

Analizando lotes de producción con ciencia de datos para conocer las brechas de rendimiento de los cultivos en Argentina

Dr. Diego Hernán Rotili

“Proyecto Brechas”



syngenta





“La era de los datos”

SI QUEREMOS TRABAJAR CON DATOS...
TENEMOS QUE APRENDER A TRABAJAR CON DATOS...

ESTO....



**NO SON DATOS...
NO ES ANÁLISIS...
NO NOS RESUELVE UN PROBLEMA CONCRETO...**



¿En qué consiste “trabajar bien” con datos?

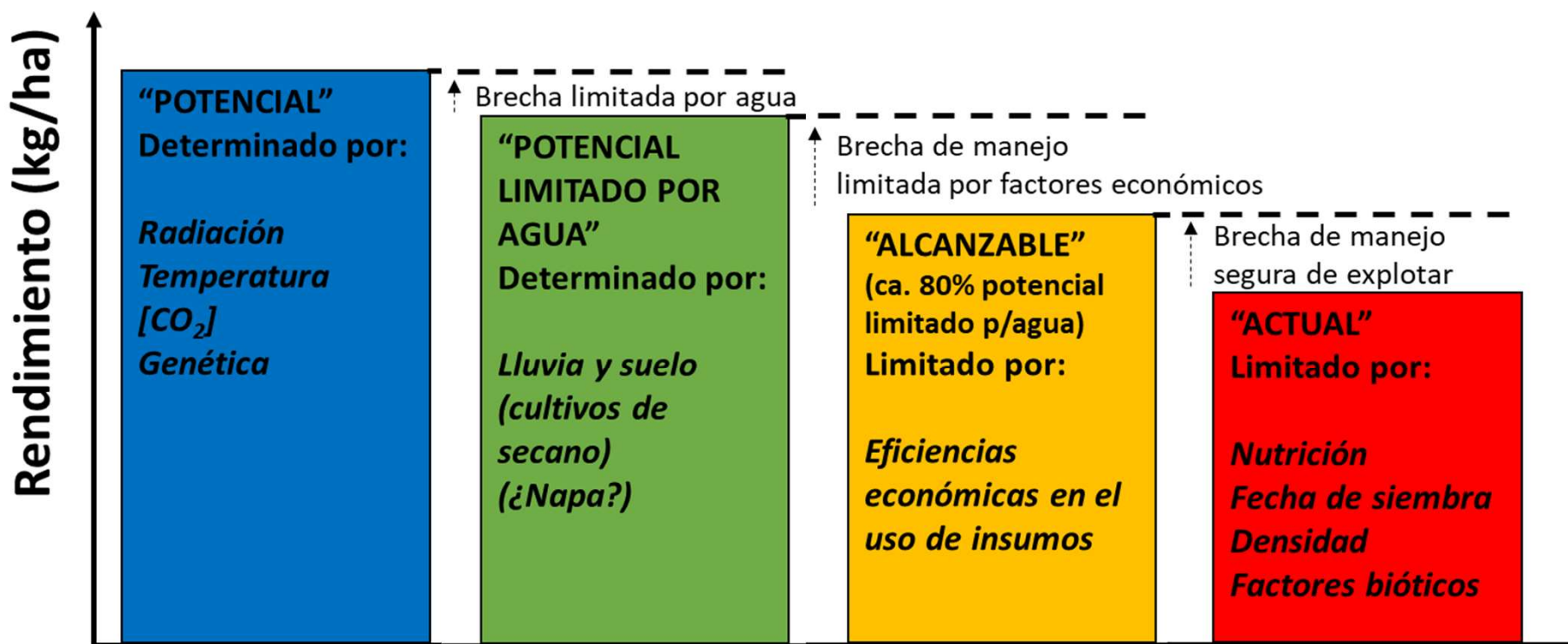
EN UNA CONSTANTE TOMA DE DECISIONES “CON SENTIDO”...

- 1) **RECONOCER UN PROBLEMA A RESOLVER**
- 2) **TENER UN OBJETIVO MEDIBLE**
- 3) **OBTENER DATOS QUE PERMITAN RESPONDER ESE OBJETIVO**
- 4) **CONOCER LAS LIMITACIONES QUE OFRECEN LOS DATOS**
- 5) **ELEGIR EL MEJOR ANÁLISIS EN FUNCIÓN DE MI OBJETIVO**
- 6) **ANALIZAR LOS DATOS**
- 7) **CONOCER LAS LIMITACIONES DEL MÉTODO DE ANÁLISIS**
- 8) **COMUNICAR LOS RESULTADOS DE FORMA ADECUADA**



El problema a resolver...

¿CUÁLES SON LAS BRECHAS DE RENDIMIENTO EN LOS CULTIVOS DE GRANO DE ARGENTINA?



Van Ittersum et al. (2013)
Field Crops Research



¿Cómo estudiar las brechas?

DIFERENTES MÉTODOS

A) Modelos de simulación

- Muchos años y lotes pueden considerarse
- Explora límites biofísicos
- No representa condiciones reales (malezas, plagas, enfermedades, maquinaria, etc.)

