



International Center for Tropical Agriculture
Since 1967 / *Science to cultivate change*

Big data para apoyar la toma de decisiones en agricultura

Datos

Métodos

Investigaciones

Hugo Dorado

h.a.dorado@cigar.org



“Hubo 5 exabytes de información creada entre comienzos de la civilización hasta el 2003, pero ahora esa información es creada cada 2 días.” Erick Schmidt, de google



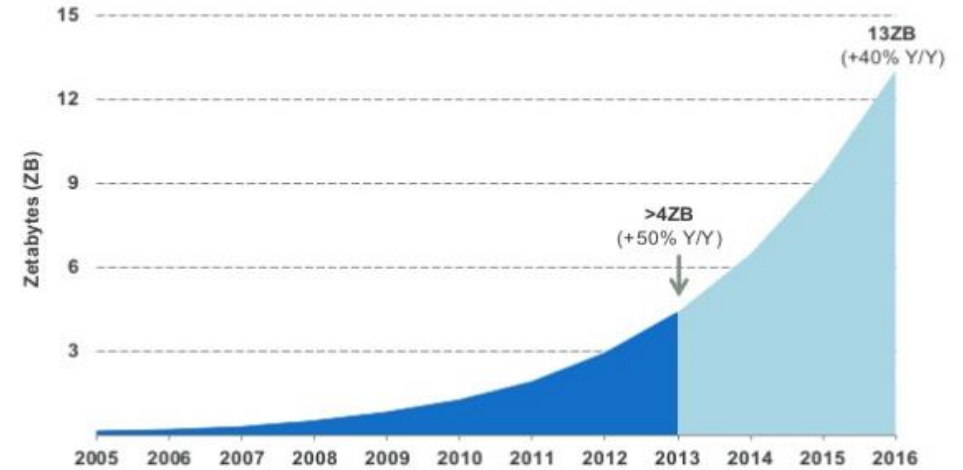
1.6 billones de usuarios activos

Más de 40.000 búsquedas por segundo de google!

6000 tweets por segundo!



<http://www.internetlivestats.com/>



@KPCB Note: 1 petabyte = 1MM gigabytes, 1 zetabyte = 1MM petabytes. Source: IDC Digital Universe, data as of 5/14.

¿Que viene pasando en la agricultura?



Cómo buscamos información en agricultura?

Registros de cosecha.



Estaciones meteorológicas de IDEAM.



Plataformas web para registro y reporte.



Aplicaciones móviles.



Percepción remota.



Drones



cenicaña



Gremios y centros de investigación.



A qué necesidad responde?

Qué sembrar, dónde, cómo, cuándo

Tradicionalmente las decisiones por agricultores han sido basadas en:

- Experiencias exitosas
- Incentivos
- Tendencias-Mercados
- Conocimiento de expertos

Diferentes de metodologías de extensión:

- Top-Down
- Agricultor por agricultor (Asistencias técnica)
- Grupos (CropCheck, GTT, ECAs, parcelas demostrativas, benchmarking)



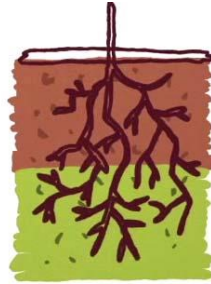
Qué proponemos ?

Identificar las características que hace que el rendimiento varíe !!!



Clima

+



Suelo

+

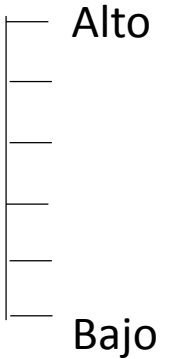


Manejo agronómico

=



Productividad



Factores no controlables

Factores controlables

PARA:

Descubrir señales (combinación de factores) – Metodologías flexibles !!!

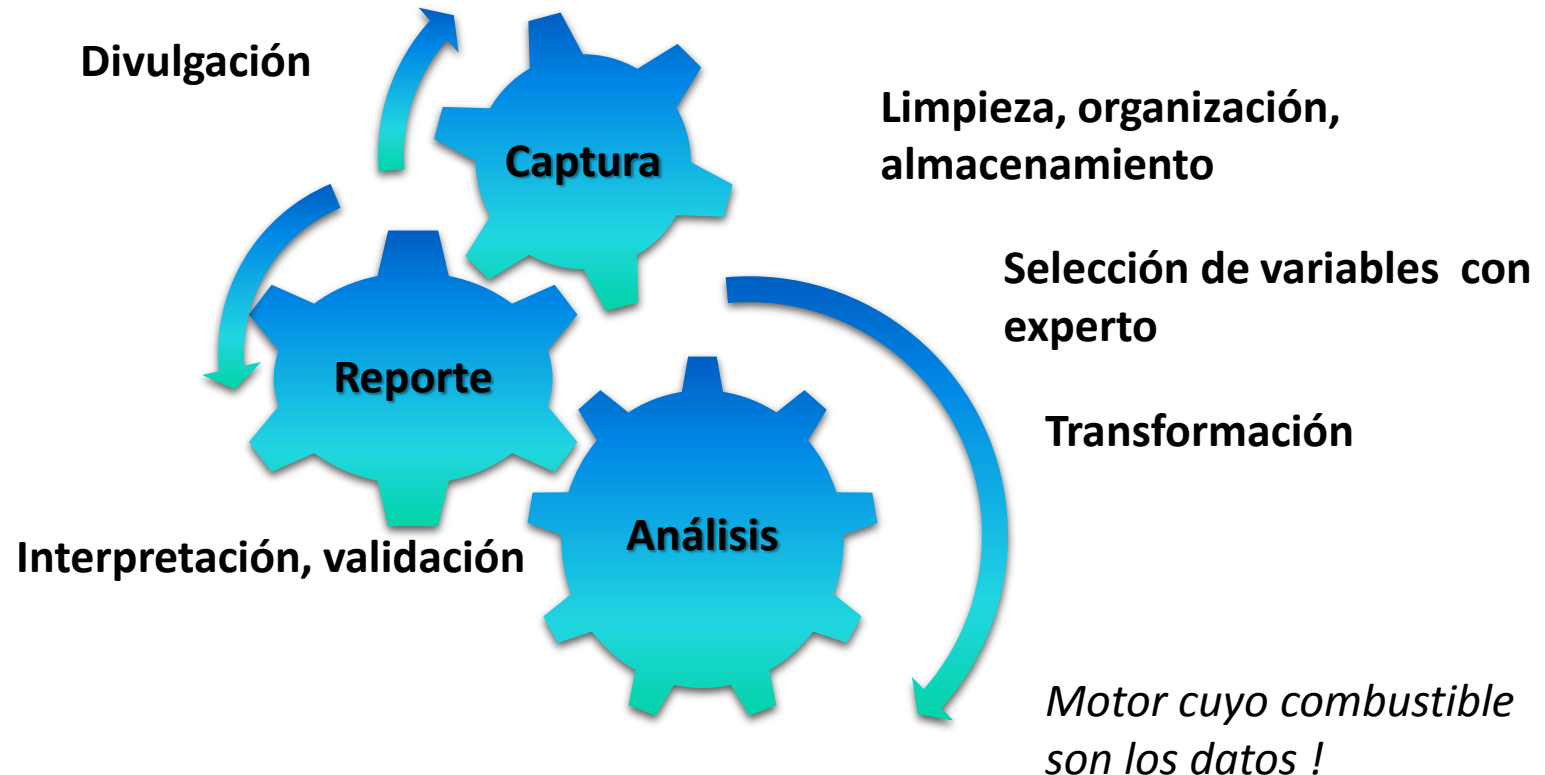
% ? + % ? + % ? = A Explicar (100 %)

