

Agricultura Especifica por Sitio

El uso de datos observacionales para optimizar la agricultura

Equipo

Daniel Jimenez, PhD, Agronomo
Luis Armando Muñoz, Msc, Biologo
Hugo Andres Dorado, Estadistico
Victor Hugo Patiño, Estadistico
Andres Aguilar, Ing. Agricola
Juan Felipe Rodriguez, Ing. de sistemas
Harold Hachicanoy, Estadistico

Participan



Apoyan



Sylvain Delerce, Msc. , 5 ago. 2015 / Jornadas ACPA 2015

Cual es la idea ?

Cada vez que un agricultor prepara un lote, siembra un cultivo, lo maneja y lo cosecha, esta **experimentando una combinación única de factores** que resulta en la producción que el registro.

Si somos capaces de caracterizar con suficiente detalle las condiciones de suelo y de clima en las que creció un cultivo, y que tenemos el registro del manejo que se le dio así como de la producción que genero...

... entonces cada cultivo representa un experimento del cual podemos aprender.

Al caracterizar **muchos cultivos**, podemos agregar el conocimiento capturado en condiciones muy diversas y **conocer la respuesta del cultivo a las variaciones del ambiente así como del manejo.**

VER:

Cock, J., Oberthür, T., Isaacs, C., Läderach, P. R., Palma, A., Carbonell, J., ... Anderson, E. (2011). Crop management based on field observations: Case studies in sugarcane and coffee. *Agricultural Systems*, 104(9), 755–769. doi:10.1016/j.agsy.2011.07.001



Agricultura Especifica por Sitio Compartiendo experiencias

Principios



Cultura de medición : lo que no se mide, no se puede manejar, y tampoco compartir o comparar.

Tomar registros permite aprender sobre sus propias experiencias y sobre las de los demás



El **conocimiento colectivo** es mucho mas poderoso que un conocimiento individual: permite multiplicar las observaciones y experimentos (= parcelas experimentales)



Hoy las **Tecnologías de Información y Comunicación (TIC)** permiten compartir, analizar información de forma muy simple e inmediata

Definición

La agricultura específica por sitio:

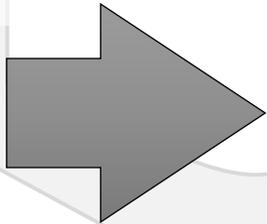
“se define como el arte de realizar las prácticas agronómicas requeridas por una especie vegetal, de acuerdo con las condiciones espaciales y temporales del sitio donde se cultiva, para obtener de ella su rendimiento potencia”

Cenicaña, 2004

El encuentro con Big Data

Unos principios de lo que es BIG DATA:

1. Todo el día, todos generamos muchos datos, **estamos “datificando” todo el tiempo** (smartphones, tarjeta de crédito, consultas internet...). Y no son datos de experimentos controlados.
2. Dejar de trabajar con muestras pequeñas para acercarse del **N=All**
3. **Relacionar BDD** de propósitos diferentes puede permitir descubrir patrones, relaciones desconocidas
4. A veces es suficiente **constatar relaciones mas no explicar** las causas de un fenómeno
5. **La cantidad de datos puede compensar la calidad** variable



- Aprender de los datos existentes y reales (no experimentales)
- Usar métodos estadísticos y computacionales para “leer” en grandes cantidades de datos
- Combinar información del pasado con información en tiempo real

VER:

Mayer-Schonberger, V., & Cukier, K. (2013). *Big Data: A Revolution That Will Transform How We Live, Work and Think.*

Tomar decisiones en agricultura

MERCADO
EXPERIENCIA DE OTROS
Tradiciones
MODA LIBROS
Incentivos



AEPS, Big Data



Agronomía soportada
por datos

Conocimiento del cultivo
Aprender de las
experiencias pasadas

Herramientas de pronóstico
Anticipar lo que viene

Con que información trabajamos ?

Datos diarios de estaciones climáticas
BDD existentes a nivel mas grueso (Worldclim, TRMM...)

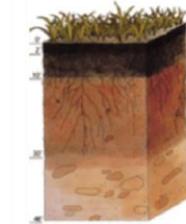
Datos de las fincas : RASTA, análisis de suelo, BDD
existentes : IGAC, Digital Soil Mapping

Datos capturados en la finca : practicas
registradas desde la siembra hasta la cosecha



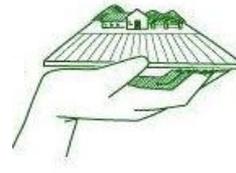
Clima % ?

+



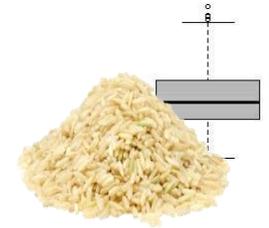
Suelo %?

+



Manejo agronómico % ?

=



Rdto/ha (100%)

Factores no controlables

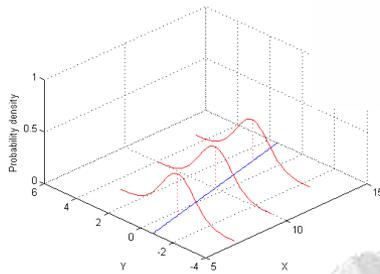
Factores
controlables

Documentar de la mejor forma posible cada componente de la ecuación

Que métodos de análisis usamos ?

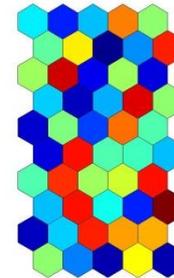
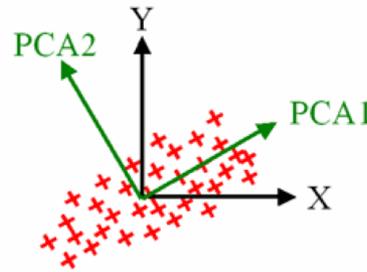
Visualizaciones y técnicas exploratorias:

- Matriz de correlación
- Scatter plot
- Boxplot
- Clustering
- Análisis factoriales



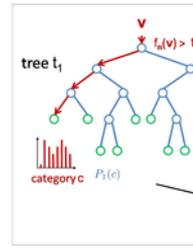
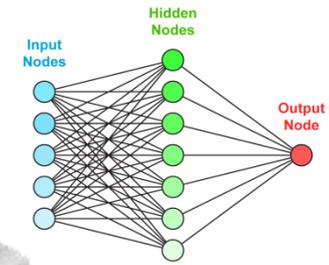
Modelos simples :

- Regresión lineal múltiple
- Modelo lineal generalizado
- Modelos mixtos



Técnicas de aprendizaje automático (Machine learning):

- Redes neuronales artificiales
- Random Forest
- Conditional Forest
- Mapas autorganizados de Kohonen



Esquema de trabajo

Fomentar la cultura de toma de datos, y facilitar la tarea usando TICS:

- involucrar al agricultor
- Aprovechar las nuevas herramientas TIC (*Aplicativo Android en GooglePlay, buscar por CIAT*)
- Descentralizar la toma de datos para lograr mayores cantidades



Modernizar el almacenamiento de información para un acceso más eficiente:

- Protocolos de limpieza y estandarización de los datos
- Motor de BDD libres tipo SQL
- Almacenamiento en la nube

Analizar los datos con nuevos enfoques

- Combinar varias metodologías, multivariados, no lineales, no paramétricos, machine learning...
- Articular modelos estadísticos con conocimiento experto para revelar patrones/tendencias en los datos
- Identificar factores limitantes de forma sitio-específica
- Combinar resultados con pronósticos climáticos para ayudar a la toma de decisión para lo que viene



Entregar información procesada a los usuarios finales de forma amigable

- Buscar la forma más adecuada de retroalimentar nuestros usuarios con resultados de análisis, pero presentados de forma intuitiva
- Involucrar los usuarios en el diseño de los aplicativos
- Trabajo de socialización en grupos
- Usar herramientas web para disponibilidad permanente