B) Experimentos a campo

- Microparcels en condiciones de campo
- Comparación de muchas prácticas en la misma condición
- + laborioso y necesidad de repetirlos muchos años

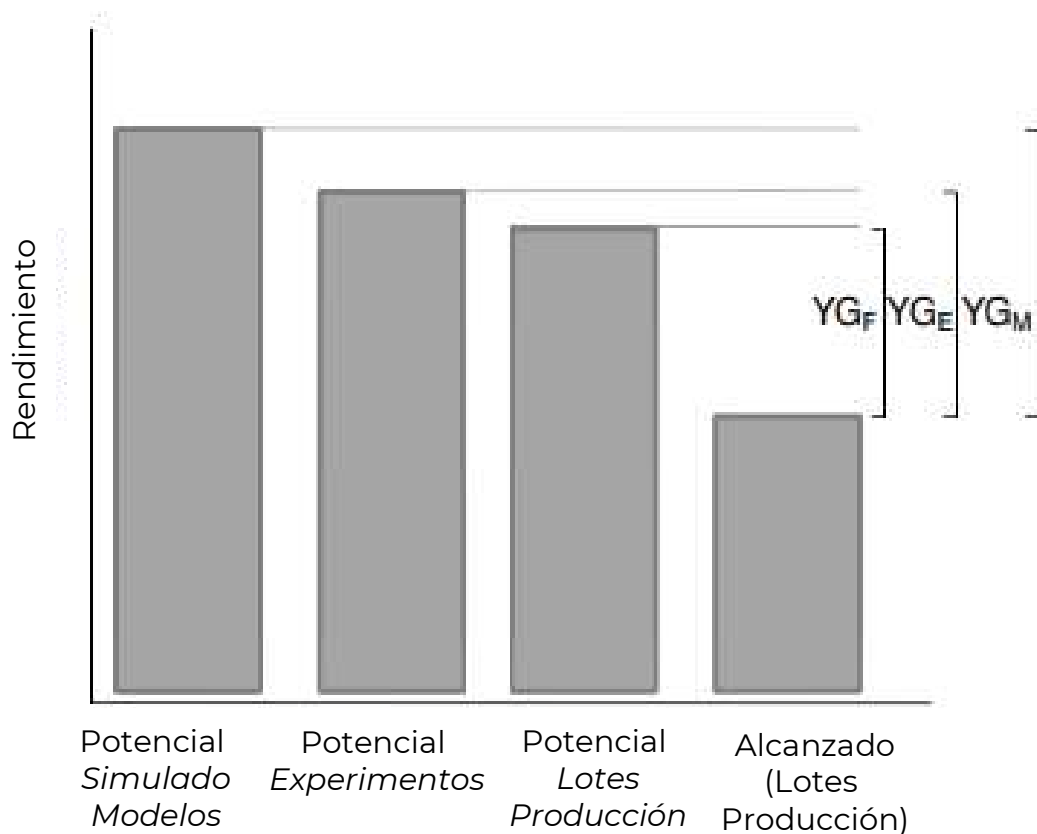
C) Lotes de producción

- Condiciones de producción muchos años disponibles
- Consideran el costo de producción y la eficiencia económica
- Pueden no contemplar condiciones de manejo superadoras



¿Cómo estudiar las brechas?

EL MÉTODO DE ANÁLISIS CONDICIONA EL RESULTADO



Pros y contras por método...

Los niveles de rendimiento (y la brecha) difiere según el método de análisis.

Los resultados de los análisis no indican verdades, sólo nos guían a acercarnos a alguna verdad en función de nuestro objetivo

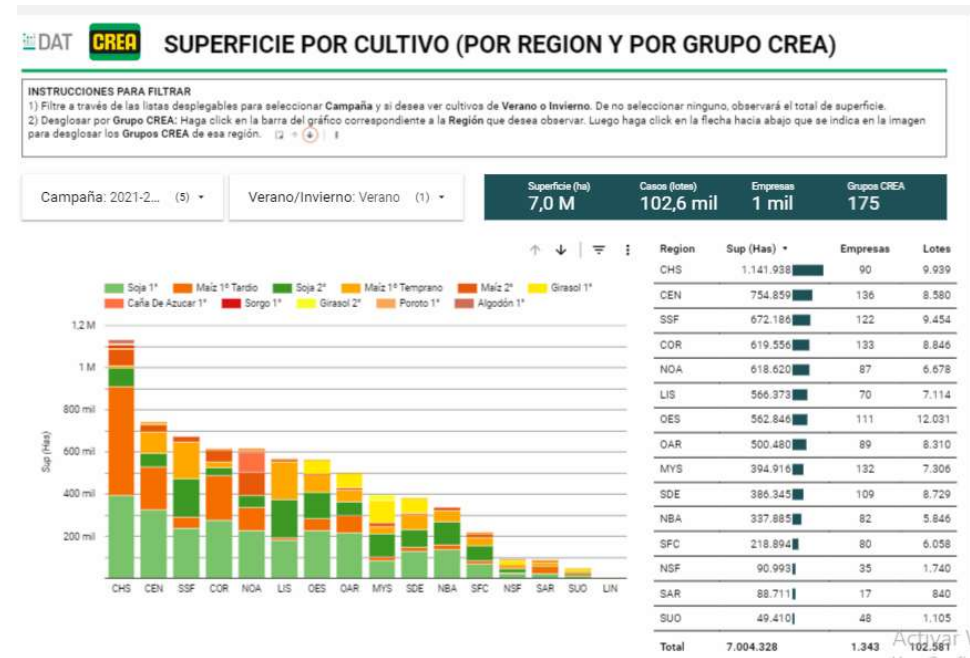
Lobell et al. (2009)
Annual Review of Environmental Resources

C) Lotes de producción

“Proyecto Brechas”



- La base contiene información de varios cultivos y hasta 25 campañas (330 mil casos > 22 M has).
- Para el proyecto analizamos la información de las últimas 5 campañas, que suman más de 6 millones de has entre maíz y soja.



