variar en el rango de 0.2 y 0.5m, y la densidad aparente del material, con valores entre 1.1 y 1.7  $\text{tn}/\text{m}^3$  como normal general. Una vez calculado el volumen y el peso de la hectárea de parcela concreta, se obtiene el contenido de cada elemento en  $\text{kg}/\text{ha}$ . Estos niveles serán tenidos en cuenta para determinar los elementos que se encuentran en déficit y en qué cantidad, adaptando así las recomendaciones genéricas de fertilización de los cultivos a las condiciones específicas del medio concreto a utilizar.

Se adjunta el archivo de PowerBI con la información comentada.

## 5. CLUSTERING

Hemos usado un modelo de clustering para intentar definir los niveles de nutrientes óptimos para conseguir un pH neutro, o lo más cercano posible a él. Vemos que con un modelo k-means para 9 clusters, el cluster nº6 tiene un pH muy próximo a 7, nuestro valor objetivo.

Teniendo en cuenta estos parámetros podemos monitorizar nuestro cultivo para asegurar que todos los valores de nutrientes permanecen en los rangos requeridos y, en caso de necesitarlo, aportar nutrientes si detectamos carencia en alguno de los valores o corregir el exceso de ser necesario mediante una fracción de riego adicional para el lavado de sales.

						K means Clustering
centroid_id	M3_K	M3_P	PH	Acidified_carbon	EC	Acidified_nitro...
5	1.746,44	43,89	7,5	2,43	2,59	0,15
4	97,93	7,44	5,74	0,74	0,05	0,05
6	477,65	160,66	6,94	1,37	0,14	0,1
1	417,53	17,2	8,18	0,78	0,3	0,06
9	98,19	7,56	6,1	0,75	0,06	0,05
7	190,37	11,3	5,96	5,12	0,18	0,3
3	182,49	11,4	5,97	2,63	0,14	0,18
2	93,14	7,03	5,9	0,78	0,05	0,05

## 6. OTROS ARCHIVOS

Tanto la presentación como los archivos de Python y mapa de Kepler en formato HTML, se adjuntan adicionalmente al presente documento.