Plataforma web AEPS:
<http://www.open-aeps.org:8080/>

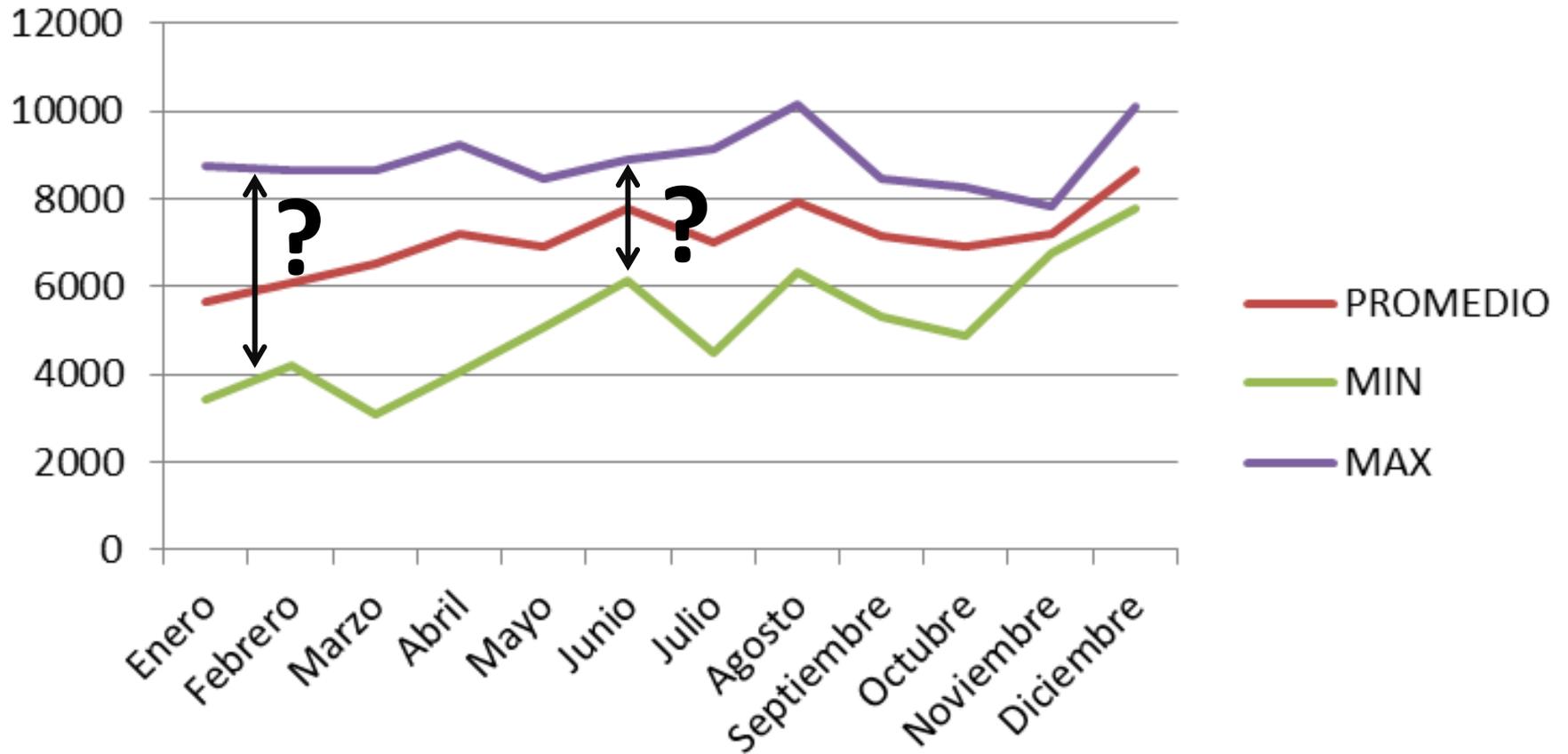


Aplicaciones

1. Detección de factores limitantes del rendimiento

Caso arroz colombia clima

Rendimientos 2011 - Tolima



Ejemplo : caso arroz Colombia

Datos del cultivo

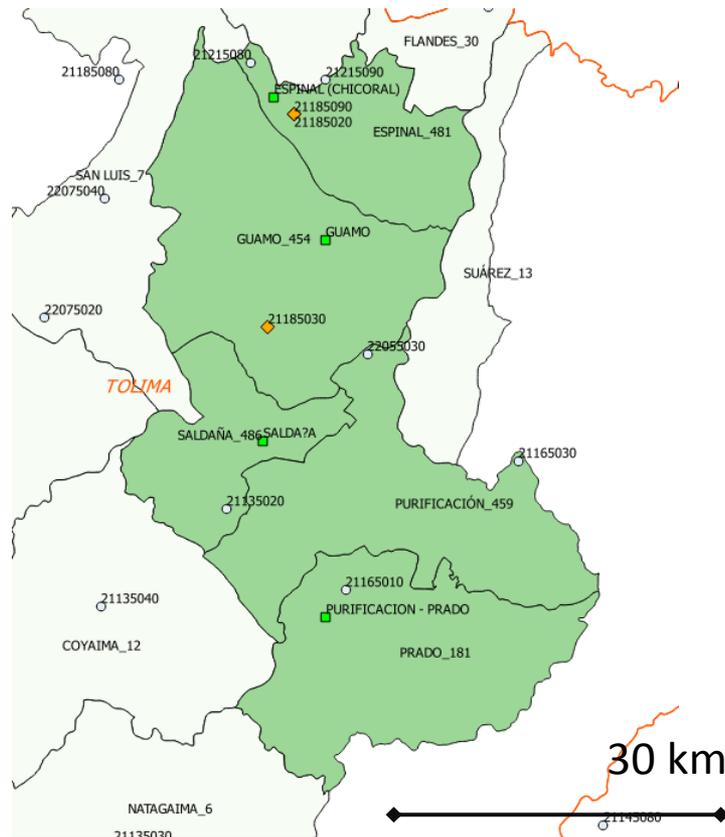
Bases de datos usadas	Propósito inicial	N
Encuesta Nacional Arroceras	Monitorear la eficiencia económica del sector	1237
Registros de cosecha	Monitorear los niveles de producción en regiones	6000
Ensayos época de siembra	Estudio técnico sobre fechas de siembra	500

Series de datos diarios de clima

Red de estaciones de IDEAM



Red agroclimática de FEDEARROZ



Datos de clima

Preparación de los datos de clima

Control de calidad de las series

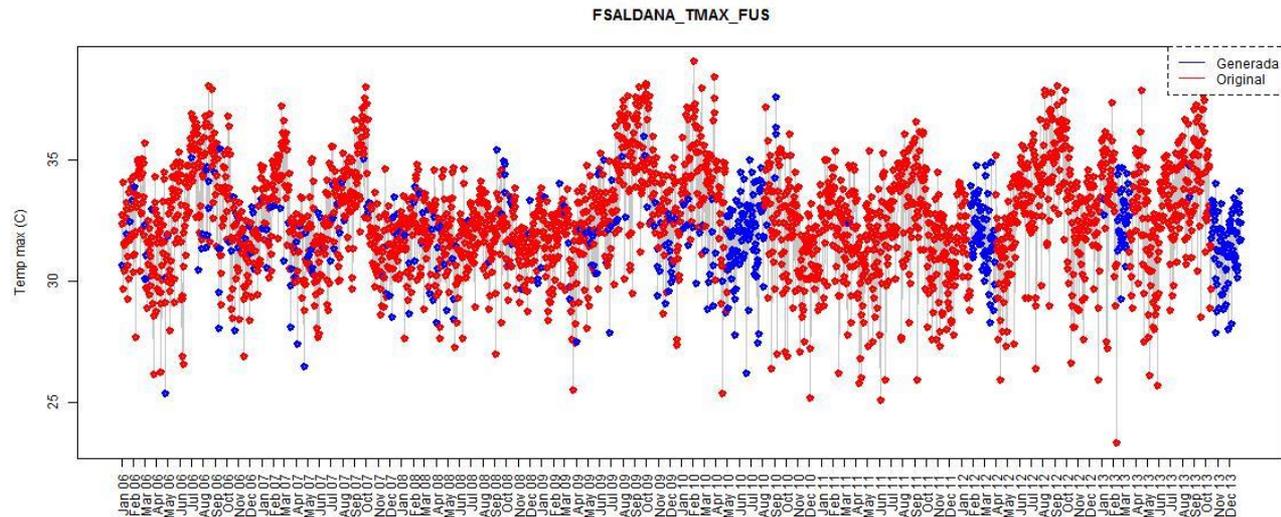
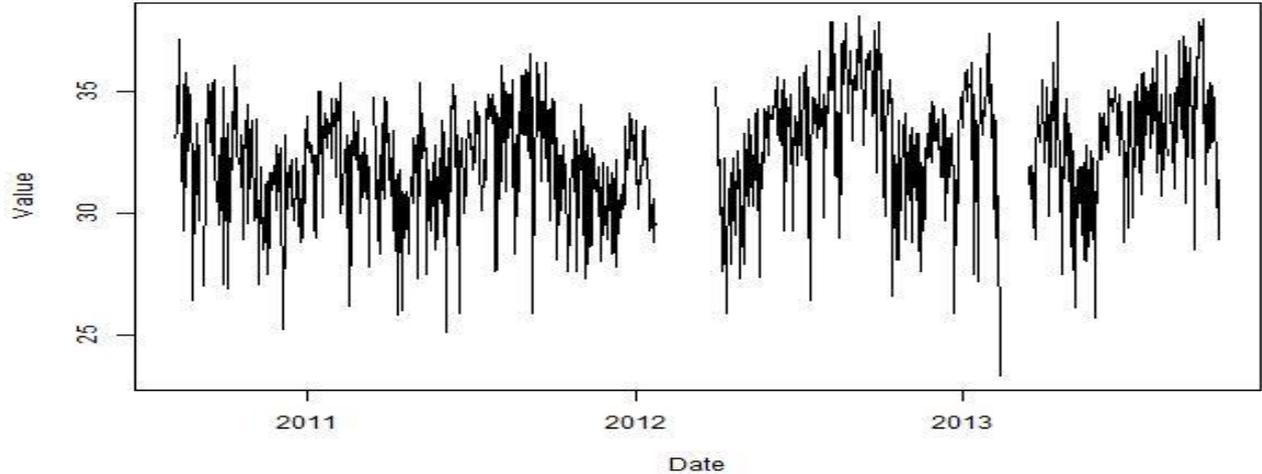