+250 variables de manejo
Se redujeron a aprox. 20 por cultivo



Primera decisión – La escala/universo de aplicación

Macrozonas:

Norte

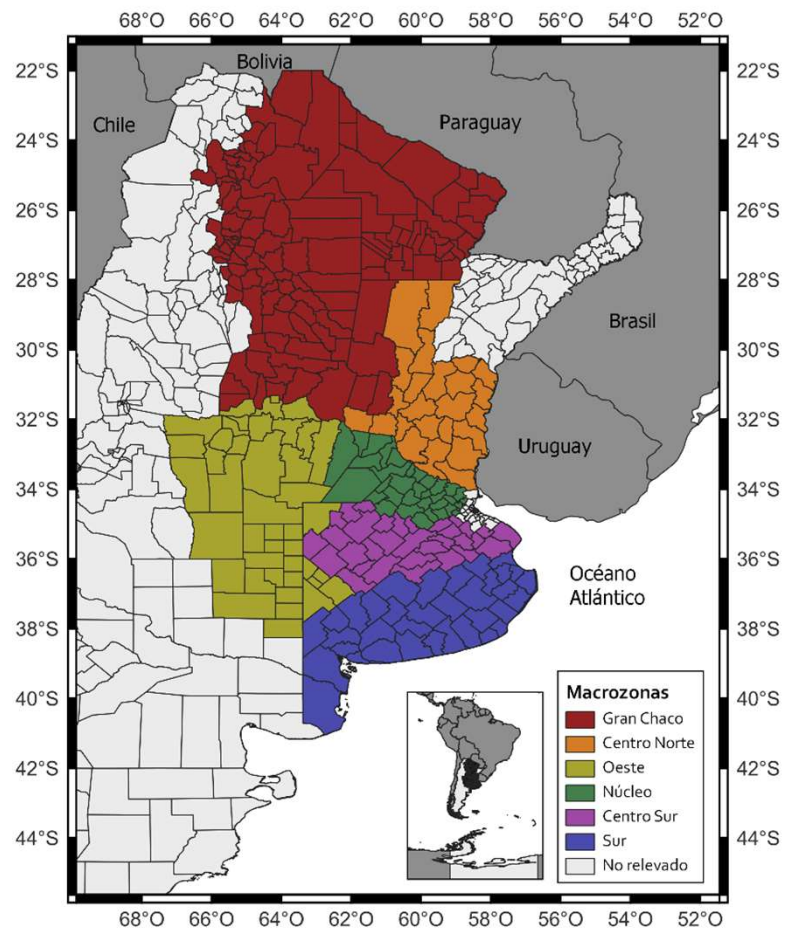
Centro Norte

Centro Oeste

Núcleo

Centro Sur

Sur



Primera limitación:

Los análisis valen a escala de “MACROZONA”.

Las conclusiones valen a esa escala.

Las aplicaciones también...



Generación de los datos

+250 variables de manejo
Se redujeron a aprox. 20 por cultivo
Ejemplo en maíz:

Rendimiento	Fecha de Siembra	P aplicado	Herbicidas postemergencia
Tenencia	Densidad	S aplicado	Fungicidas
Superficie	Distanciamiento	K aplicado	Insecticidas
Antecesor	Fertilización	Malezas resistentes	Napa
Híbrido	N aplicado	Herbicidas barbecho	Potencial Ambiente

Segunda decisión – Los ambientes

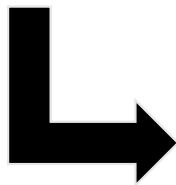
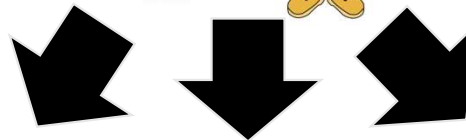
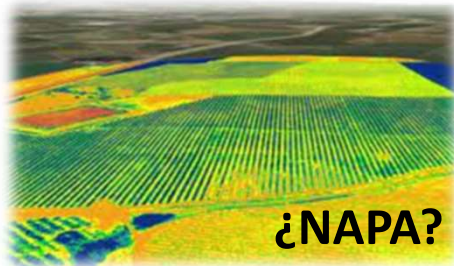


Segunda limitación:

Los análisis valen para los cultivos en secano

En cada “ambiente” subjetivo definido por el productor

↓
AMBIENTE PERCIBIDO
 (“POTENCIAL PRODUCTIVO”)



