



Centro Internacional de Agricultura Tropical
Desde 1967 *Ciencia para cultivar el cambio*

Uso de minería de datos en agricultura para hacer frente al cambio climático.

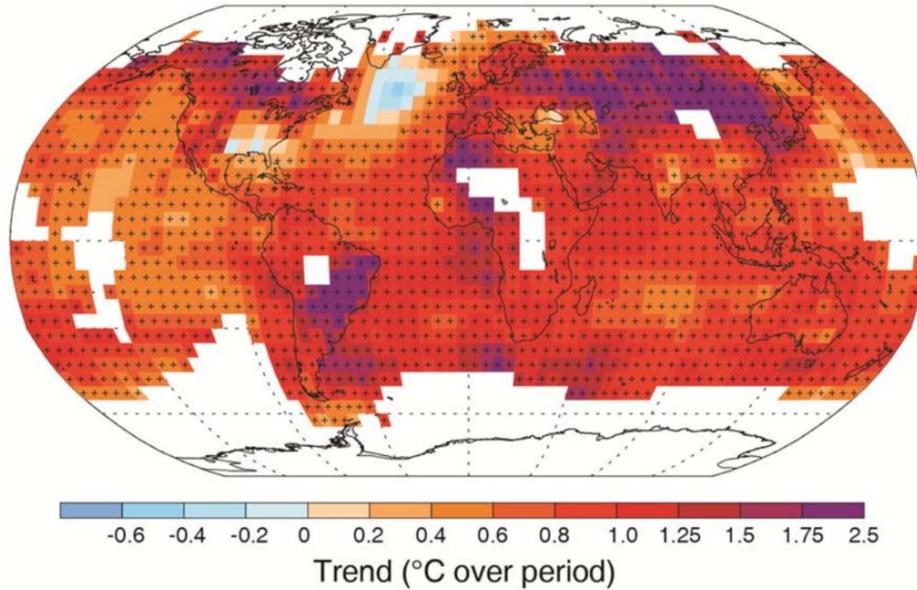
M.Sc. Hugo Andrés Dorado B.
Investigador Asociado del CIAT
h.a.dorado@cgiar.org



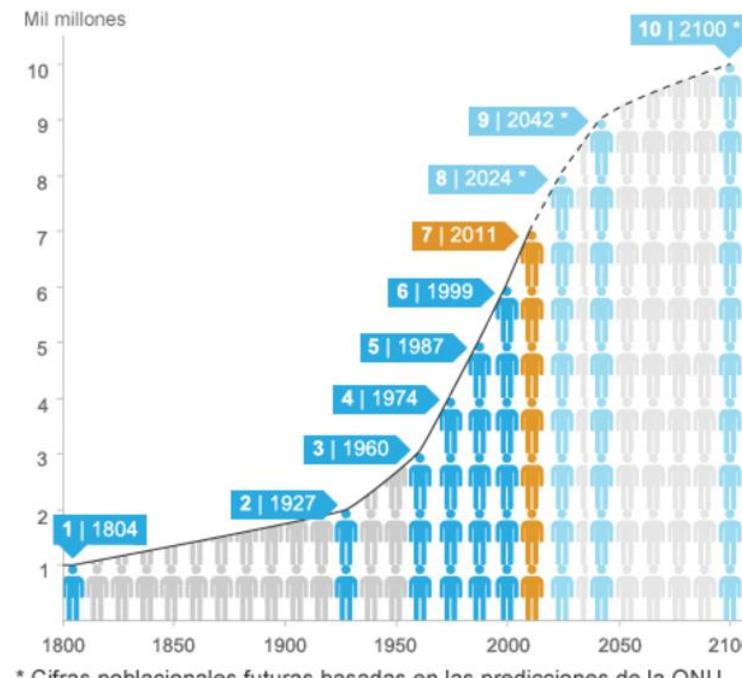
El CIAT es un Centro de Investigación de CGIAR

Cambio climático y crecimiento poblacional

Observed change in average surface temperature 1901–2012



Crecimiento de la población mundial: alcanzando 7 mil millones



FUENTE: Fondo de Población de la ONU

Según el reporte de la FAO, “Agricultura mundial en la perspectiva del año 2050”

Necesitamos **incrementar la producción mundial agrícola significativamente** antes de 2050.