Homogeneización a escala diaria

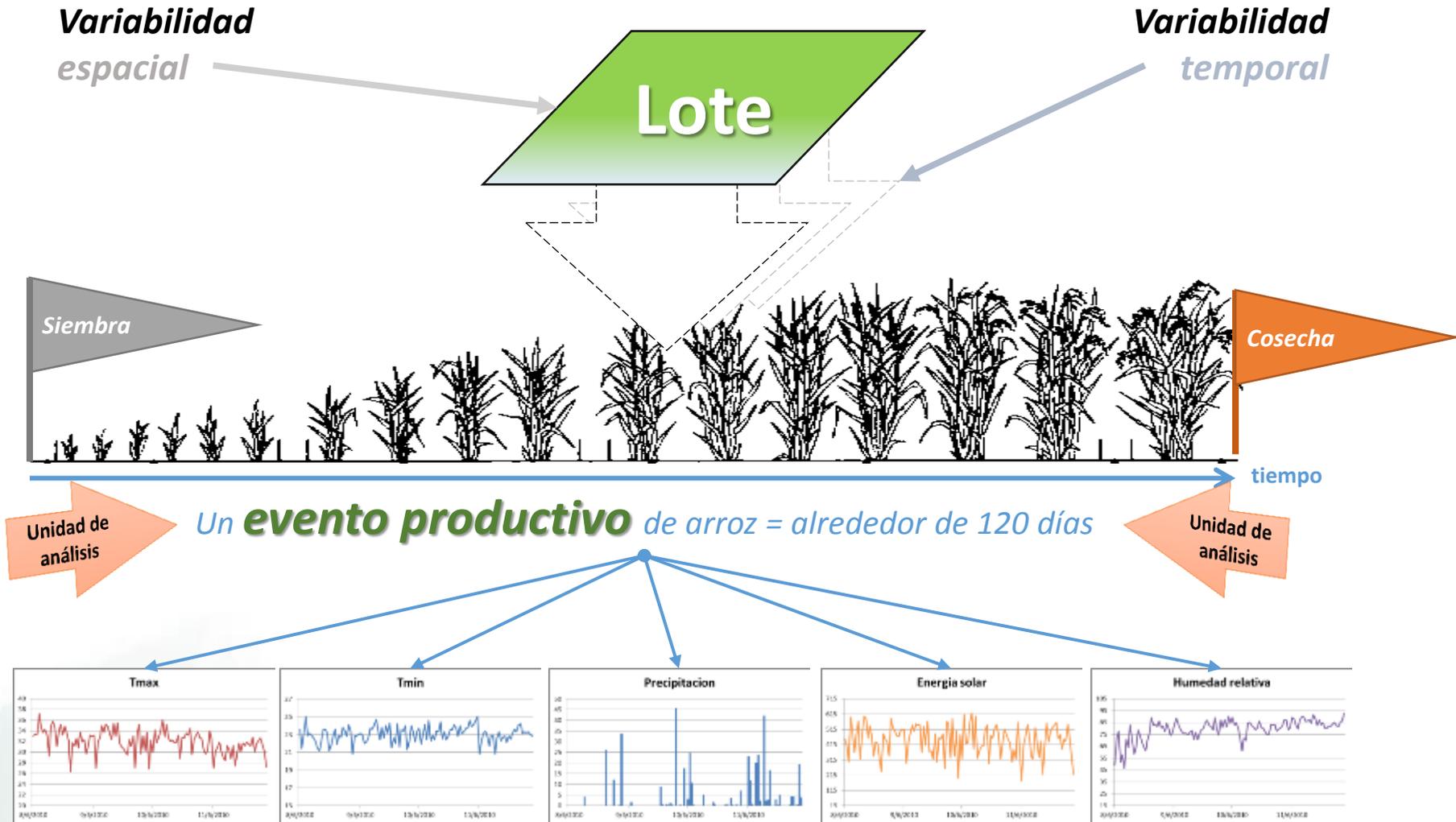


Estimación de faltantes para relleno

RMAWGEN
Random Forest

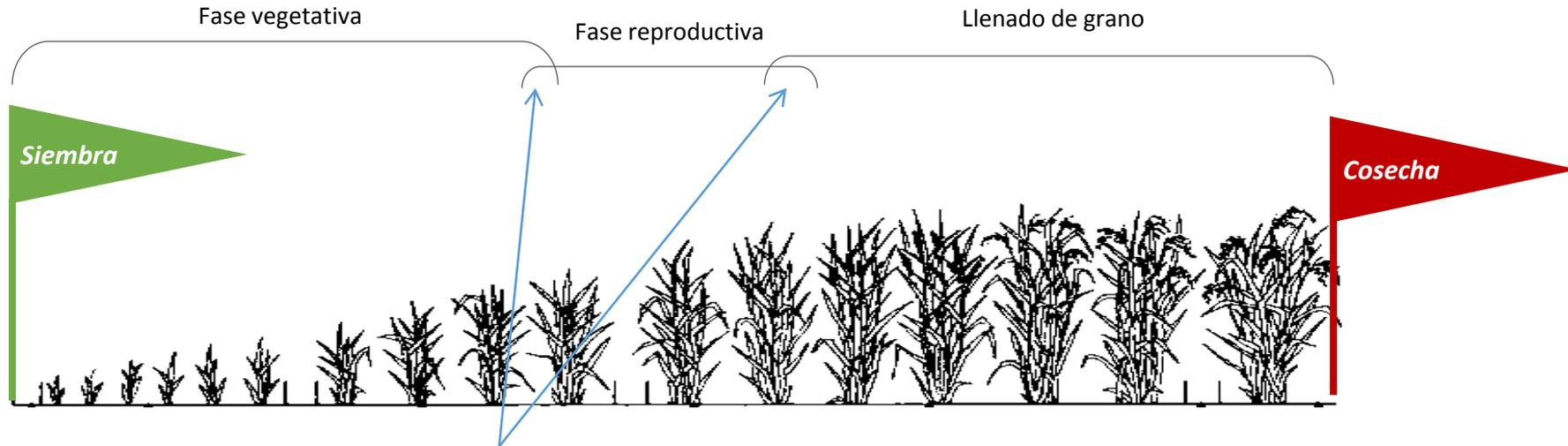


Unidad de análisis



Serie climática completa para 5 variables

Análisis por etapa fisiológica



Pueden haber diferencias entre los materiales !

Fase vegetativa

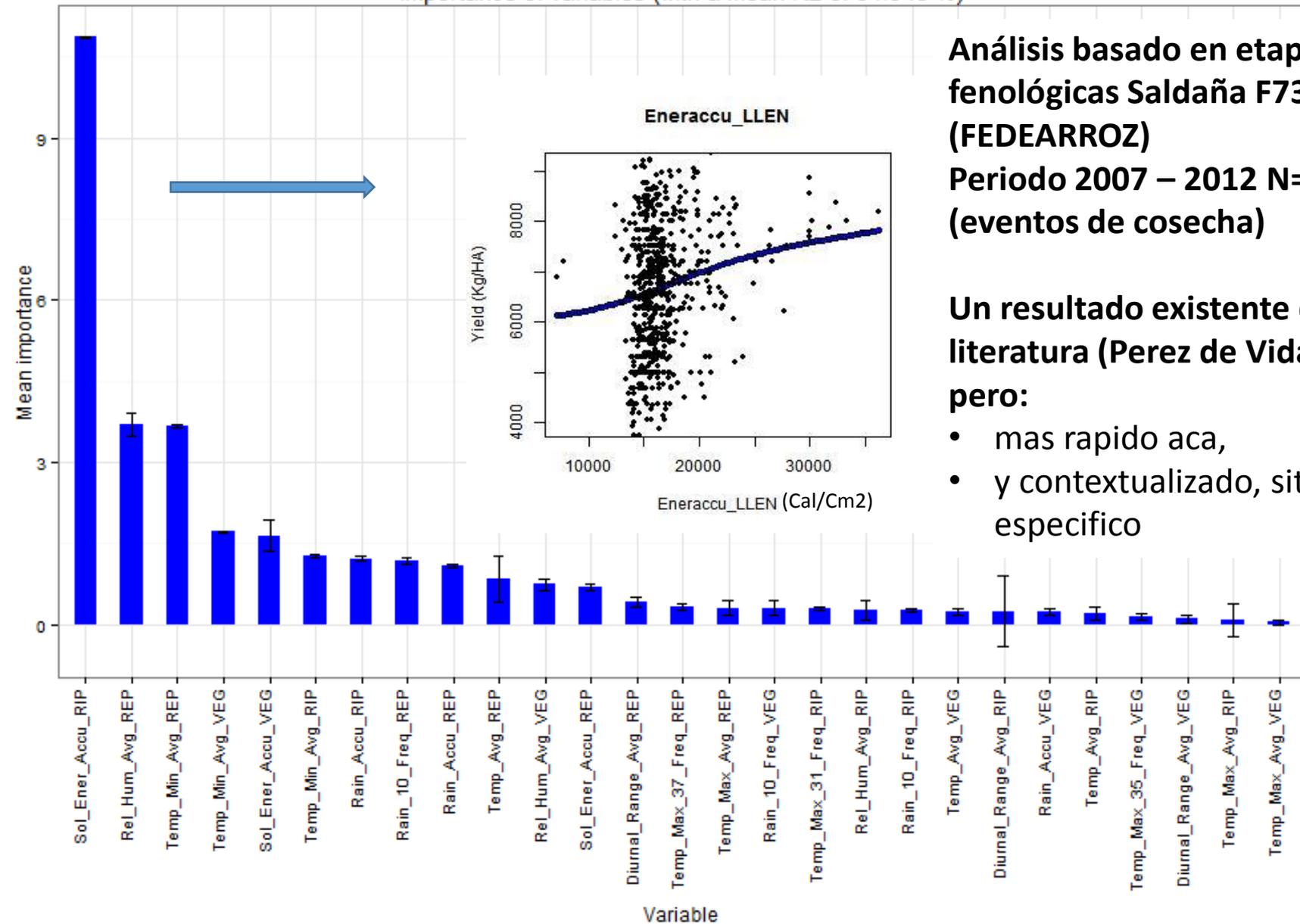
Reproductiva

Llenado grano

Var	Significancia	Unidad
Tmax	Tmax de la ventana	°C
Tmin	Tmin de la ventana	°C
TEMPmax34freq	Frecuencia de días con Tmax > 34°C	Adimensional
TEMPavg	Temperatura media de la ventana	°C
GDaccu11	GD acumulados en la ventana (base 11)	°C
Eneraccu	Energía acumulada en la ventana	Cal.cm ⁻²
VPD_avg	VPD media de la ventana	Pa
RHUM_avg	Humedad relativa media de la ventana	%
RANGO_Diurno	Rango diurno medio de la ventana	°C
RANGO_Diurno10freq	Frecuencia de días con rangos < 10°C	Adimensional

Identificar factores limitantes por fases

Importance of variables (with a mean R2 of 34.948 %)



Análisis basado en etapas fenológicas Saldaña F733 (FEDEARROZ)

Periodo 2007 – 2012 N= 267 (eventos de cosecha)