NAPA

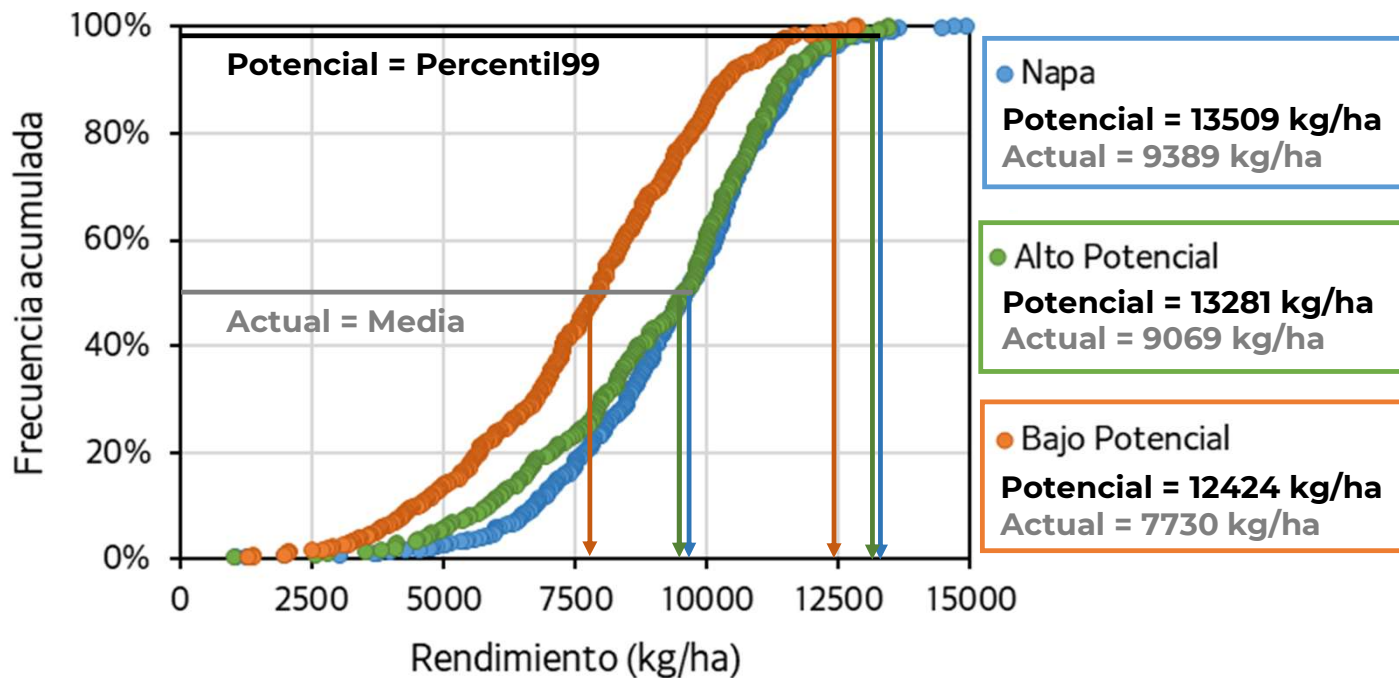
ALTO

BAJO

Tercera decisión – Rendimientos

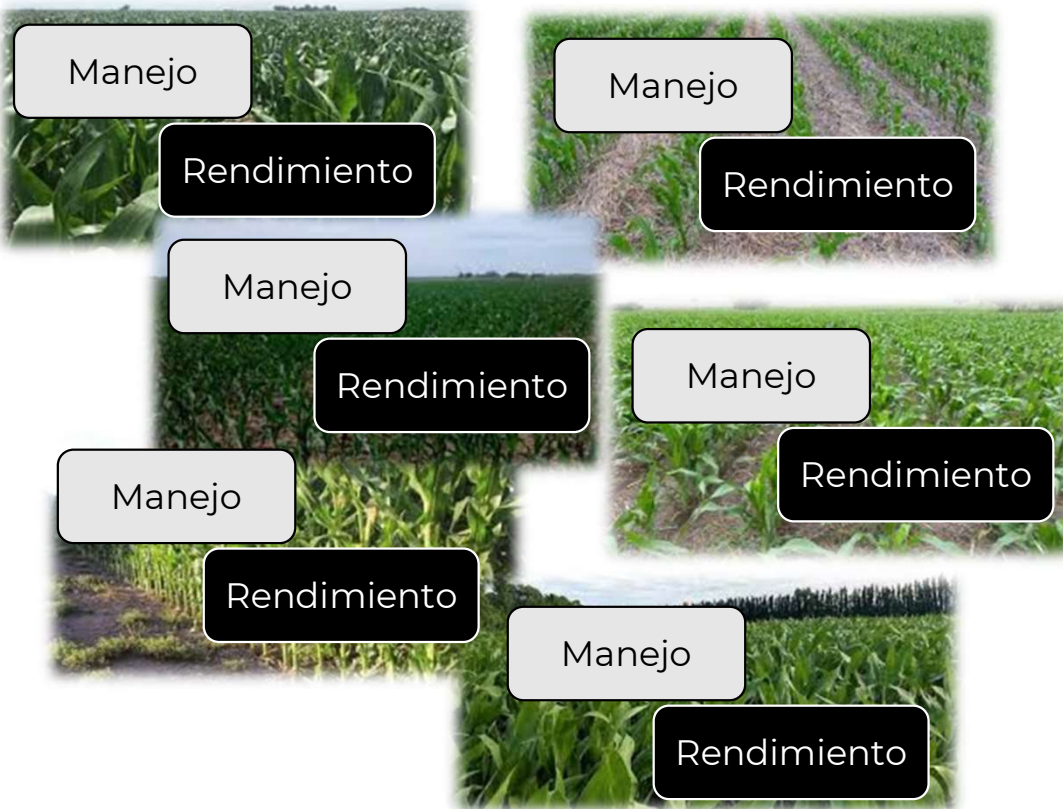
¿Cómo estimamos los niveles de rendimiento? Potencial y actual

Maíz temprano - Macroregión Centro Oeste



Tercera decisión – Rendimientos

¿Cómo estimamos los niveles de rendimiento? Alcanzable



Rendimiento Alcanzable:

Aquel rendimiento:

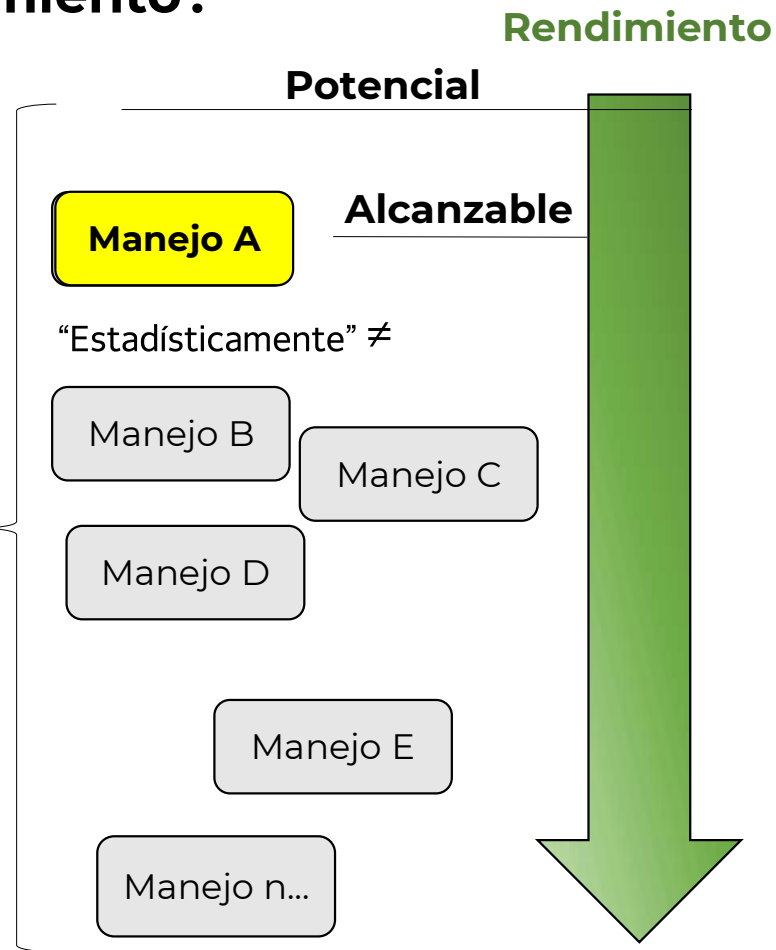
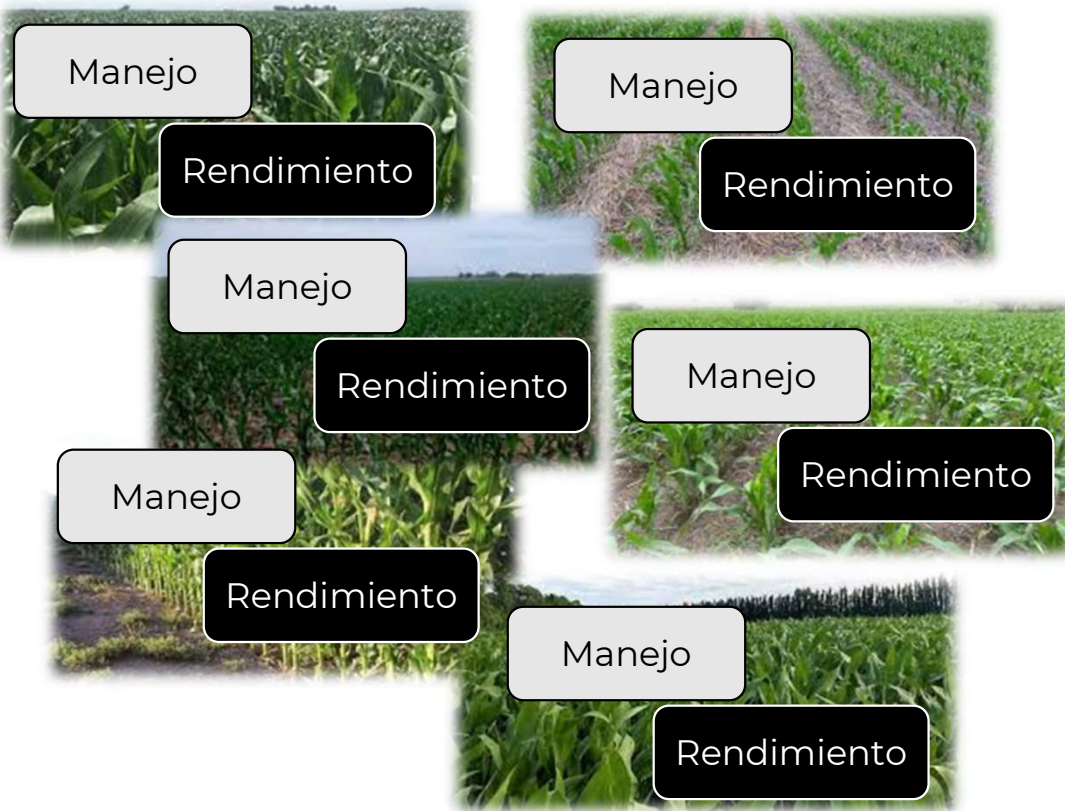
- (i) alcanzable por productores.
- (ii) en un ambiente de producción determinado,
- (iii) con un ambiente climático promedio, y
- (iv) cuando se adoptan las tecnologías agronómicas de procesos y de insumos económicamente óptimas.