Cómo hacerlo?

punto de partida

Información primaria: capturada a través de diferentes fuentes –
Plataforma en línea

Información secundaria : Bases de datos existentes



Cómo? - métodos de análisis



Métodos estadísticos tradicionales

Regresión lineal múltiple (OLS)

Análisis factoriales (PCA, MCA, CATPCA)

Modelo lineal generalizado (GLM)

Modelos mixtos



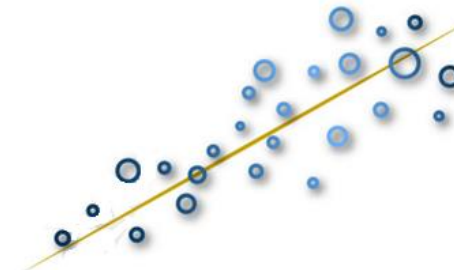
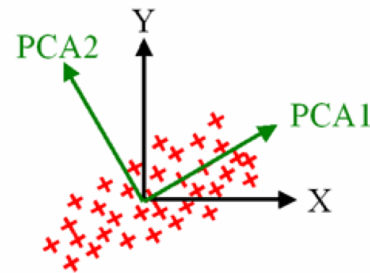
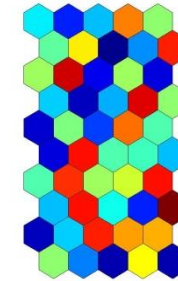
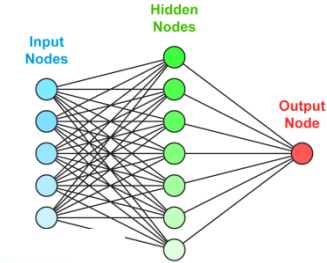
Métodos basados en Machine learning

Redes neuronales artificiales (supervisadas, no supervisadas)

Random Forest

Conditional Forest

Lógica difusa



Casos de estudio

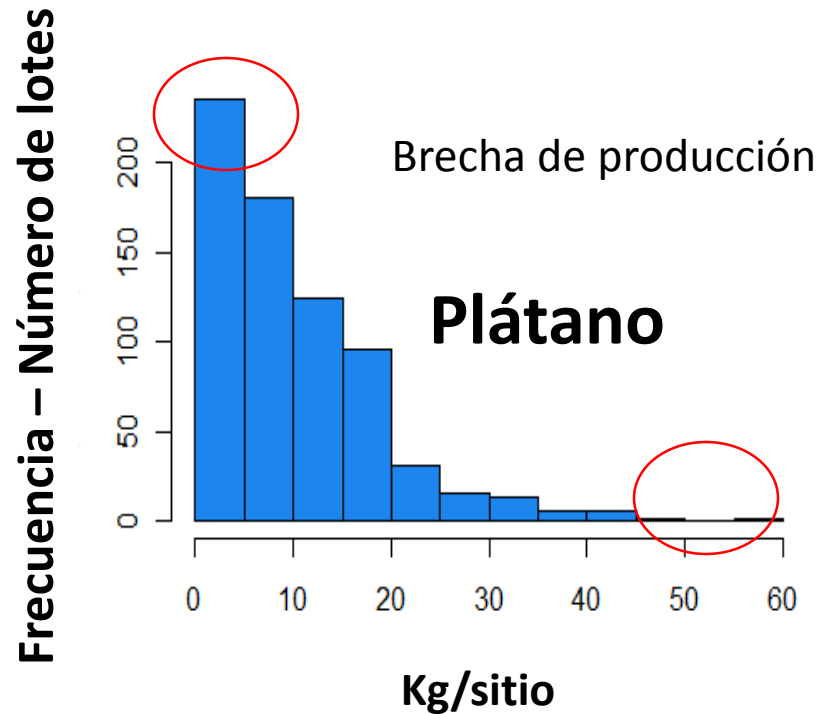


- **(Caso 1) Clima, Suelo y Manejo** Plátano en Colombia.
- **(Caso 2) Suelo y manejo** : Maíz en Córdoba
- **(Caso 3) Manejo agronómico y Suelo:** Banano en Magdalena (Información secundaria)