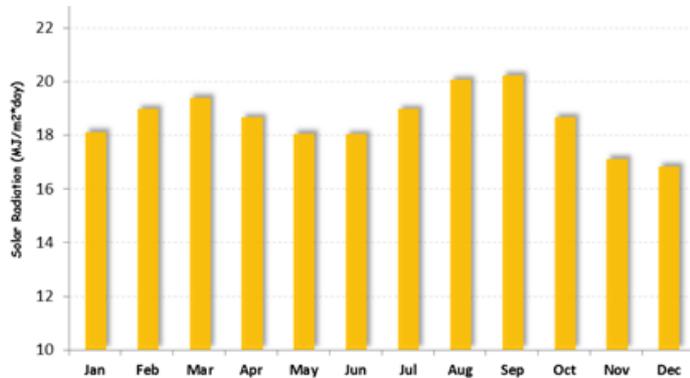
Un resultado existente en la literatura (Perez de Vida 2013), pero:

- mas rapido aca,
- y contextualizado, sitio especifico

Del resultado al impacto

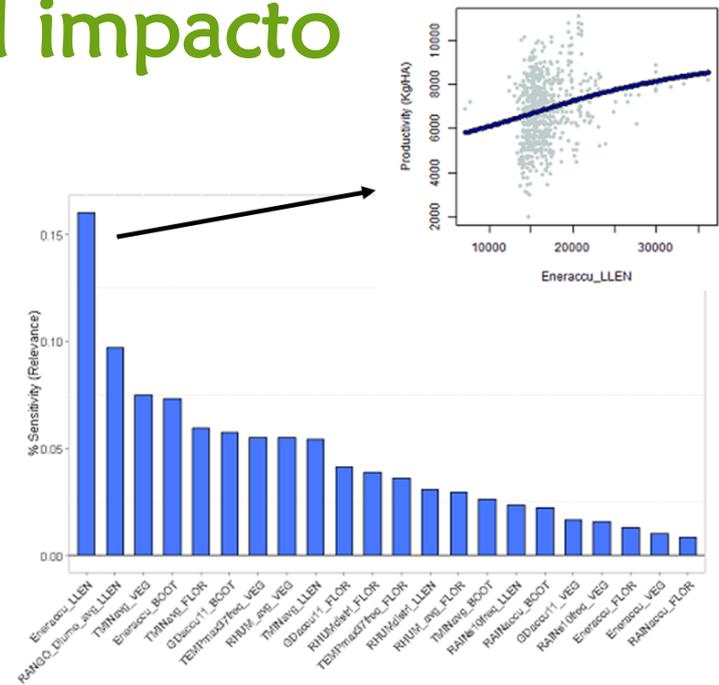
1. La radiación solar en fase de llenado de grano es el factor limitante para F733 en Saldaña (Análisis de redes neuronales)

Variación interanual de la disponibilidad de radiación solar en Saldaña (Tolima)



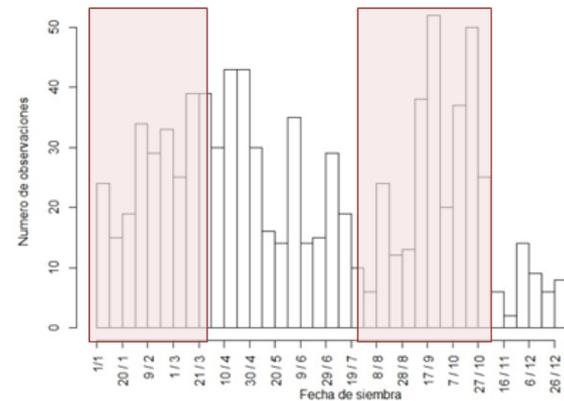
3. Por ahora se siembra arroz todo el año en esta localidad:

→ Ajustar las fechas de siembra para aprovechar los dos picos de radiación



2. Hay dos periodos de máxima oferta en el año

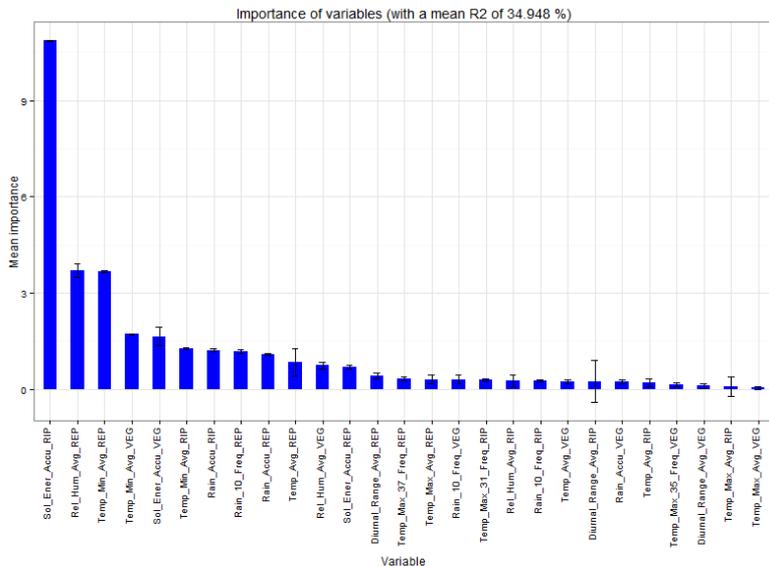
Distribución de las siembra en Sadaña, a lo largo del año



Respuesta diferencial entre materiales

Análisis basado en etapas fenológicas Saldaña - RIEGO (FEDEARROZ) – **Periodo 2007 – 2012**

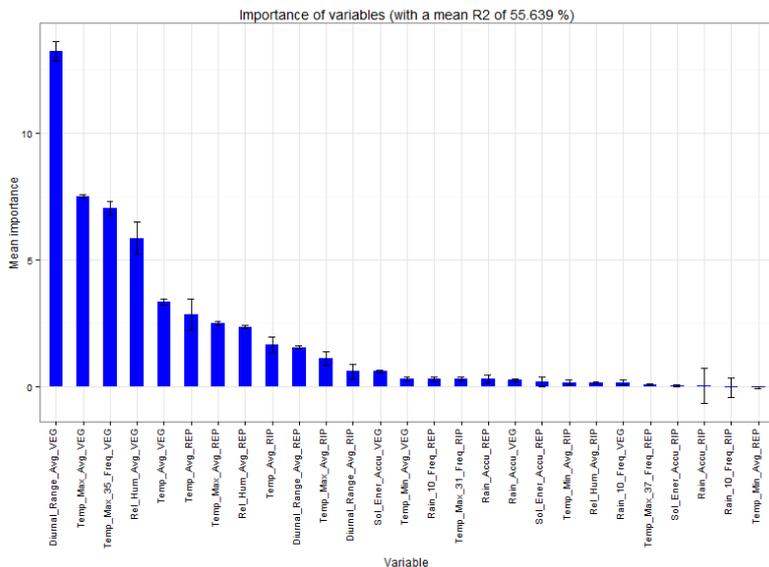
Fedearroz 733
Cforest 2014 (N=267)



Para **FEDEARROZ 733**, el clima explica el **34.9%** del rendimiento. Las variables que mas explican son de las **fases REP y RIP: energía solar mas que todo**

Factores diferentes!
Las variedades responden de manera diferente al clima!

Cimarron Barinas
Cforest 2014 (N=78)

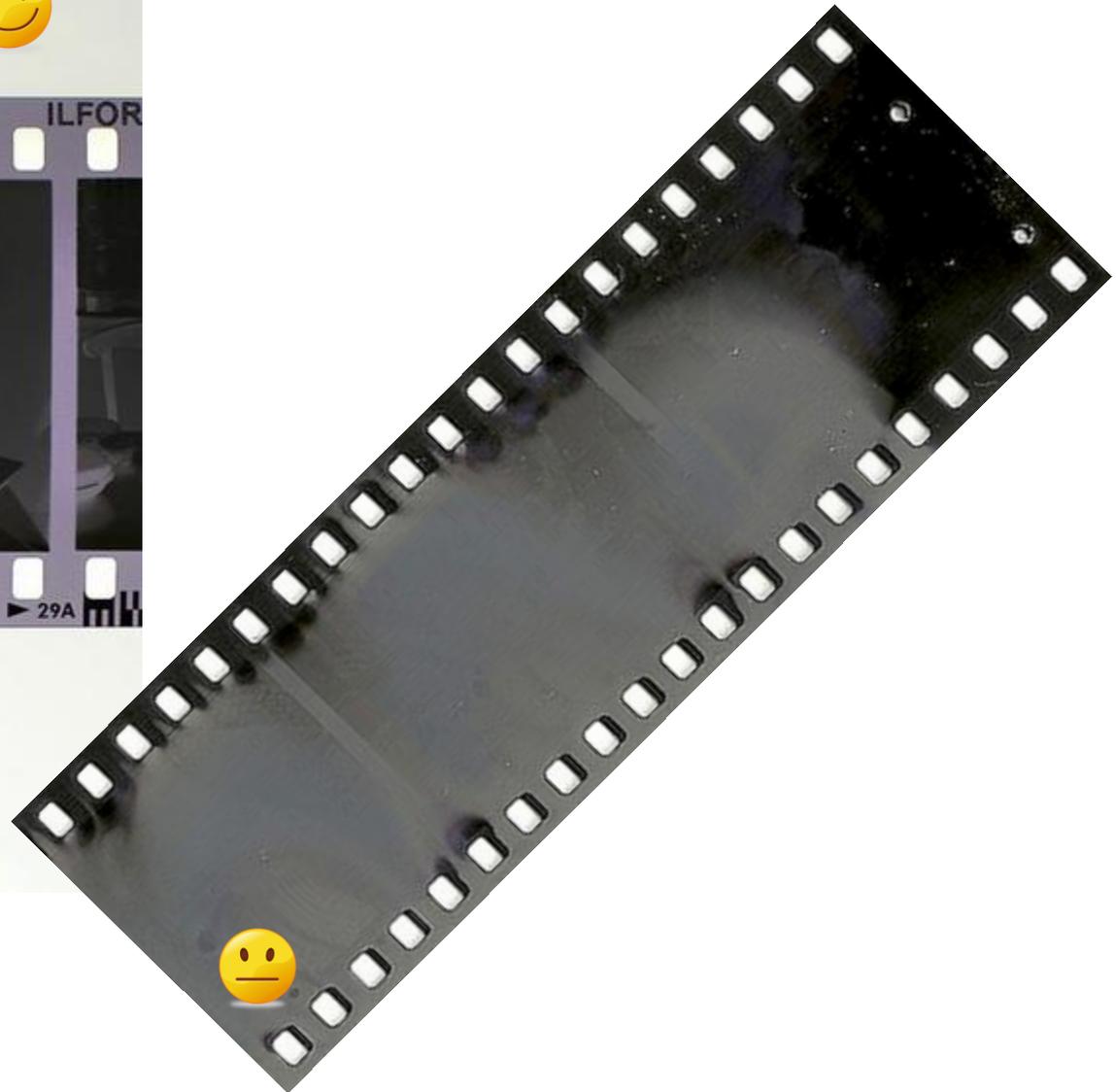


Para **Cimarrón Barinas**, el clima explica el **55.6%** del rendimiento. Las variables que mas explican son de la **fase vegetativa** y tienen que ver con la **amplitud térmica y la temperaturas máximas**.