Fischer (2015)
Field Crops Research



Tercera decisión – Rendimientos

¿Cómo estimamos los niveles de rendimiento Alcanzable

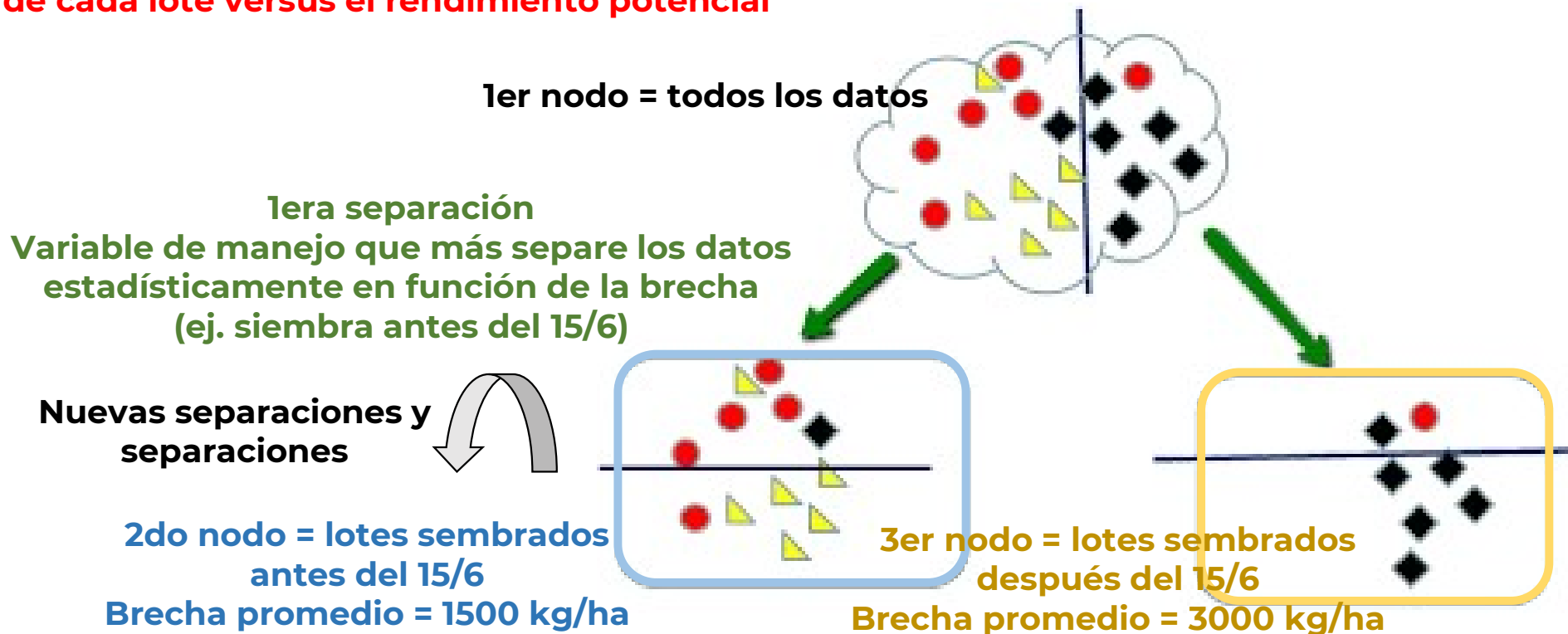


Tercera decisión – Rendimientos

¿Cómo estimamos los niveles de rendimiento?

Alcanzable = la técnica (ej. árboles de partición – *machine learning*)

Primero = Elegir variable respuesta = brecha de rendimiento de cada lote versus el rendimiento potencial



Tercera decisión – Rendimientos

¿Cómo estimamos los niveles de rendimiento?
Alcanzable = la técnica (ej. árboles de partición – *machine learning*)

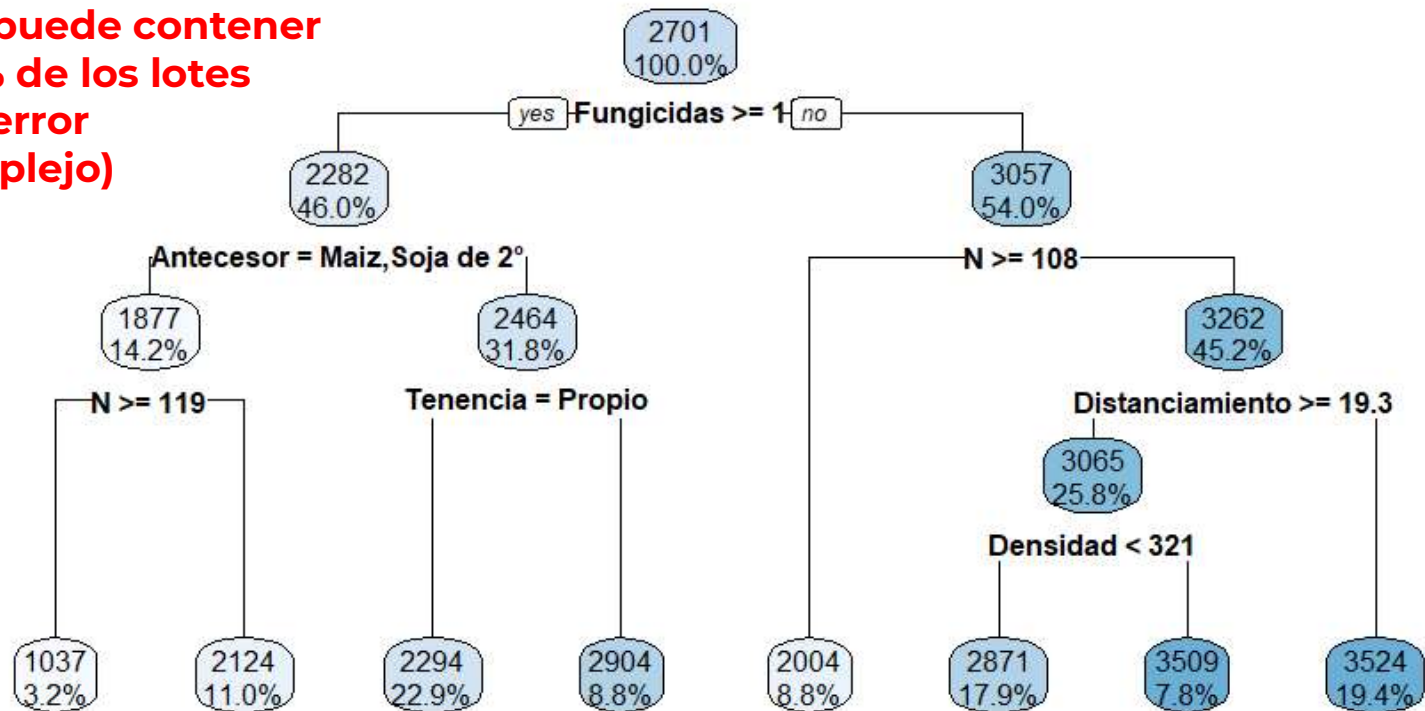
Restricciones elegidas:

- (i) el último nodo no puede contener menos del 3% de los lotes
- (ii) Minimización del error (método complejo)

Ej.:
Trigo

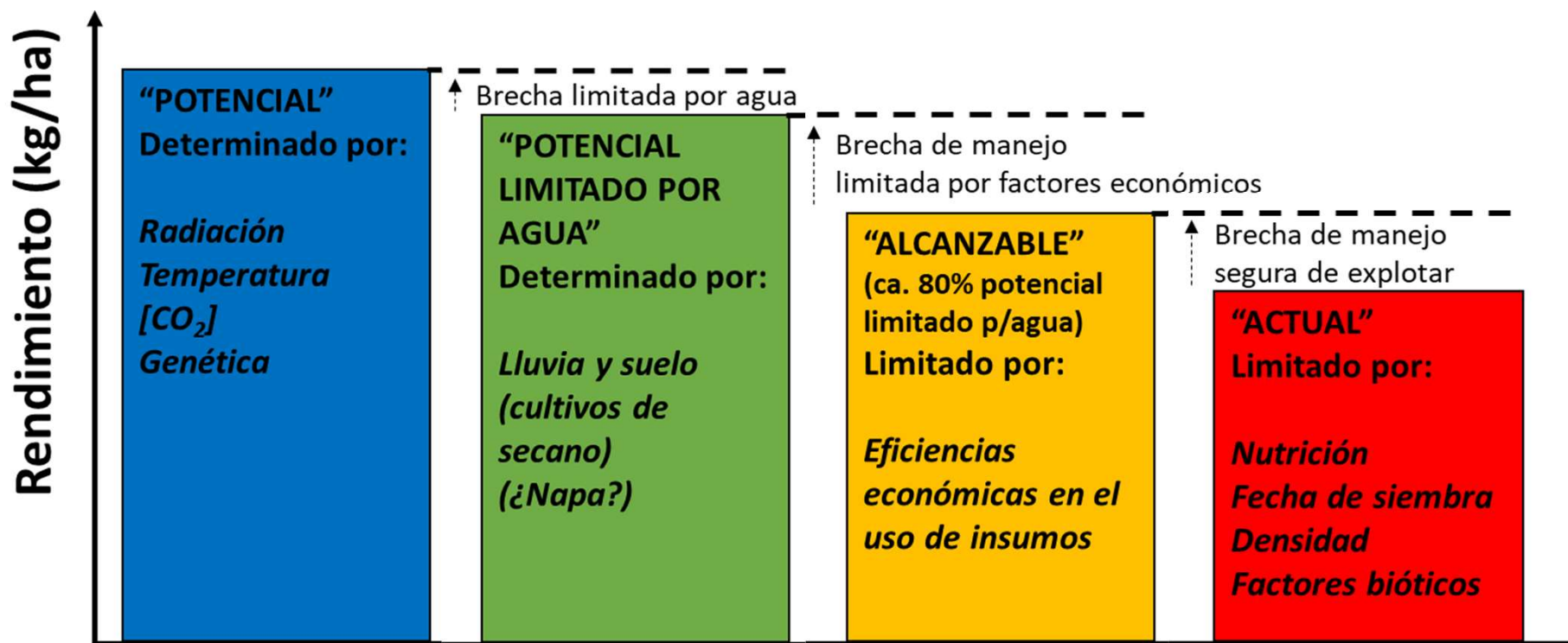
Macrozona
Centro Oeste

Lotes con Napa





El problema a resolver...¿respuestas?