La agricultura debe adaptarse a nuevas las condiciones climáticas

Agricultura en el 2000

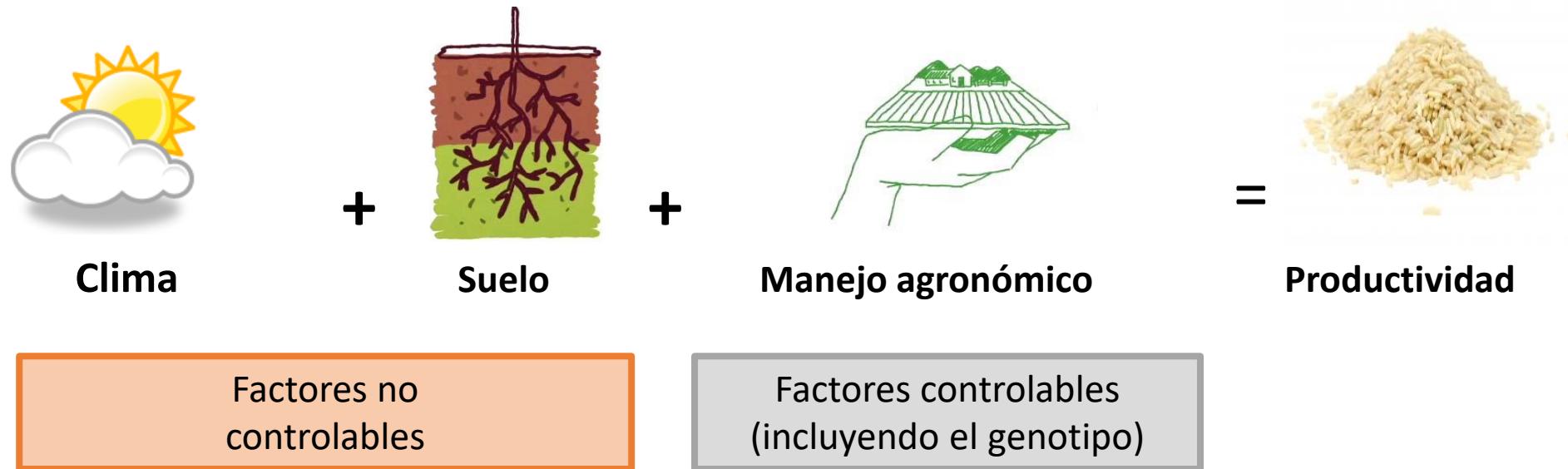


Agricultura en el 2019, colectando y diseminando datos a gran escala



¿Cuáles son los factores que afectan el rendimiento?

Aspectos que afectan el rendimiento

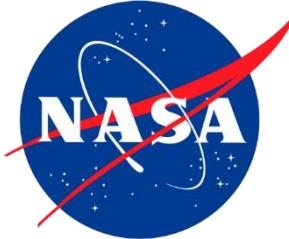


Identificar factores o combinaciones de factores que conducen a altas o bajas producciones

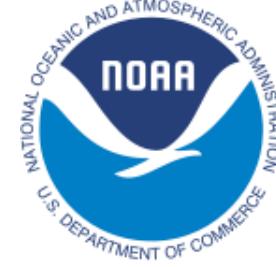
% ? + % ? + %? = A Explicar (100 %)

Algunas fuentes de información (Datos secundarios y primarios)

Clima



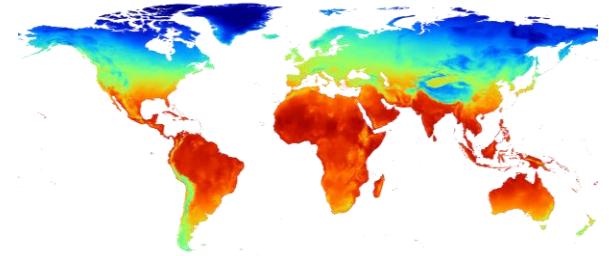
<https://power.larc.nasa.gov/>



<https://www.ncdc.noaa.gov/cdo-web/>



<https://www.senamhi.gob.pe/>



<http://worldclim.org/version2>

Suelo



OpenTopography

High-Resolution Topography Data and Tools

<https://opentopography.org>



SOILGRIDS

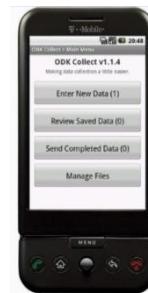
A system for automated global soil mapping