Repuesta al clima de cada material en cada región = insumos para Fito mejoradores + ayuda a la elección para los agricultores

Detectar señales...

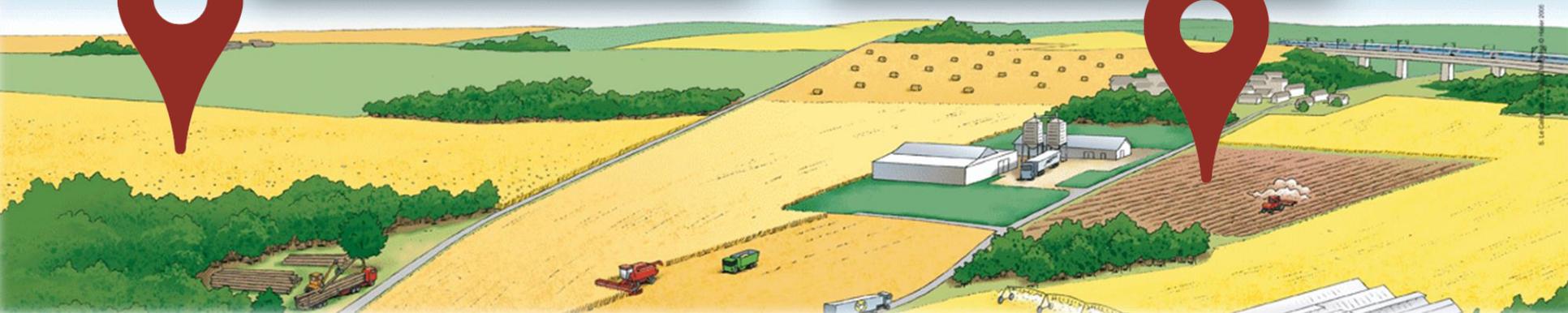
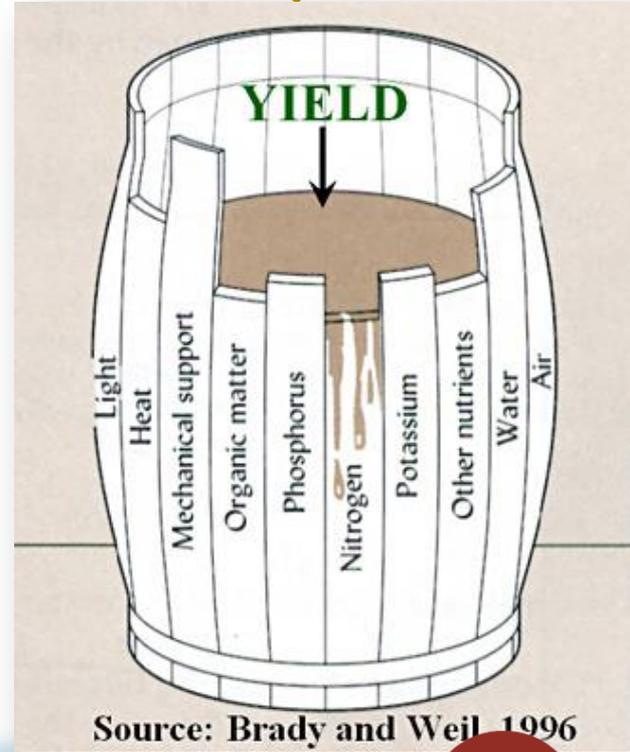
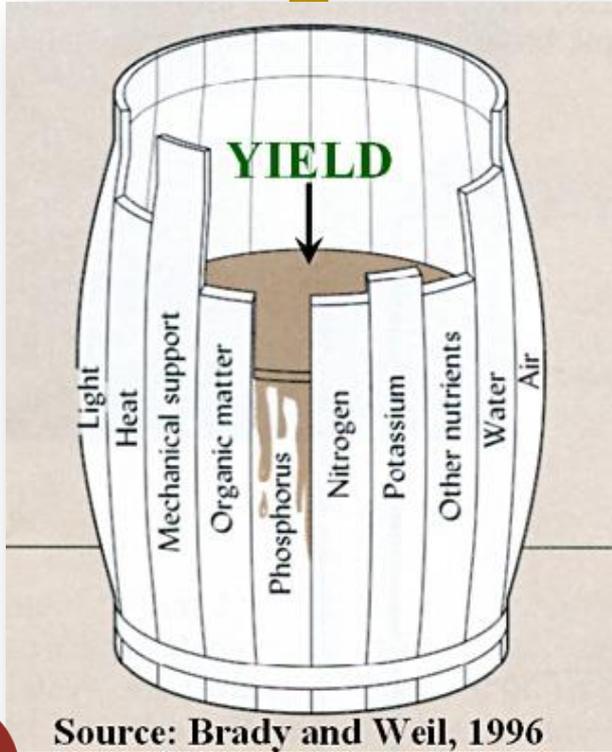


...cuando la

Buscando los factores limitantes...

...Nada Nuevo, todo por hacer...

Mejora en el manejo





Aplicaciones

2. Identificación de secuencias climáticas favorables/desfavorables

Agrupamiento según condiciones climáticas

Ejemplo : Espinal, Riego



Hierarchical clustering

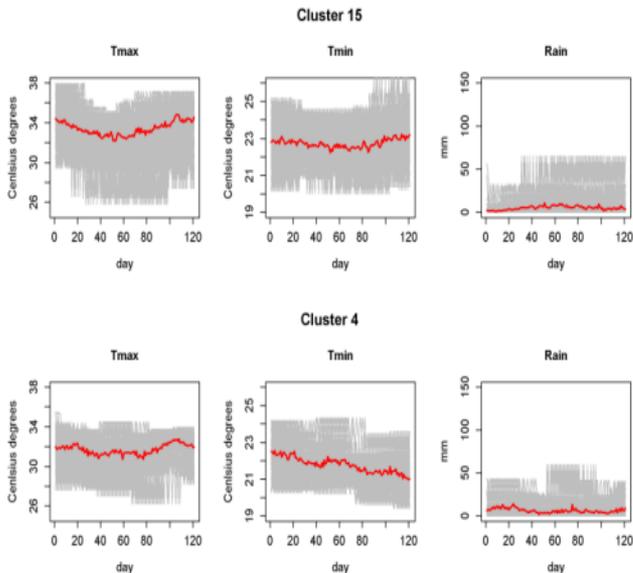
Análisis de agrupamiento de los eventos de cosecha basado en las secuencias de las 5 variables de clima que los caracterizan (Tmin, Tmax, Prec, Solar radiation, humedad relativa)

38 clusters

Cluster	Numero de eventos
15	139
17	136
4	116
20	116
21	81
24	63
11	44
16	44
10	42
26	42
9	41
18	34
23	32
31	32
12	30
3	27
5	22
38	22
22	21
13	20
25	20
32, 30, 37, 14, 29, 19, 6, 2, 27, 1, 8, 33, 35, 28, 7, 34, 36	<20
Total	1239

Cada cluster representa una secuencia climática de 120 días que ocurrió en el pasado en muchos cultivos de arroz.