Caso de estudio **Plátano en Colombia: Clima, Suelo y Manejo agronómico**



Hipótesis

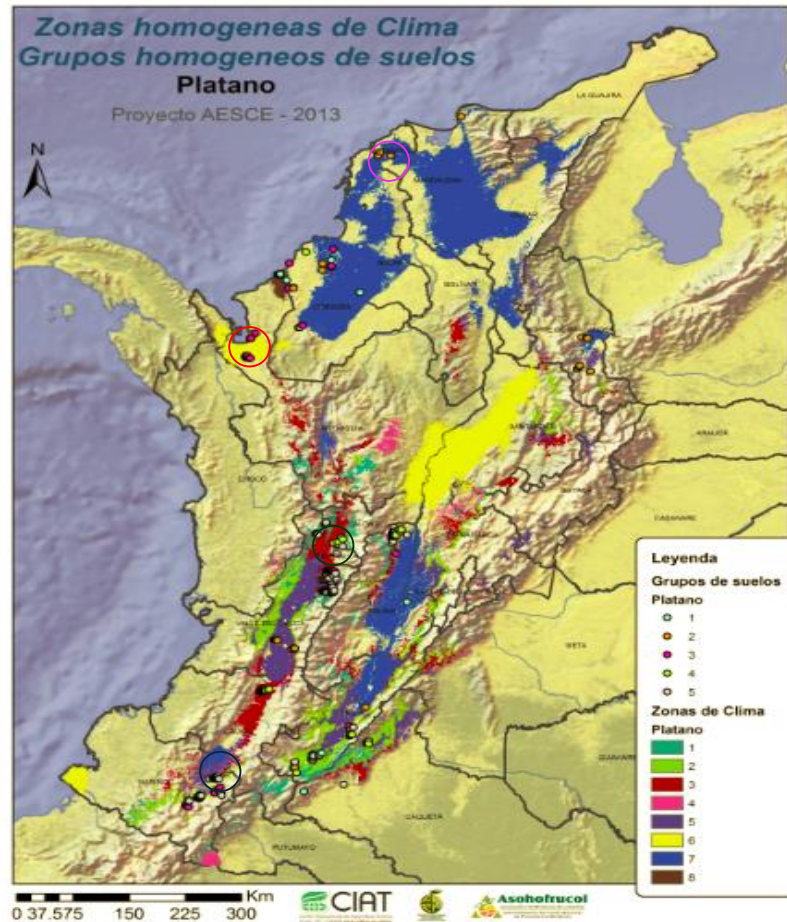
Más de 800 lotes productivos – representando heterogeneidad de condiciones



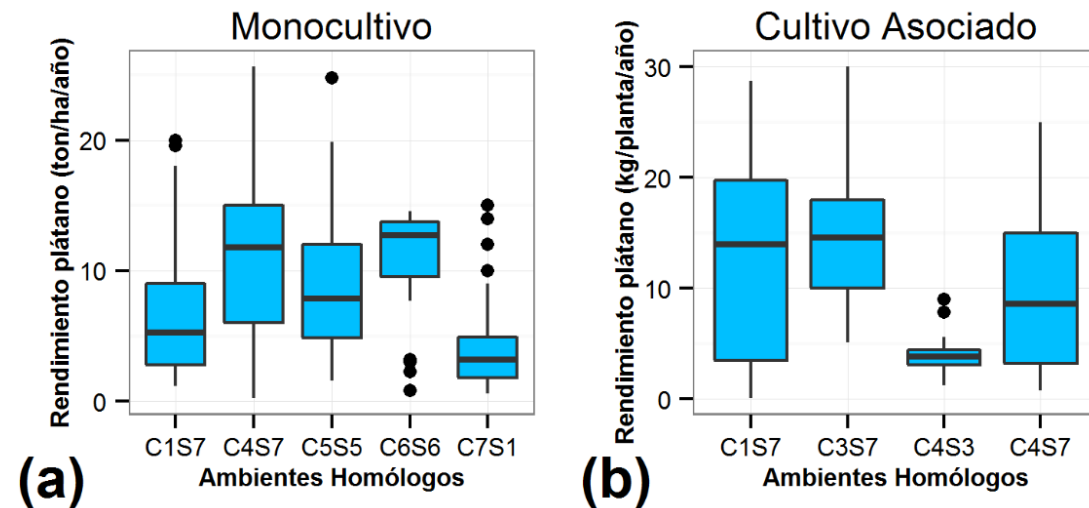


Caso de estudio Plátano en Colombia: Clima, Suelo y Manejo agronómico

Análisis factoriales: PCA- CATPCA. Modelos mixtos y BLUPS



Suelo y manejo: Caracterizado directamente de los lotes.
Clima: Worldclime



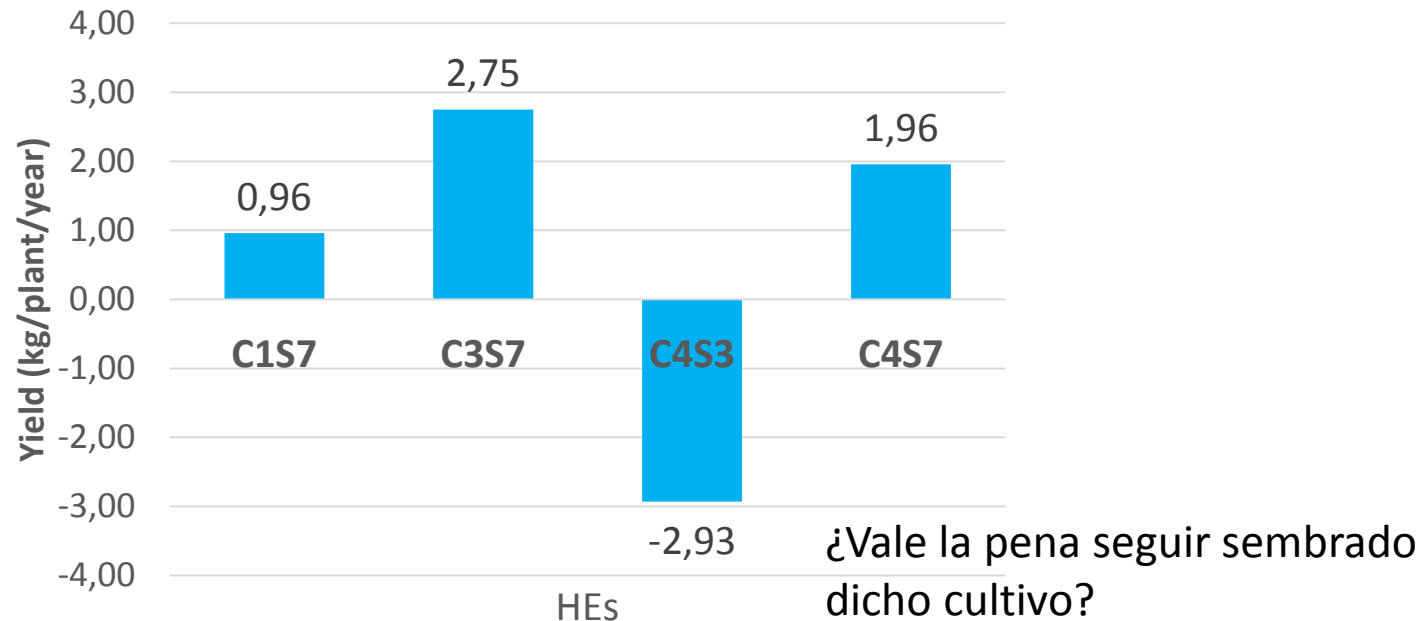
Productores con **las mismas condiciones** tenían una gran variación en rendimiento