<https://soilgrids.org>

Cultivo



Statistics



<https://opendatakit.org>

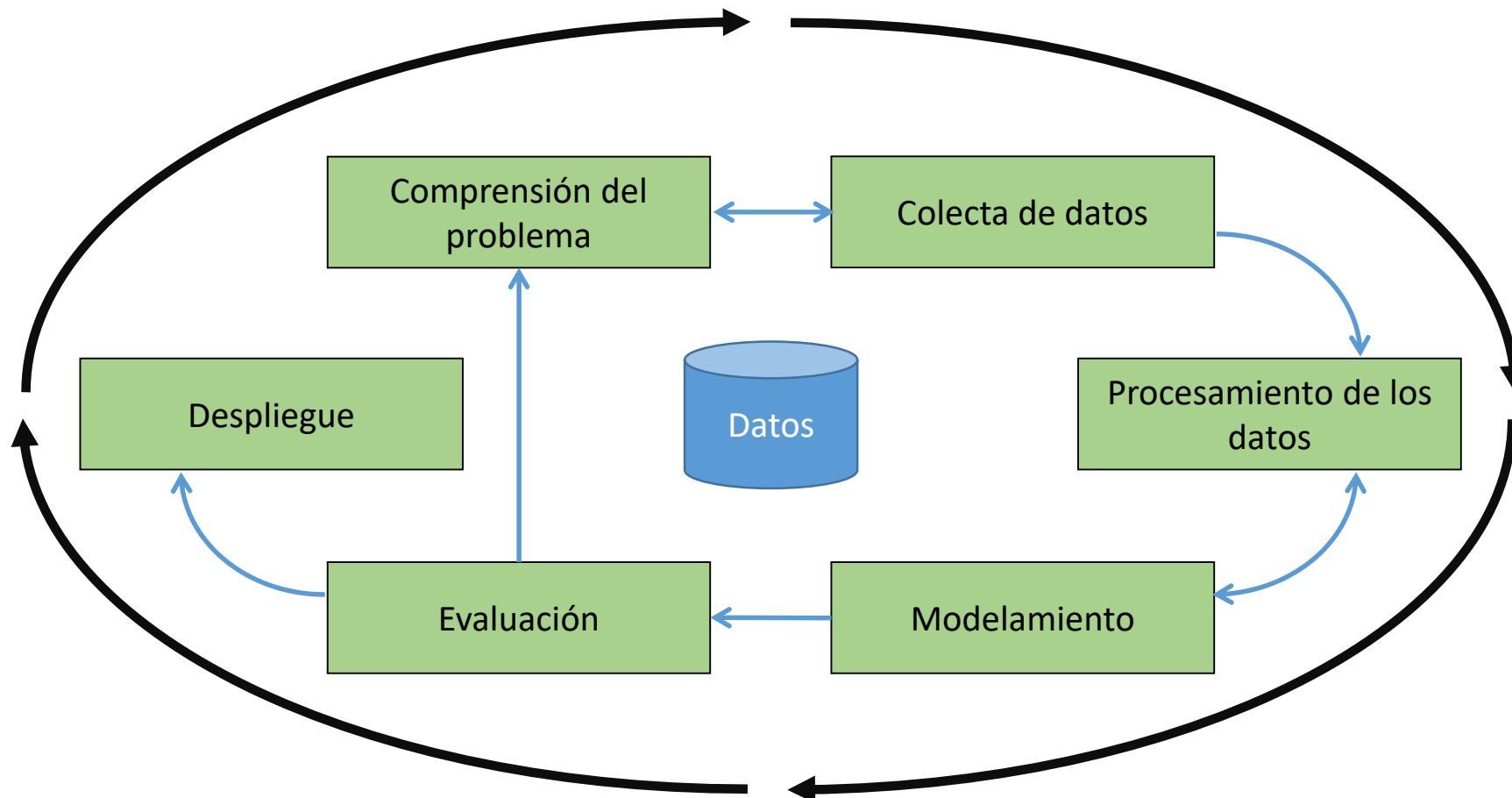
Instituciones públicas o privadas

[ciat.cgiar.org](http://www.fao.org/statistics/en/) <http://www.fao.org/statistics/en/>

Forjando la sostenibilidad alimentaria futura



Flujo análisis de información



Basado en

CRISP-DM
CROSS INDUSTRY STANDARD PROCESS FOR DATA MINING

Procesamiento y limpieza de la información

Distintos formatos de fecha y caracteres

AuxVar	Y	Lote	FechaSiembra	FechaCosech	Tecnico	Área	Cod. Act.
CHARCO-10_Lf	1	CHARCO-10	12/6/2014	4/26/2015	Silvio_Valle	12.3	9_1_Fitosanil
CHARCO-10_2_	1	CHARCO-10	12/6/2014	4/26/2015	Silvio_Valle	12.3	12_2_Fitosanil
CHARCO-10_Co	1	CHARCO-10	12/6/2014	4/26/2015	Silvio_Valle	12.3	45_Control_G
CHARCO-10_Co	1	CHARCO-10	12/6/2014	4/26/2015	Silvio_Valle	12.3	8_Control_P
CHARCO-10_Co	1	CHARCO-10	12/6/2014	4/26/2015	Silvio_Valle	12.3	36_Control_P
CHARCO-10_Ds	1	CHARCO-10	12/6/2014	4/26/2015	Silvio_Valle	12.3	19_Desinfecci
CHARCO-10_Fel	1	CHARCO-10	12/6/2014	4/26/2015	Silvio_Valle	12.3	27_Fertilizaci
CHARCO-10_Fel	1	CHARCO-10	12/6/2014	4/26/2015	Silvio_Valle	12.3	27_Fertilizaci
CHARCO-10_Fel	1	CHARCO-10	12/6/2014	4/26/2015	Silvio_Valle	12.3	27_Fertilizaci
CHARCO-10_Fo	1	CHARCO-10	12/6/2014	4/26/2015	Silvio_Valle	2	31_Foqueo_F
CHARCO-10_Fo	1	CHARCO-10	12/6/2014	4/26/2015	Silvio_Valle	2.5	31_Foqueo_F
CHARCO-10_Ge	1	CHARCO-10	12/6/2014	4/26/2015	Silvio_Valle	12.3	57_Germinaci
CHARCO-10_Pri	1	CHARCO-10	12/6/2014	4/26/2015	Silvio_Valle	12.3	6_Primera_F
CHARCO-10_Sc	1	CHARCO-10	12/6/2014	4/26/2015	Silvio_Valle	12.3	7_Seunda_F