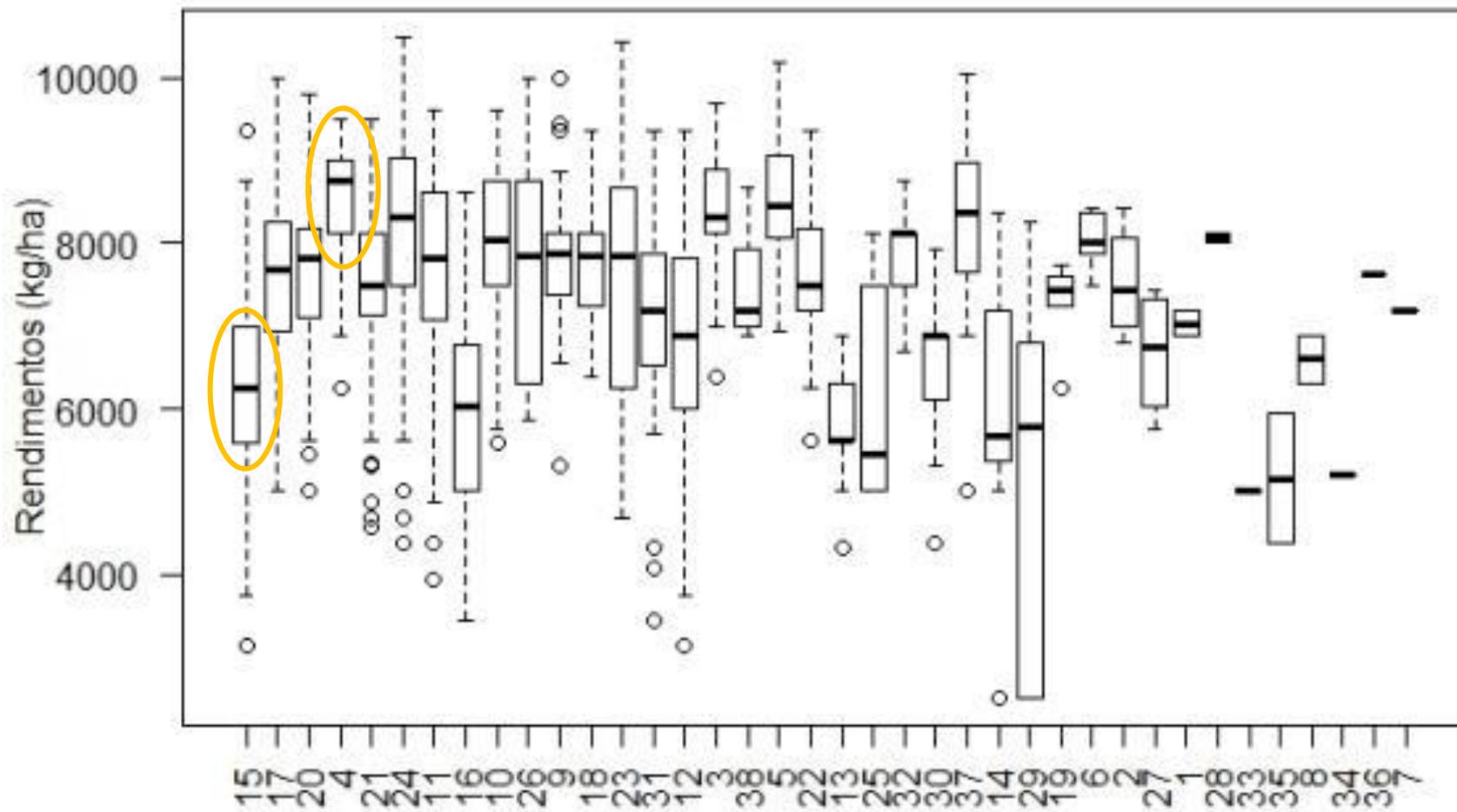
Si comparamos los clusters con un pronóstico, podemos aprender de las experiencias pasadas que se asemejen al pronóstico.



Visualización de los clusters de clima. En gris, todos los eventos individuales que pertenecen al cluster, en rojo, el perfil del centroide del cluster

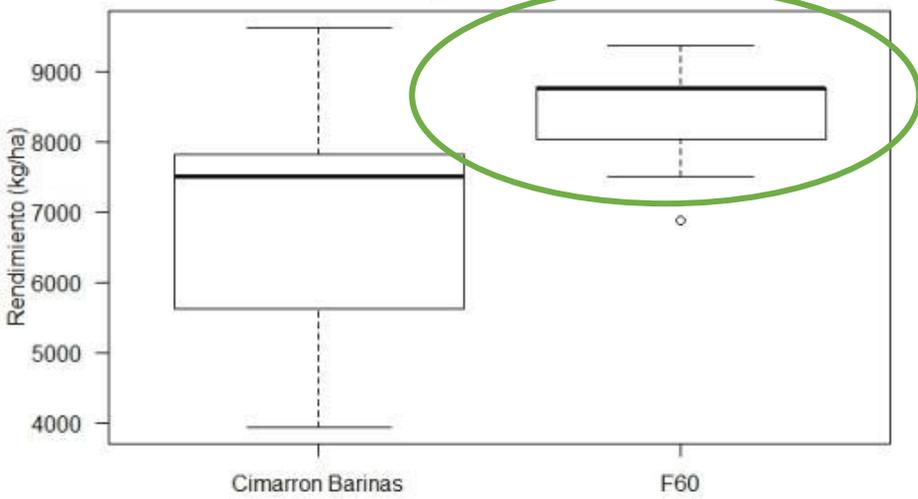
Secuencias climáticas favorables/desfavorables

Analisis Espinal-Guamo 05/2014
Variabilidad de rendimientos por cluster de clima

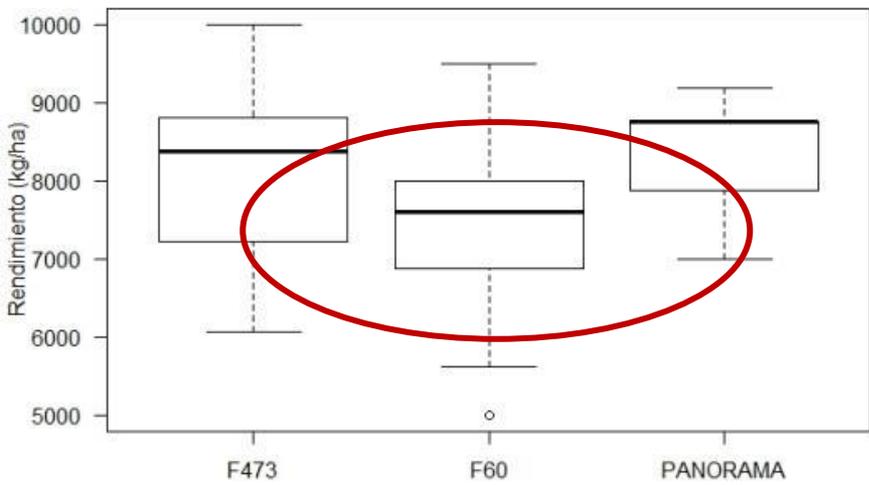


En cada secuencias, practicas óptimales

Distribucion de rendimientos por material sembrado, cluster 11
Análisis Espinal-Guamo 05/2014



Distribucion de rendimientos por material sembrado, cluster 17
Análisis Espinal-Guamo 05/2014



Preparando para el futuro : La relevancia de las redes de estaciones agroclimáticas

Todo se basa en los datos, y la precisión de los mismos. Mas estaciones, es mas exactitud a la hora de caracterizar el clima que experimentaron los cultivos.