Caso de estudio Plátano en Colombia: Clima, Suelo y Manejo agronómico

Análisis factoriales: PCA- CATPCA. Modelos mixtos y BLUPS

Cultivo asociado



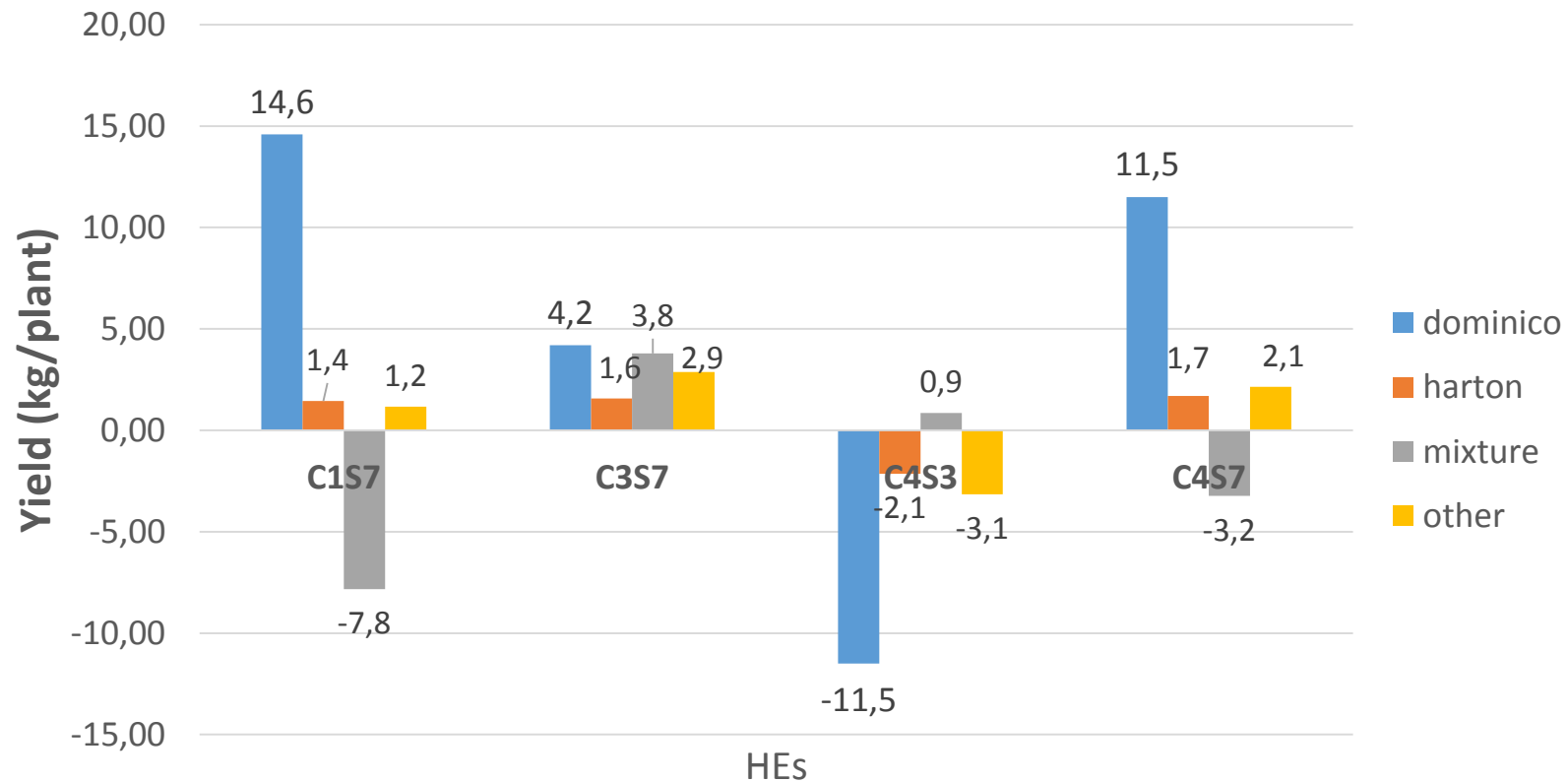
Efectos de las zonas homólogas en el modelo mixto



Caso de estudio Plátano en Colombia: Clima, Suelo y Manejo agronómico

Cultivo asociado

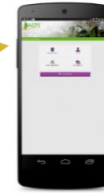
Modelos mixtos estimación de modelos mixtos



Efectos de las variedades en las diferentes zonas homólogas



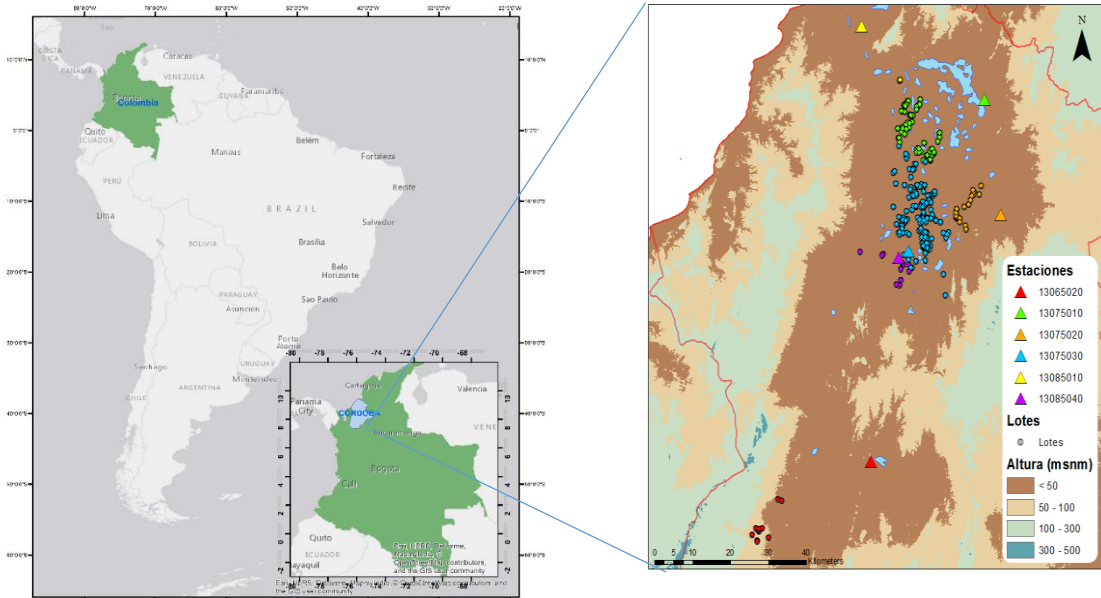
El caso de Maíz en Córdoba – Productividad- FENALCE; Cómo sembrar ?