Cronologico

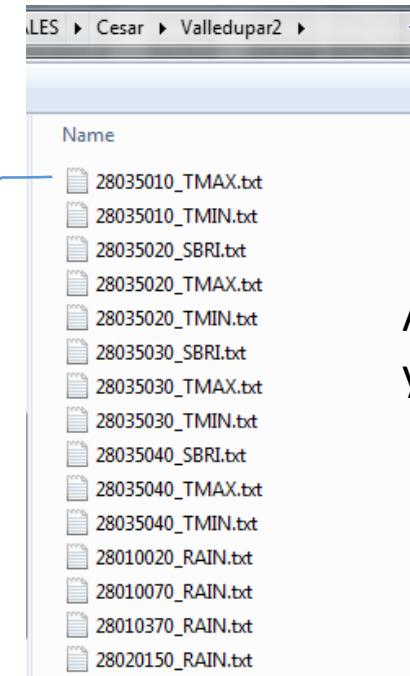
ProdAplicados

MonitoreoPLAGAS

Coordenadas mal registradas

37	5°08'27.5"	-75°54'31.3"	RISARALDA	APIA
38	5°08'42.3"	-75°55'02.2"	RISARALDA	APIA
39	5°78'41.0"	-75°05'02.8"	RISARALDA	APIA
40	5°67'16.8"	-75°04'17.0"	RISARALDA	APIA
41	5°08'17.8"	-75°54'18.4"	RISARALDA	APIA
42	5°09'41.8"	-75°55'26.4"	RISARALDA	APIA
43	5°09'41.6"	-75°55'26.1"	RISARALDA	APIA
44	5°09'35.3"	-75°55'10.0"	RISARALDA	APIA

<https://bigdata.cgiar.org/data-cleaning-the-unsexy-but-essential-aspect-of-data-science/>



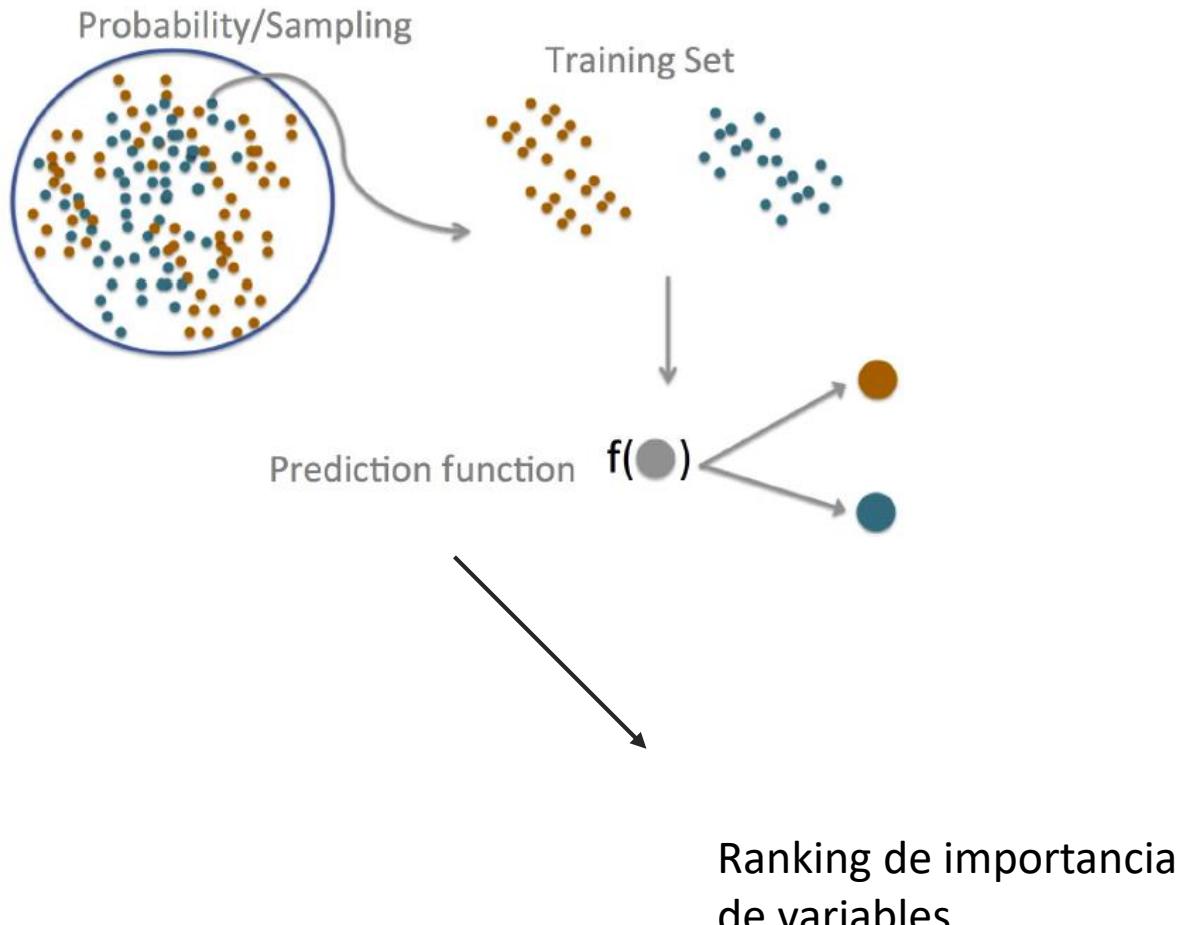
Archivos desagregados
y en otros formatos