Rain80Acum.html

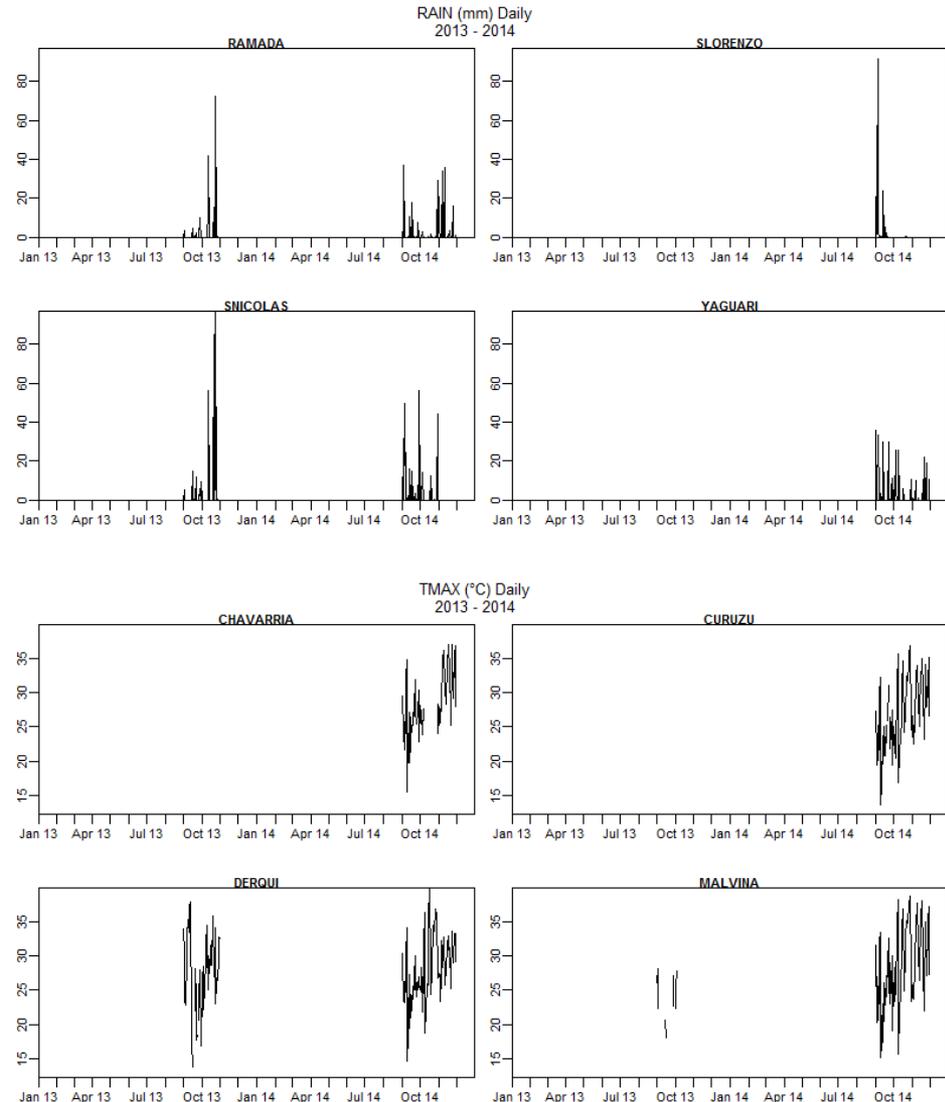


tmin2014.html



Una red en proceso de estabilización

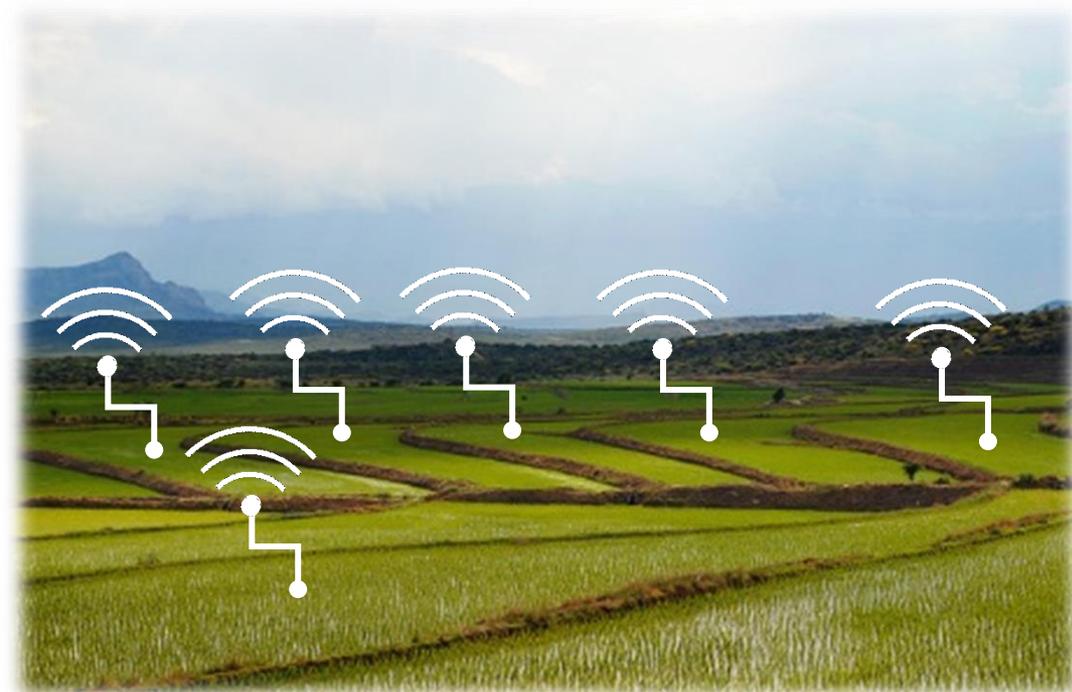
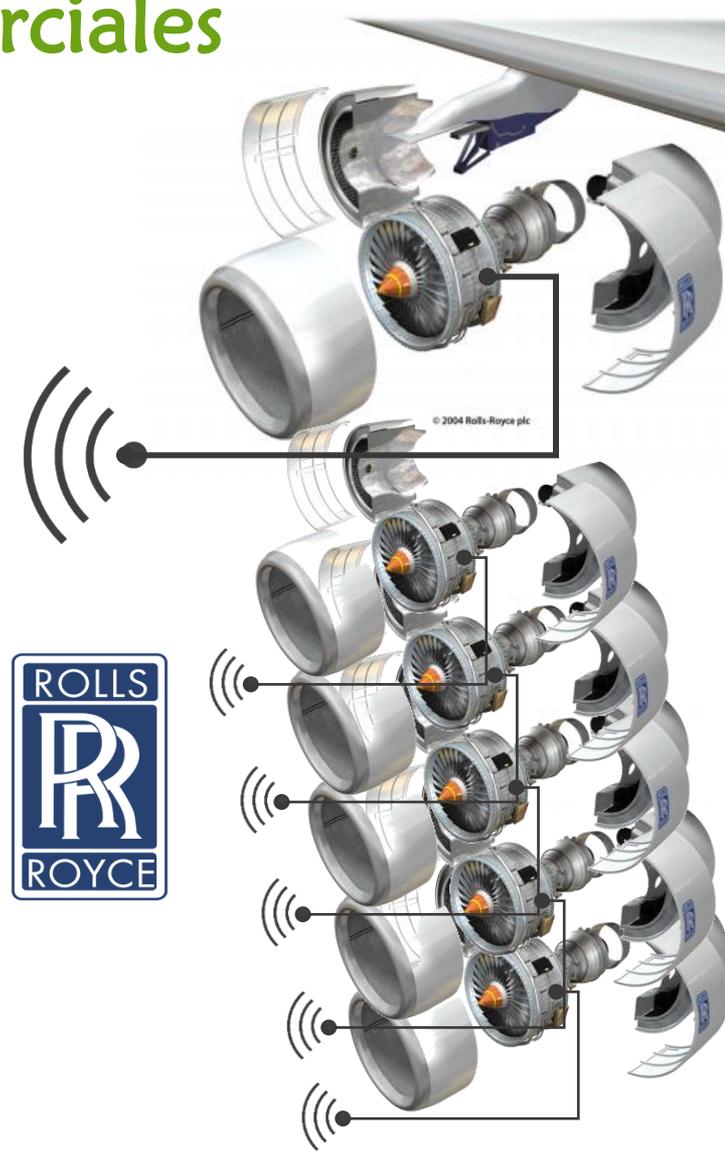
Todavía muchos datos faltantes.
Lo mas difícil es mantener una red en funcionamiento.
La conexión remota de los equipos de hoy ayudan bastante: se reportan.





Perspectivas

Monitoreo continuo de desempeño de los materiales en condiciones comerciales





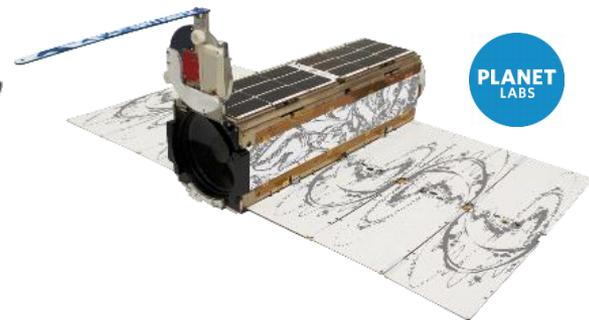
Mejorar sistemas de captura de datos



<https://play.google.com/store/apps/details?id=com.aepsmovil.aepsmovil>

<http://www.open-aeps.org:8080/>

Las maquinas lo hacen mejor...



<https://www.planet.com/>

Noticias del mundo

AIG KNOWLEDGE & INSIGHTS CENTER
Explore AIG's new Knowledge & Insights Center.
FIND OUT WHAT'S NEXT IN RISK MANAGEMENT

Want to go ad free?

ENERGY NEWS

3 COMMENTS

Harvesting Data Helps Farms Grow Yields

640Labs, acquired by Monsanto, provides analytics that can help farmers increase production.

By David Talbot on January 23, 2015

Matt Schweigert owns 7,000 acres of corn and soybean fields in Cuba City, Wisconsin. His 25 tractors, combines, and other farm machinery are fitted with 640Labs technology; sensors that track GPS location and fertilizer application volume, the volume of fertilizer sprayed, and the

AIG KNOWLEDGE & INSIGHTS CENTER

Are local policies fit for your region?



Economía
La FIFA a la Justicia - Inseguridad ciudadana
GALES 26.40 27.20 28.95 30.65 1.70 2.70 7.85 9.05
PANORAMA
Google comenzó la cosecha de datos
El gigante de Internet ha empezado a invertir en una red de agricultura computarizada que analiza datos sobre semillas y suelos para ayudar a los cultivadores a obtener mejores rendimientos de sus cosechas y ahorrar dinero.
FOTO
el clima y el rendimiento de los cultivos puede mejorar productividad del campo.
Jue may 21 2015 04:05
Farmers Business Network Inc. obtuvo US\$ 15 millones en una ronda de inversión liderada por Google Ventures, la división de capital de riesgo del gigante tecnológico. La empresa pretende utilizar esta inversión de capital para expandir su servicio de análisis de campos a más lugares de Estados Unidos, según informó el Wall Street Journal.

Forbes / Tech

HULT International Business School

Highly Ranked MBA.

LEARN MORE

OCT 2, 2013 @ 8:13 AM 86,421 VIEWS

Monsanto Buys Climate Corp For \$930 Million



Bruce Upbin
FORBES STAFF

I manage our technology coverage. FULL BIO
Opinions expressed by Forbes Contributors are their own.

FOLLOW



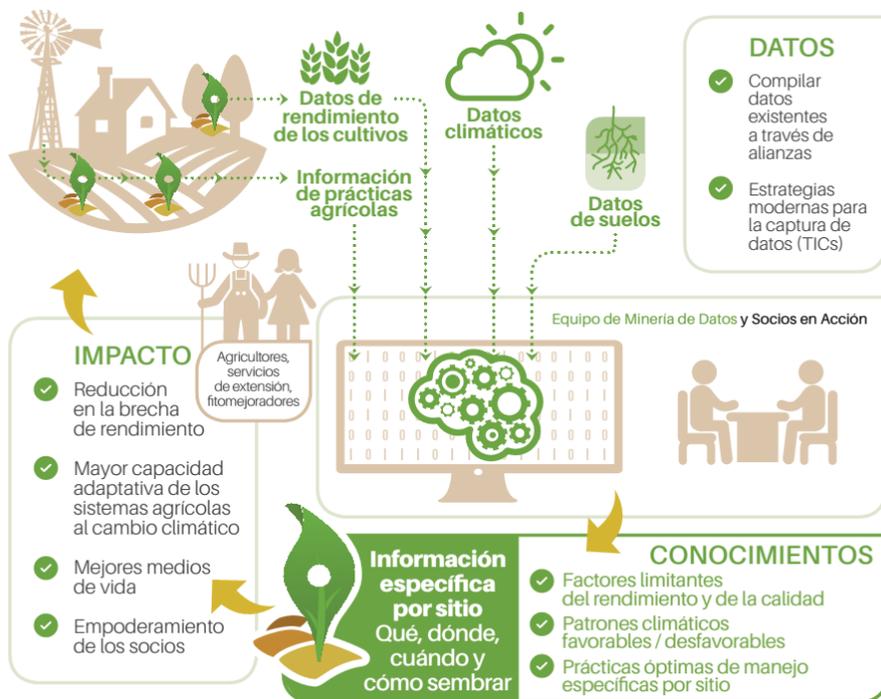
Who knew betting on the weather could make a billion dollars? It just did for The Climate Corporation, which underwrites weather insurance for farmers. Monsanto broke the news this week that it had bought the company for approximately \$930 million. The idea is to sell more data and weather-related products to farmers, and to combine that with Monsanto's seed and chemical divisions. Climate Corp (formerly known as Climate Data) was founded in 2007.