Clima



- Temperatura máxima
- temperatura mínima
- Precipitación acumulada
- Radiación Solar
- Humedad relativa



**238 eventos ciclos de producción en dos años
Tiempo de datos (2014-2015)**

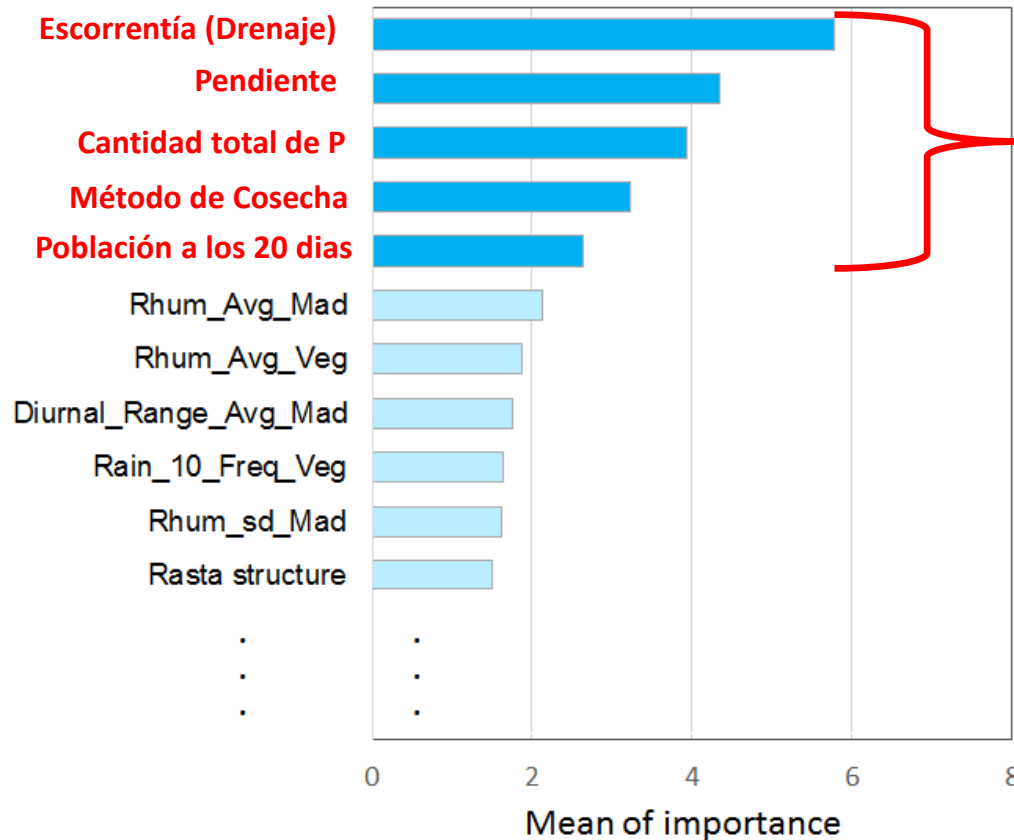




Cómo sembrar ? El caso de Maíz en Córdoba - Productividad

Arboles condicionales

$R^2 = 45.79$



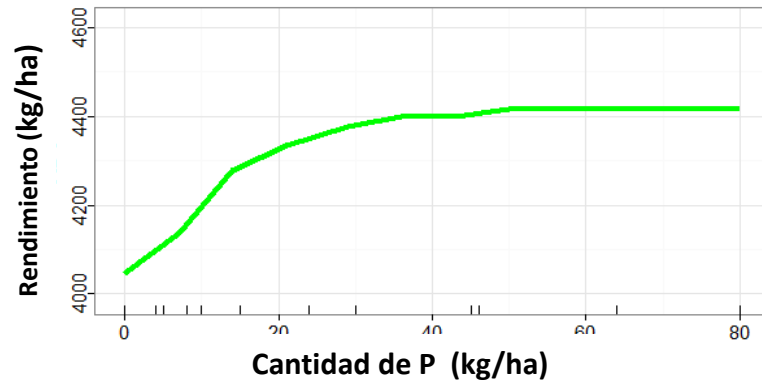
Los factores más importantes asociados con la variación en rendimiento de maíz en Córdoba



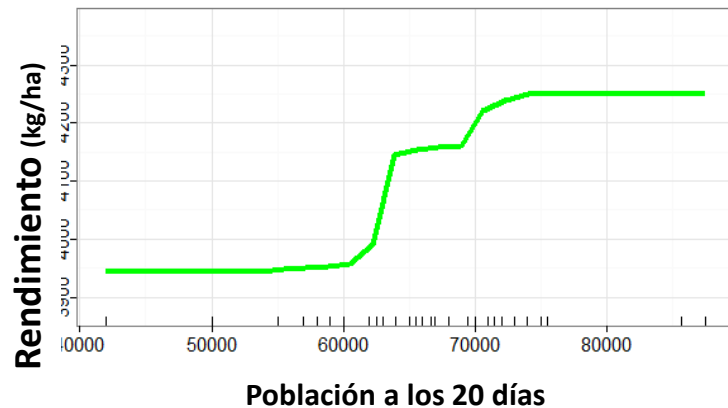
2014-2015 238



Cómo sembrar ? El caso de Maiz en Córdoba - Productividad



25 – 30 kg P /ha cantidad apropiada para maiz en Córdoba.



Población a los 20 días, al menos 65000 plantas/ha in Córdoba



Una agricultura climáticamente inteligente guiada por datos y que complementa conocimiento tradicional

El caso de Maíz Córdoba

Conocimiento tradicional – Paquete tecnológico

Profundidad efectiva suelo > 30 cm

pH 5.5-6.5

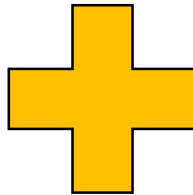
Distancia entre plantas 0.17m – 0.2m

Nutrientes requeridos N, P, K, Mg, S

Población a los 20 d: 50000 – 70000 DDE

Control arvenses: al menos 1 control (8 DAS – 2 DDS)

Control enfermedades: Al menos 1 (10 Días antes de floración)



Agricultura climáticamente inteligente guiada por datos (Big Data -minería de datos, aprendizaje automático y profundo)