Date	value
19800101	NA
19800102	NA
19800103	NA
19800104	NA
19800105	NA
19800106	NA
19800107	NA
19800108	NA
19800109	NA
19800110	NA
19800111	35.2
19800112	NA
19800113	NA
19800114	36.2
19800115	35.2
19800116	NA

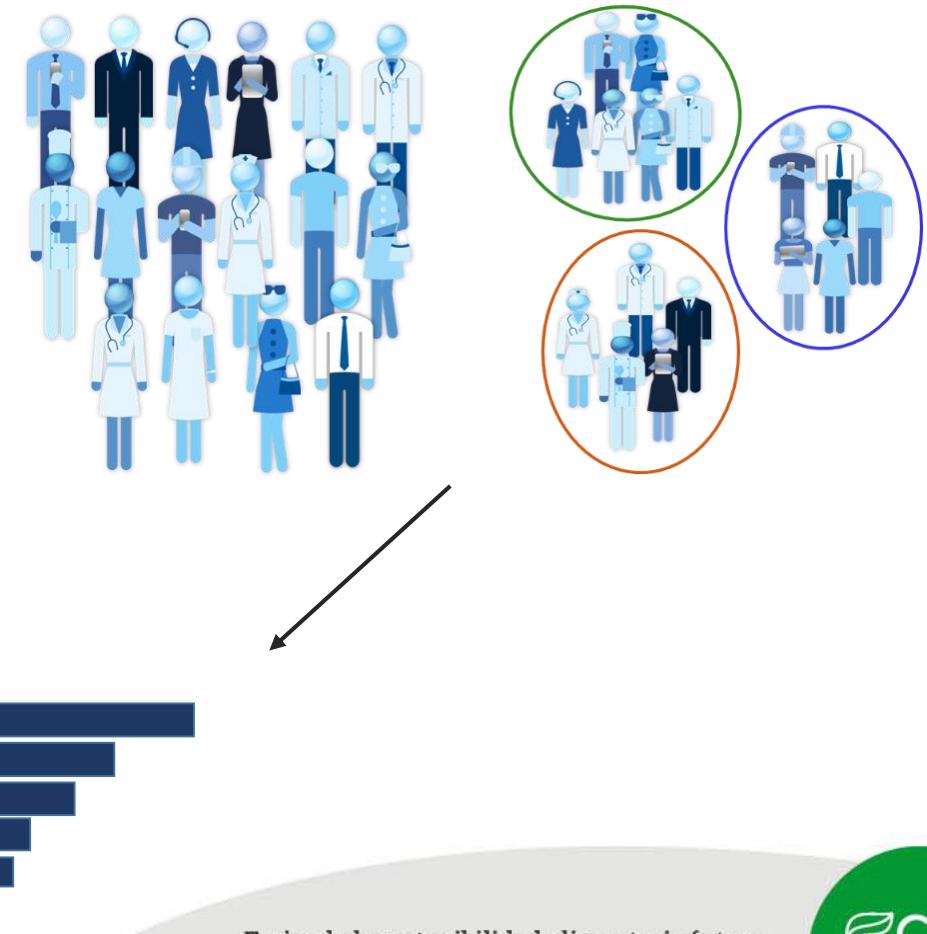
Valores perdidos

Objetivos del modelamiento de datos

- Predicción



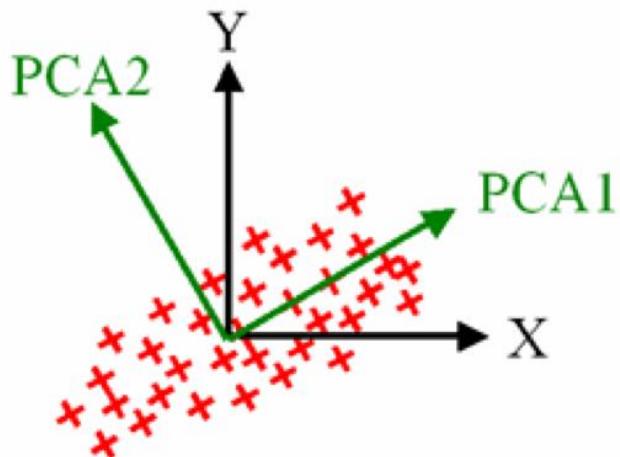
- Agrupación



Métodos de análisis

Métodos tradicionales

- Regresión lineal multiple (OLS)
- Análisis factoriales (PCA, MCA, CATPCA)
- Modelo lineal generalizado (GLM)
- Modelos mixtos



Métodos basados en aprendizaje automático

- Redes neuronales artificiales
- Random Forest - Conditional Inference Forest
- CART
- Support Vector Machine
- Optimización



Diseminación de información a usuarios finales

Directamente hacia los productores por medios presenciales.

- Asistencia técnica
- Escuelas de campo (ECAs)
- Mesas técnicas agroclimáticas



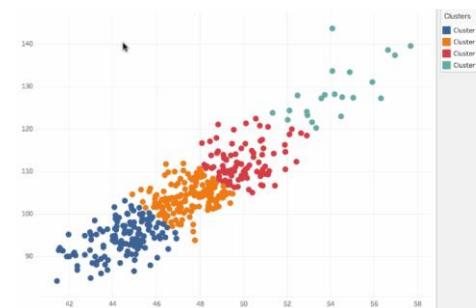
Directamente hacia los productores por medios digitales

- Aplicaciones móviles.