Agricultura Específica por Sitio Poniendo los datos al servicio de la agricultura

La minería de datos comerciales existentes, en cultivos, suelos y clima, así como el análisis de nuevos datos no-experimentales, contribuye a lograr una agricultura más productiva y resiliente ante el cambio climático.

Nuestros objetivos:

- Democratizar el uso de datos e información en la agricultura.
- Promover la agronomía basada en datos y el manejo específico por sitio.



Enfoque probado con 11 cultivos y ganadería

- Caña de azúcar
- Mora andina
- Limón
- Plátano
- Aguacate
- Lulo
- Frijol
- Mango
- Maíz
- Arroz
- Banano

Datos de **48000** cosechas analizadas

4 socios completamente empoderados trabajando de manera autónoma con la metodología

El enfoque se está implementando en **5** países

Software y aplicación móvil **para Android gratuitos**

Premios Internacionales

Global Pulse de Naciones Unidas 2014: Big Data Climate Challenge

Innovation Challenge del Banco Mundial 2014: Big Data for Development

Nuestro equipo

- Daniel Jiménez**, Ph.D. Líder del Equipo
Propicia el encuentro entre la Agricultura y el machine learning.
- Sylvain Delerce**, M.Sc. Coordinador del Equipo de Análisis
Mantiene los pies del equipo sobre la tierra.
- Luis Armando Muñoz**, M.Sc. Coordinador de Proyectos y Experto en Gestión del Conocimiento
Se asegura que el equipo hable el idioma humano al difundir los resultados entre los usuarios finales.
- Hugo Andrés Dorado** Estadístico en proceso de mutación a científico de datos
Maestro en R para diseñar algoritmos. Pone los servidores a sudar.
- Andrés Aguilar**, Ingeniero Agrícola
La última contratación. Ya vuela tan alto como el resto del equipo.
- Victor Hugo Patiño**, Especialista en Clima
Capaz de recorrer todos los terrenos de las series climáticas.
- Juan Felipe Rodríguez**, Informático
Arquitecto de nuestro sistema de información. Hace que los datos estén disponibles para que el equipo juegue 24/7.
- James Cock**, Mentor del Equipo
Pionero del concepto de agricultura específica por sitio en el sector de la caña de azúcar en Colombia.

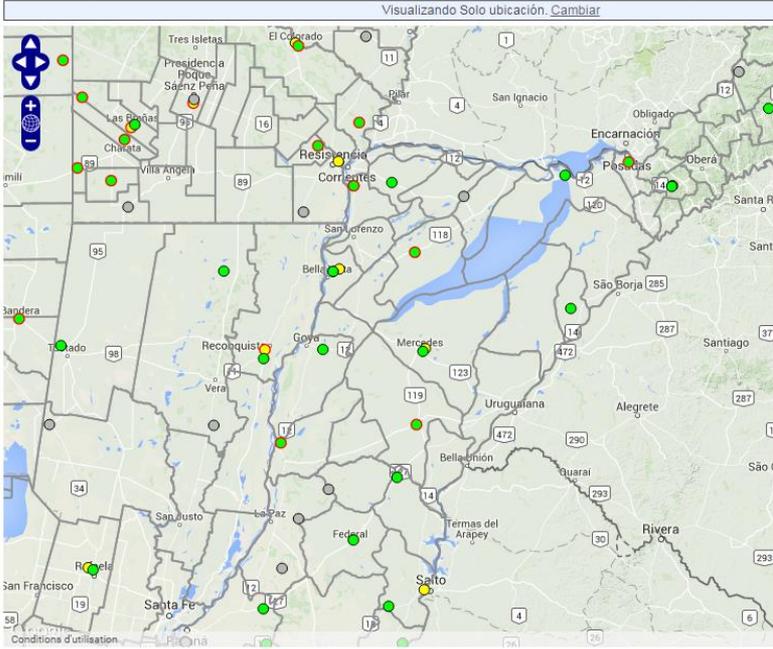
Selección de artículos científicos

- Cock J, Oberthur T, Isaacs C, Läderach PR, Palma A, Carbonell J, Anderson, E. (2011). Crop management based on field observations: Case studies in sugarcane and coffee. *Agricultural Systems* 104(9):755-769. doi:10.1016/j.agsy.2011.07.001
- Jiménez D, Cock J, Jarvis A, García J, Sattizabal HF, Van Damme P, Barreto-Sanz MA. (2011). Interpretation of commercial production information: A case study of lulo (*Solanum quitoense*), an under-researched Andean fruit. *Agricultural Systems* 104(3):258-270. doi:10.1016/j.agsy.2010.10.004
- Jiménez D, Cock J, Sattizabal HF, Barreto-Sanz MA, Pérez-Urbe A, Jarvis A, Van Damme P. (2009). Analysis of Andean blackberry (*Rubus glaucus*) production models obtained by means of artificial neural networks exploiting information collected by small-scale growers in Colombia and publicly available meteorological data. *Computers and Electronics in Agriculture* 69(2):198-208. doi:10.1016/j.compag.2009.08.008

www.ciat.cgiar.org/dapa www.open-aeps.org

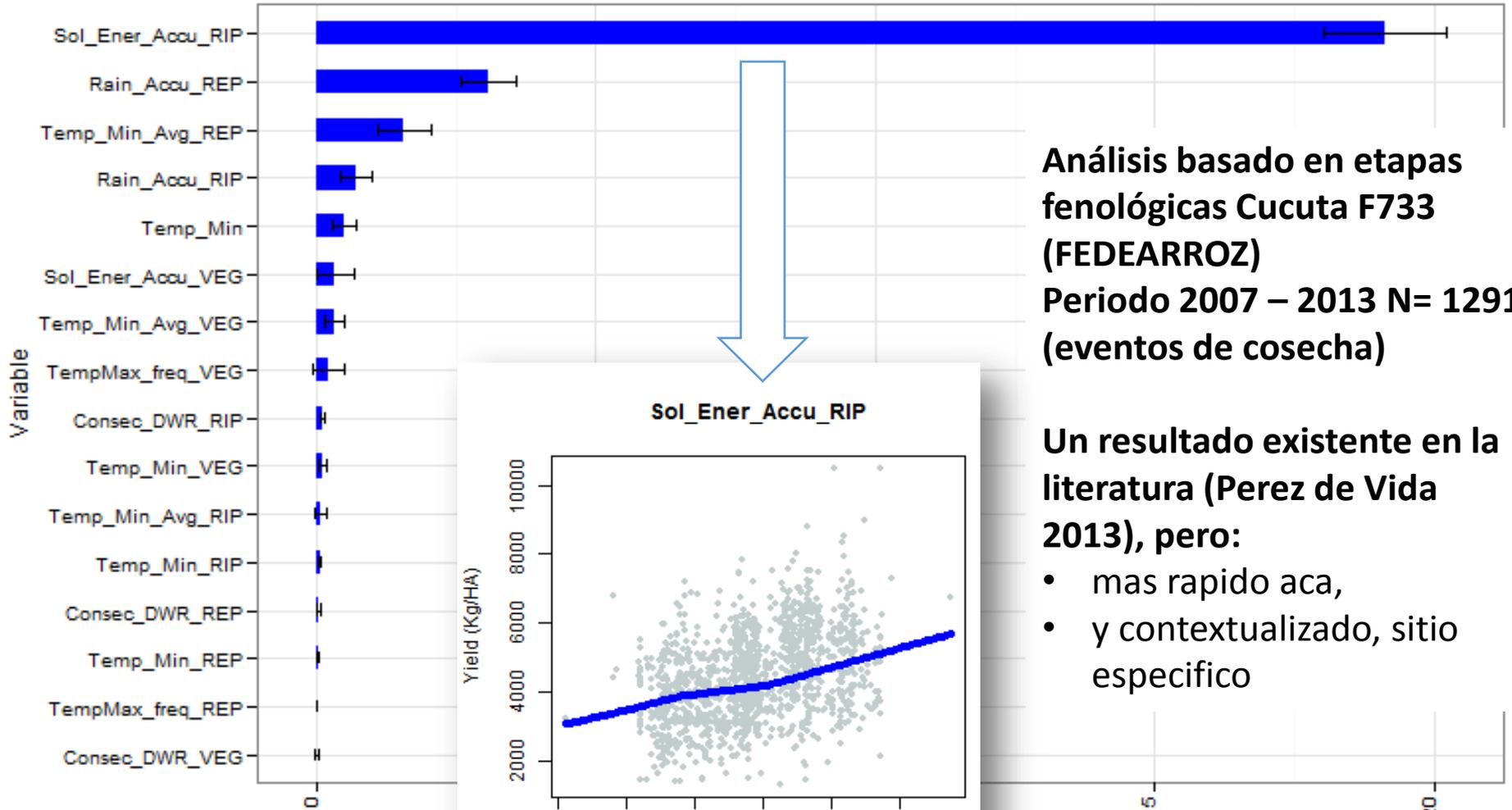
CONTACTO ✉ d.jimenez@cgiar.org

Julio 2015



Identificar factores limitantes por fases

Importance of variables (with a mean R2 of 30.21 %)



Análisis basado en etapas fenológicas Cucuta F733 (FEDEARROZ) Periodo 2007 – 2013 N= 1291 (eventos de cosecha)

Un resultado existente en la literatura (Perez de Vida 2013), pero:

- mas rapido aca,
- y contextualizado, sitio especifico