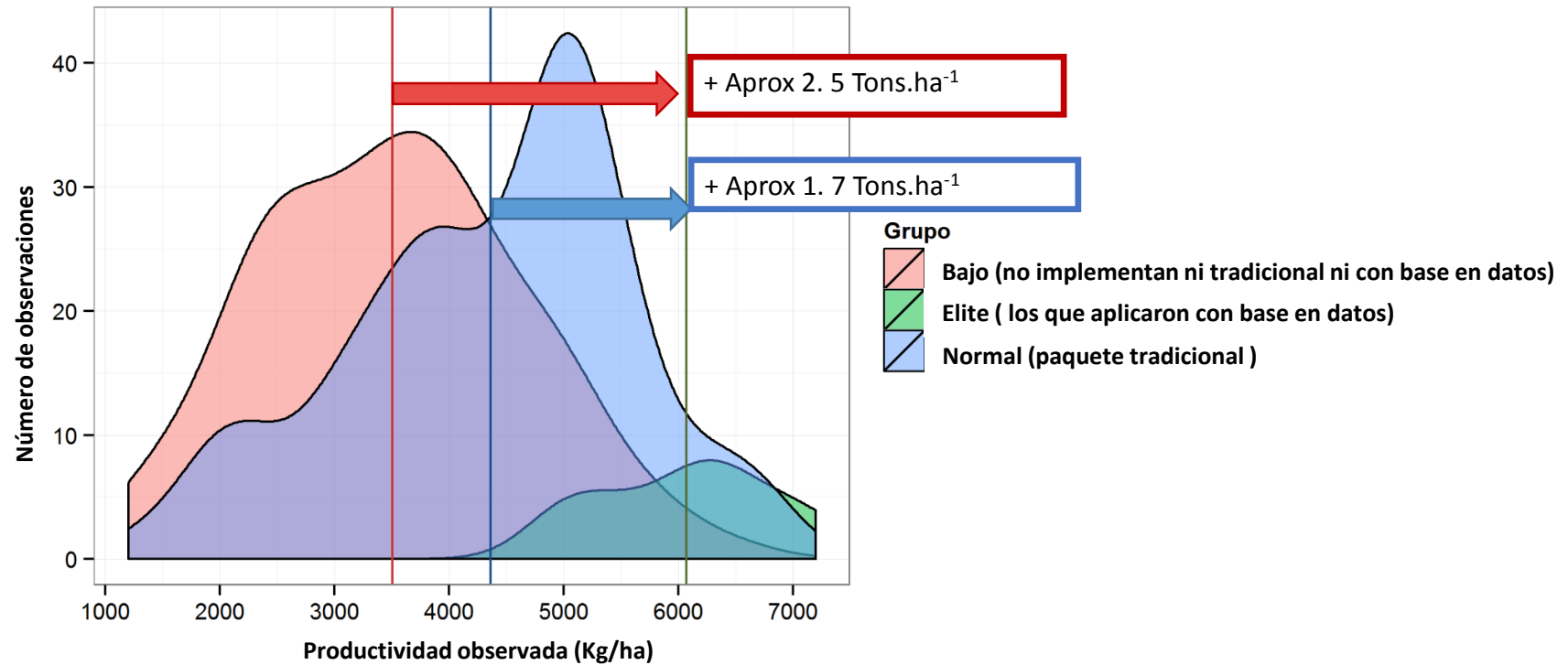
Fósforo total aplicado : 25 - 30 kg /ha

Población a los 20 d: Al menos 65000 DDE

Ambos acercamientos con base en observaciones , el Big Data revela cosas que desconocíamos



Cómo sembrar ? El caso de Maiz en Córdoba - Productividad



Distribuciones de rendimientos observados por los tres subgrupos de manejo en Córdoba- Colombia

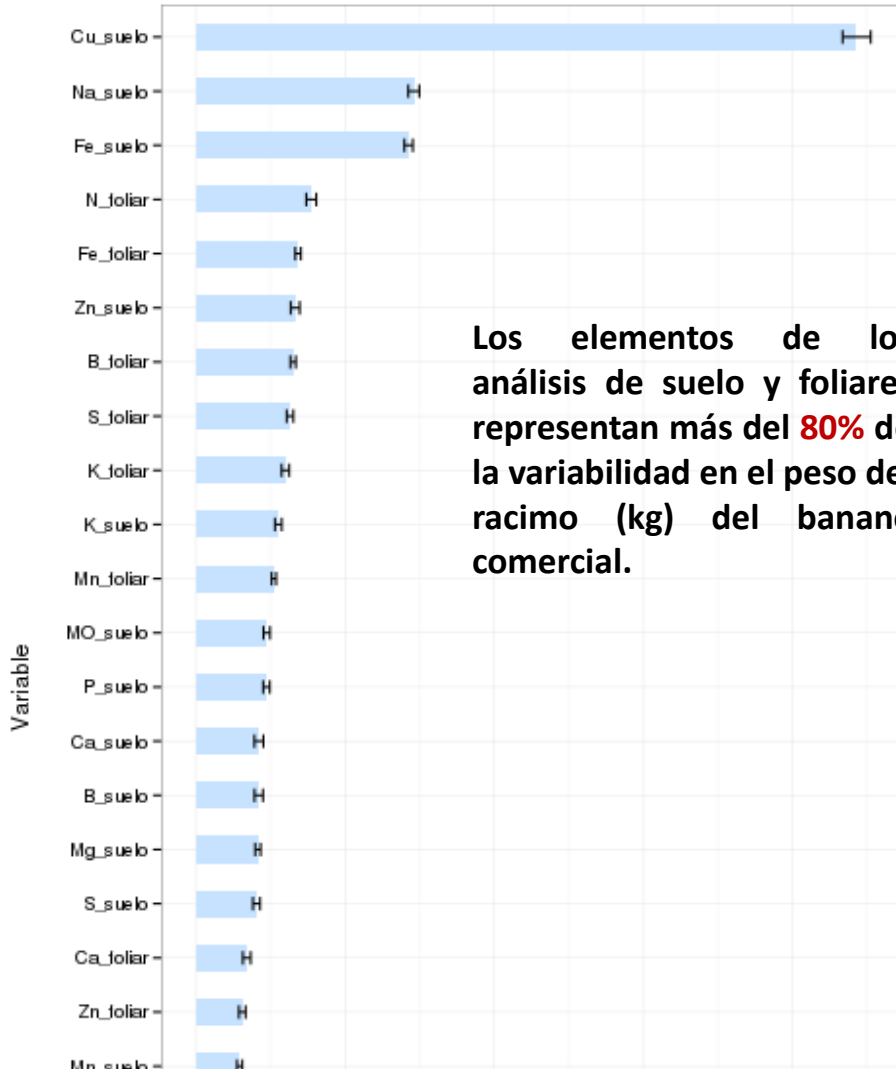
Caso de estudio: Banano en Magdalena, manejo agronómico