- Chatbot. (Asistente de texto)
- IVR (Interactive voice response), llamadas
- Mensajes de texto



Otros usuarios (Gobierno, investigadores,...)

- Visualizaciones interactivas.
- Blogs.
- Notes de prensa.
- Publicaciones



Casos de estudio en arroz - Colombia

Proyecto.



MINAGRICULTURA



Preguntas de investigación.

¿Qué efecto tiene el clima sobre el rendimiento en arroz?



Fuente de datos

- Encuesta nacional de arroz.
- Registros de cosecha.
- Experimentos de fecha de siembra.
- Estaciones meteorológicas (IDEAM).

Observaciones

1237

6000

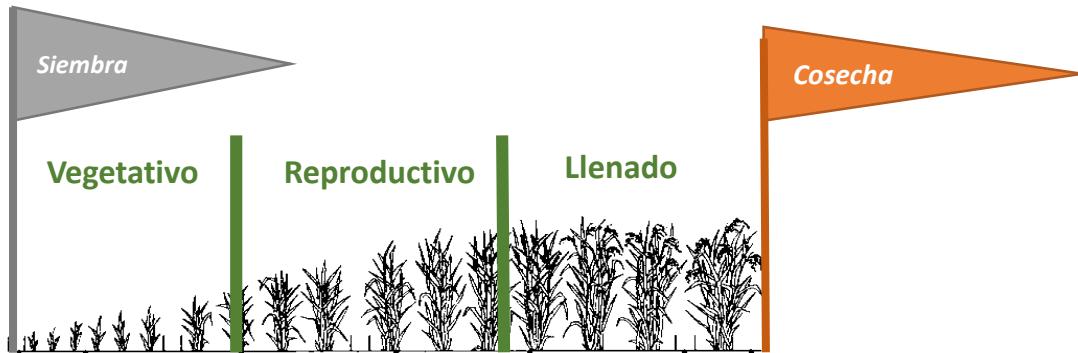
600

27

7837

Efectos del clima sobre variedades de arroz

Unidad de observación – Ciclo de cultivo



Variables climáticas por etapa (X_1, X_2, X_3, \dots)

y

Variables involucradas

- Fecha de siembra - Cosecha
- Variedad (F733, F60, Lagunas y F174)
- Localidad (Saldaña, Villavicencio)
- Temperatura promedio
- Energía solar acumulada
- Precipitación acumulada
- Número de días con precipitación mayor a 10 mm
- Promedio de humedad relativa
- Rendimiento en kg/ha