Random Forest

Datos: 2006 – 2010, N = 2188 datos de entrada – Banano de Exportación subgrupo Cavendish

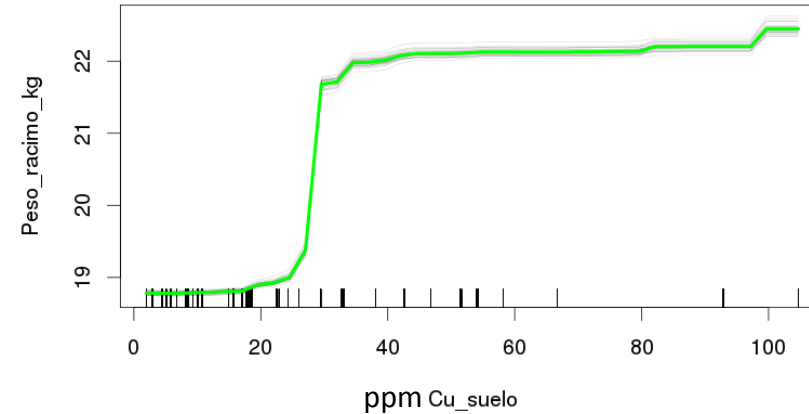


Importance of variables (with a mean R2 of 81.864 %)



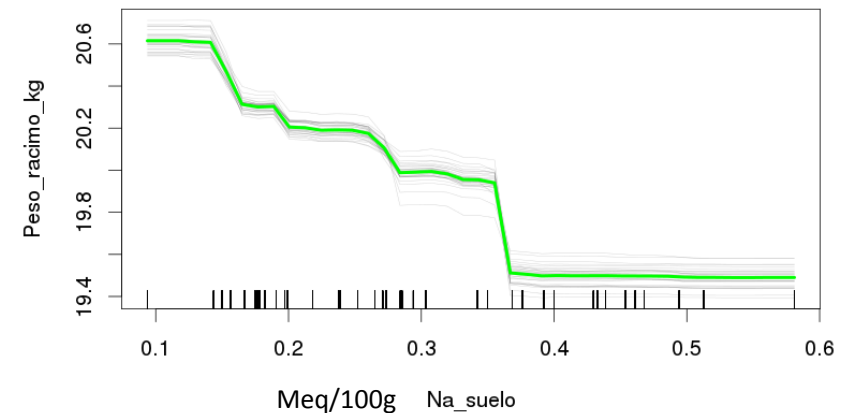
Los elementos de los análisis de suelo y foliares representan más del **80%** de la variabilidad en el peso del racimo (kg) del banano comercial.

Individual influence of Cu_suelo (with 30 profiles)



Cobre en el suelo podría ser un factor importante en el peso del racimo

Individual influence of Na_suelo (with 30 profiles)



Igualmente, **Sodio en el suelo**

Qué hemos logrado recientemente?- Socios capacitados

FENALCE

FEDEARROZ



CAMBIOS INSTITUCIONALES.

- Almacenamiento de datos a través de la plataforma AEPS (más de 2000 eventos de maíz y frijol registrados)

- Compra de un servidor
- Manejo de datos conjunto entre Área técnica y económica
- Bases de datos históricas organizadas para análisis AEPS



- 3 personas capacitadas en realizar análisis
- 30 técnicos sensibilizados en el enfoque AEPS y interpretando resultados de análisis

PERSONAL EMPODERADO.



- 6 personas capacitadas en realizar análisis
- 100 técnicos sensibilizados en el enfoque AEPS y interpretando resultados de análisis

Qué hemos logrado ?- Productores adoptando el enfoque

Reconocimiento mundial por el trabajo conjunto
MADR-CIAT-FEDEARROZ

Mejor iniciativa para
adaptación al clima

UNITED NATIONS
CLIMATE
SUMMIT 2014 | **BIG DATA**
CLIMATE
CHALLENGE
WINNERS



Qué hemos logrado ?- Publicaciones

RESEARCH ARTICLE

From Observation to Information: Data-Driven Understanding of on Farm Yield Variation

Daniel Jiménez^{1*}, Hugo Dorado¹, James Cock¹, Steven D. Prager¹, Sylvain Delerce¹, Alexandre Grillon², Mercedes Andrade Bejarano³, Hector Benavides⁴, Andy Jarvis¹

1 Decision and Policy Analysis (DAPA), International Center for Tropical Agriculture (CIAT), Cali, Colombia, **2** University of Applied Sciences of Western Switzerland (HEIG-VD), Yverdon-les-bains, Switzerland, **3** School of Statistics, Universidad del Valle, Cali, Colombia, **4** City government of Pereira, Secretary of rural development, project site-specific agriculture, Pereira, Colombia

Publish | About | Browse

OPEN ACCESS PEER-REVIEWED

RESEARCH ARTICLE

Assessing Weather-Yield Relationships in Rice at Local Scale Using Data Mining Approaches

Sylvain Delerce , Hugo Dorado, Alexandre Grillon, María Camila Rebolledo, Steven D. Prager, Victor Hugo Patiño, Gabriel Garcés Varón, Daniel Jiménez

Published: August 25, 2016 • <http://dx.doi.org/10.1371/journal.pone.0161620>

[Harvard Dataverse](#) > [CIAT - International Center for Tropical Agriculture Dataverse](#) >

Rice cropping events data for Saldaña (lowland irrigated rice) and Villavicencio

Federación Nacional de Arroceros (Fedearroz), 2015, "Rice cropping events data for Saldaña (lowland irrigated rice) and Villavicencio (rainfed rice), Colombia", [doi:10.7910/DVN/MGUTG3](https://doi.org/10.7910/DVN/MGUTG3), Harvard Dataverse, V3

<http://journals.plos.org/plosone/>



Próximos pasos!

Análisis espectrales de la vegetación a través de fotografías aéreas.