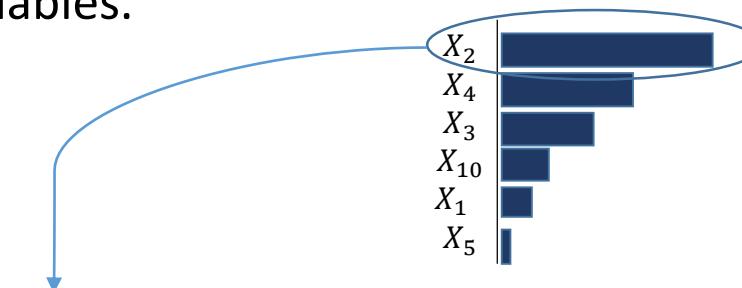
Efectos del clima sobre variedades de arroz

Metodología utilizada.

Random forest - importancia de variables (Breiman, 2001)

- ✓ Mejor desempeño en calidad de predicción (RMSE).
- ✓ Soportar relaciones no lineales.
- ✓ Equilibrar la influencia por correlación entre las variables.

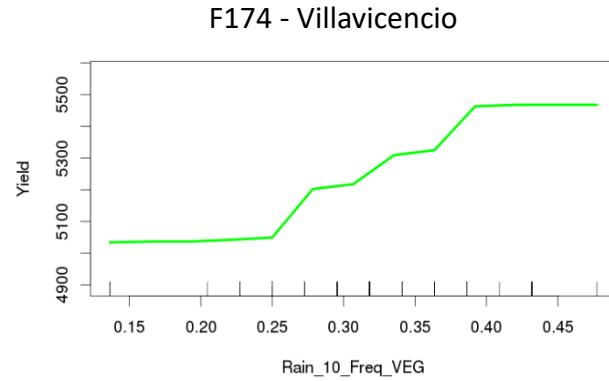
Resultados.



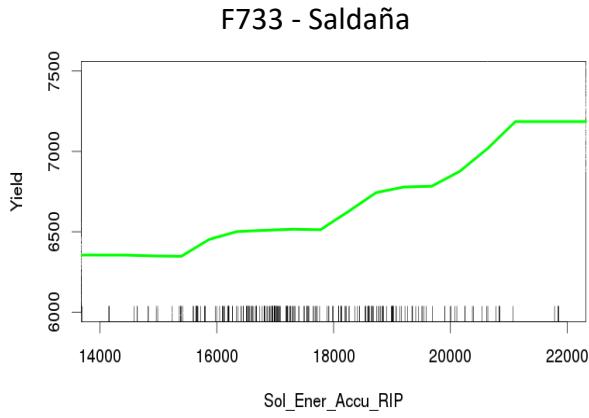
Modelo	Observaciones	Desempeño	Variable mas relevante
Saldaña - F733	267	29.90%	Energia solar acumulada en fase de llenado de grano
Saldaña - F60	150	46.60%	Temperatura promedio en fase reproductiva
Saldaña -Lagunas	187	6.96%	Temperatura promedio en fase de llenado de grano
Villavicencio - F174	134	28.10%	Días con lluvia en fase vegetativa

Efectos del clima sobre variedades de arroz

Dependencias parciales. – Condiciones climáticas asociadas a altos rendimientos



Precipitación en el **40%** de los **días** de la fase vegetativa



Radiación solar acumulada en todo el ciclo de llenado de grano mayor a **21.000 cal/cm²**

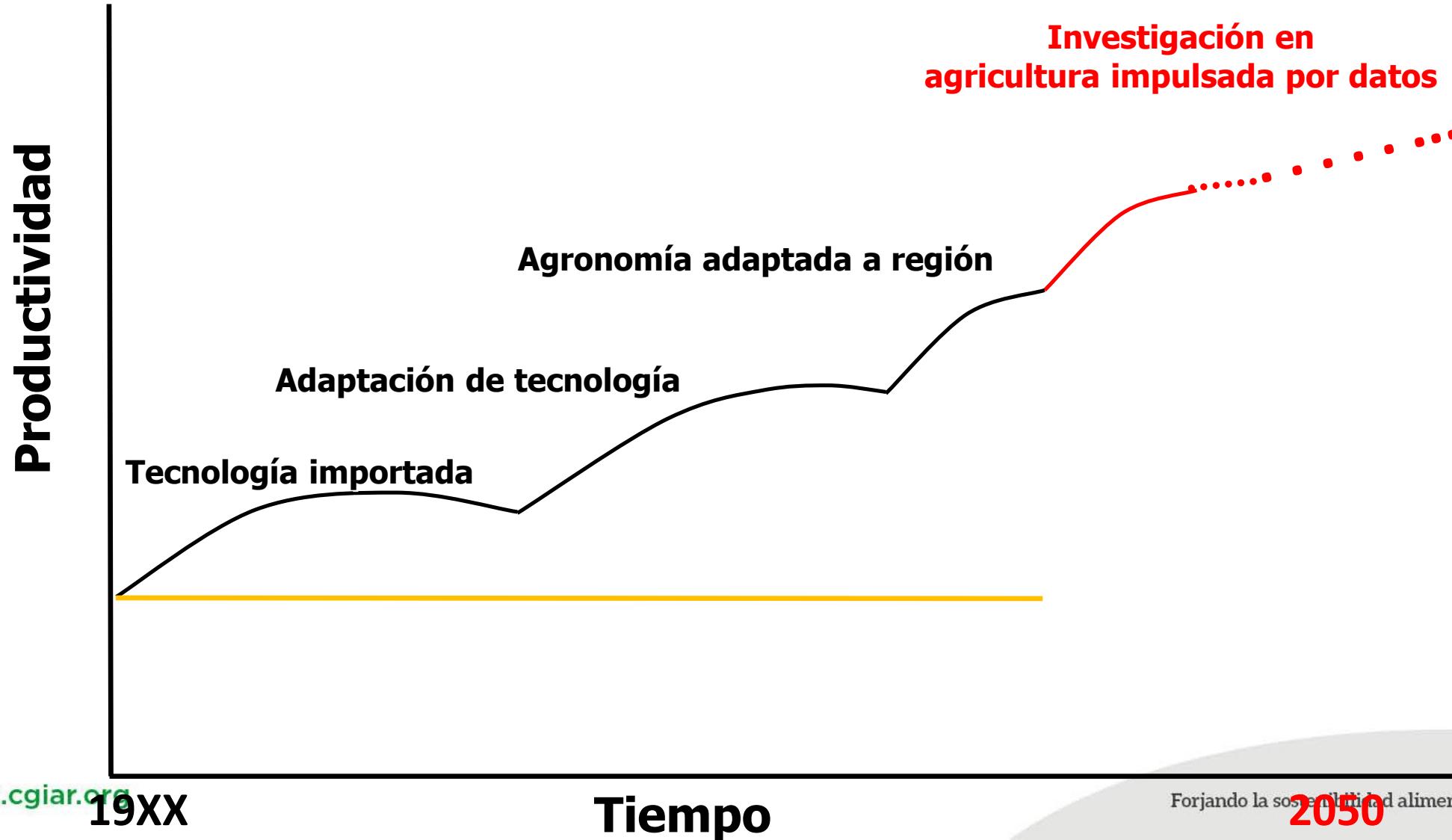
Mecanismos para tomar mejores decisiones frente al cambio climático.

Productores: Plantear fechas de siembra estratégicamente en el año para acomodarse a las mejores condiciones según la variedad que se vaya a sembrar.

Mejoradores: Probar la efectividad y resistencia a factores climáticos de variedades desarrolladas a partir de datos de productores.

Mejores decisiones – Mayor productividad!!

Evolución de la productividad en países en vía de desarrollo



Herramientas y software

Procesamiento de datos



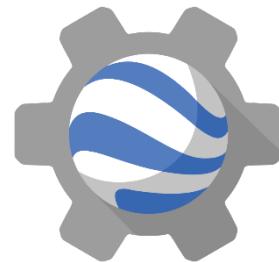
Análisis de datos



Datos espaciales



Quantum Gis

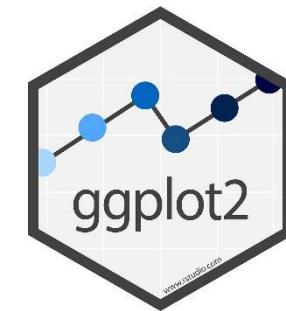


Google earth engine

Computación en la nube



Visualización de datos



Conclusiones

- Uno de los **sectores** más **afectados** por el **cambio climático** es la **agricultura** y dado que tenemos el reto de alimentar una población que cada vez es más grande, debemos encontrar **estrategias** que permitan adaptarse a **nuevas condiciones climáticas**.
- A través del **análisis de datos** podemos descubrir ciertos **patrones** que nos permiten encontrar la manera en que el clima afecta los procesos productivos y con ello tomar **mejor decisiones**.
- Existen **recursos en línea disponibles** y de manera gratuita para implementar y formarse en análisis de datos.

¡Gracias!



NOS ENORGULLECE
HABER CELEBRADO 50 AÑOS
DE INVESTIGACIÓN AGRÍCOLA
PARA EL DESARROLLO

Centro Internacional de Agricultura Tropical - CIAT

Sede Principal y Oficina Regional
para Suramérica y el Caribe

 +57 2 445 0000

Km 17 Recta Cali-Palmira
A.A. 6713, Cali, Colombia

 ciat@cgiar.org

 ciat.cgiar.org



El CIAT es un Centro de Investigación de CGIAR