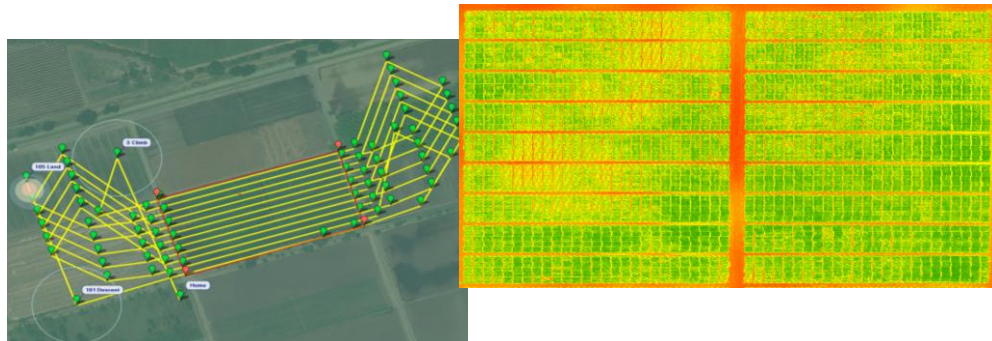
PHANTOM 3 ADVANCED



Autonomía	2:30 horas
Velocidad máxima	50 [km/h]
Altura de vuelo	150 [m]
Radiometría	10 [km]
Usos	

Autonomía
Velocidad máxima
Altura de vuelo
Radiometría
Usos

2:30 horas
50 [km/h]
150 [m]
10 [km]

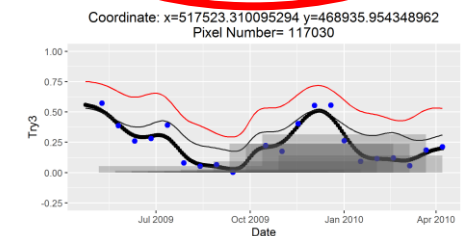
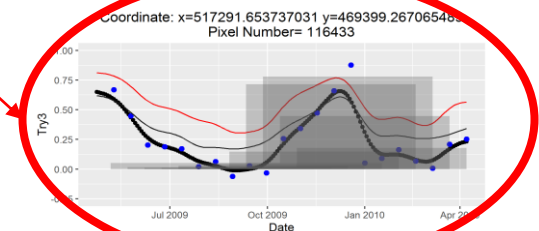
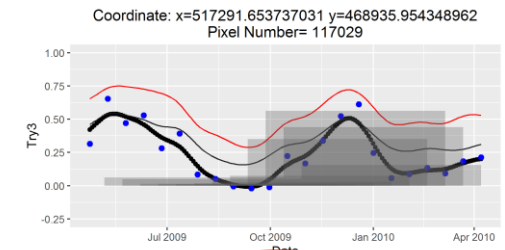
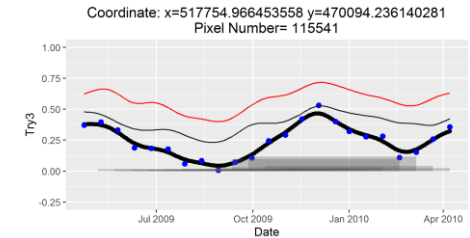


Percepción remota para identificar fechas de siembra.

Satélite: MODIS
Producto: MOD13Q1
Resolución: 250 m
Frecuencia: 16 días

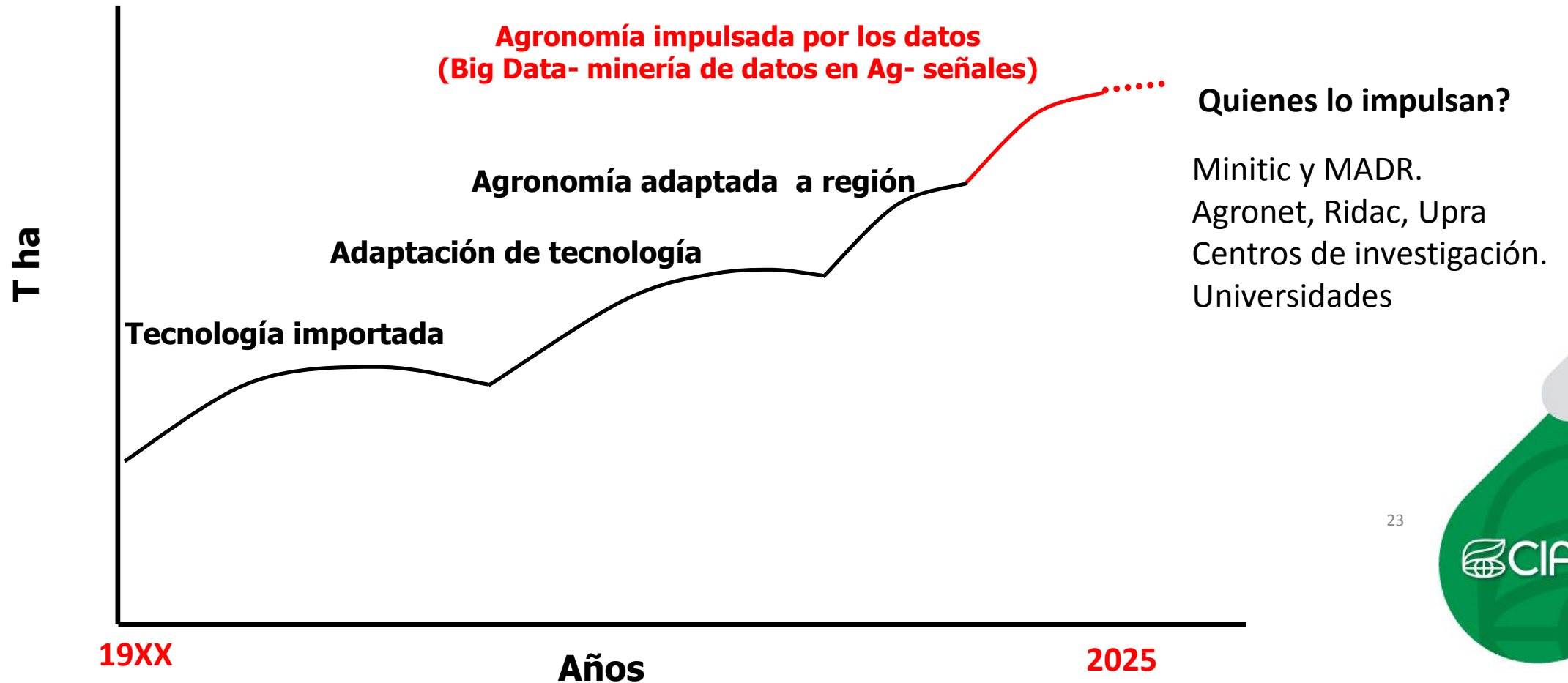


Arroz en Tolima



Oportunidades con Big Data y Agricultura climáticamente inteligente

Propuestas ciencia + otros actores



Gracias!!!



Member of the
CGIAR Consortium

www.ciat.cgiar.org
www.cgiar.org



CGIAR

Science for a food secure future