Potencial en seco =

Percentil99

Alcanzable/Económico =

Manejo común

Menor brecha vs. potencial

Árbol de partición

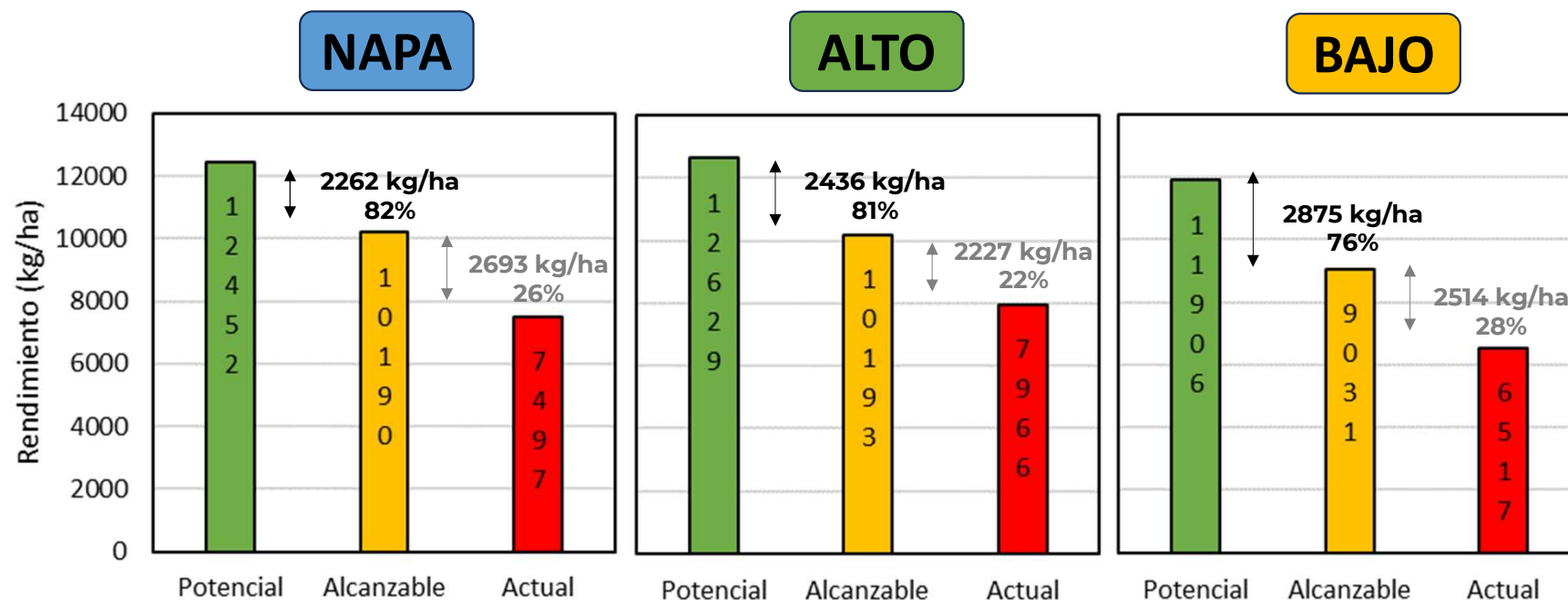
Actual =

Promedio



El problema a resolver...¿respuestas?

Ej.: Maíz Tardío Macrozona Centro Oeste



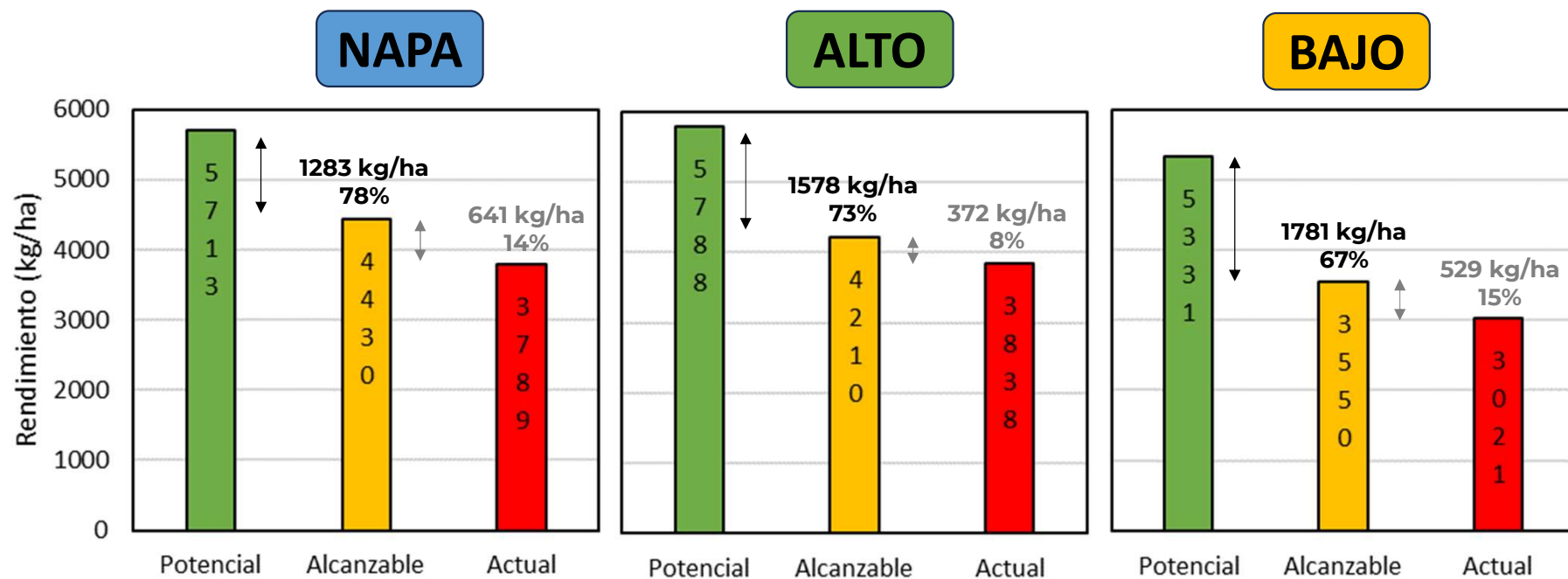
Varias interpretaciones a comunicar:

- (i) El alcanzable estuvo “alrededor” del 80% del potencial, pero esa relación disminuyó a medida que el ambiente fue más limitante en general.
- (ii) Hay bastante rendimiento que el ambiente podría “otorgar” pero no es rentable explotar.
- (iii) Hay mucho rendimiento explotable aún a través del manejo.



El problema a resolver...¿respuestas?

Ej.: Soja 1^{ra} Macrozona Centro Oeste

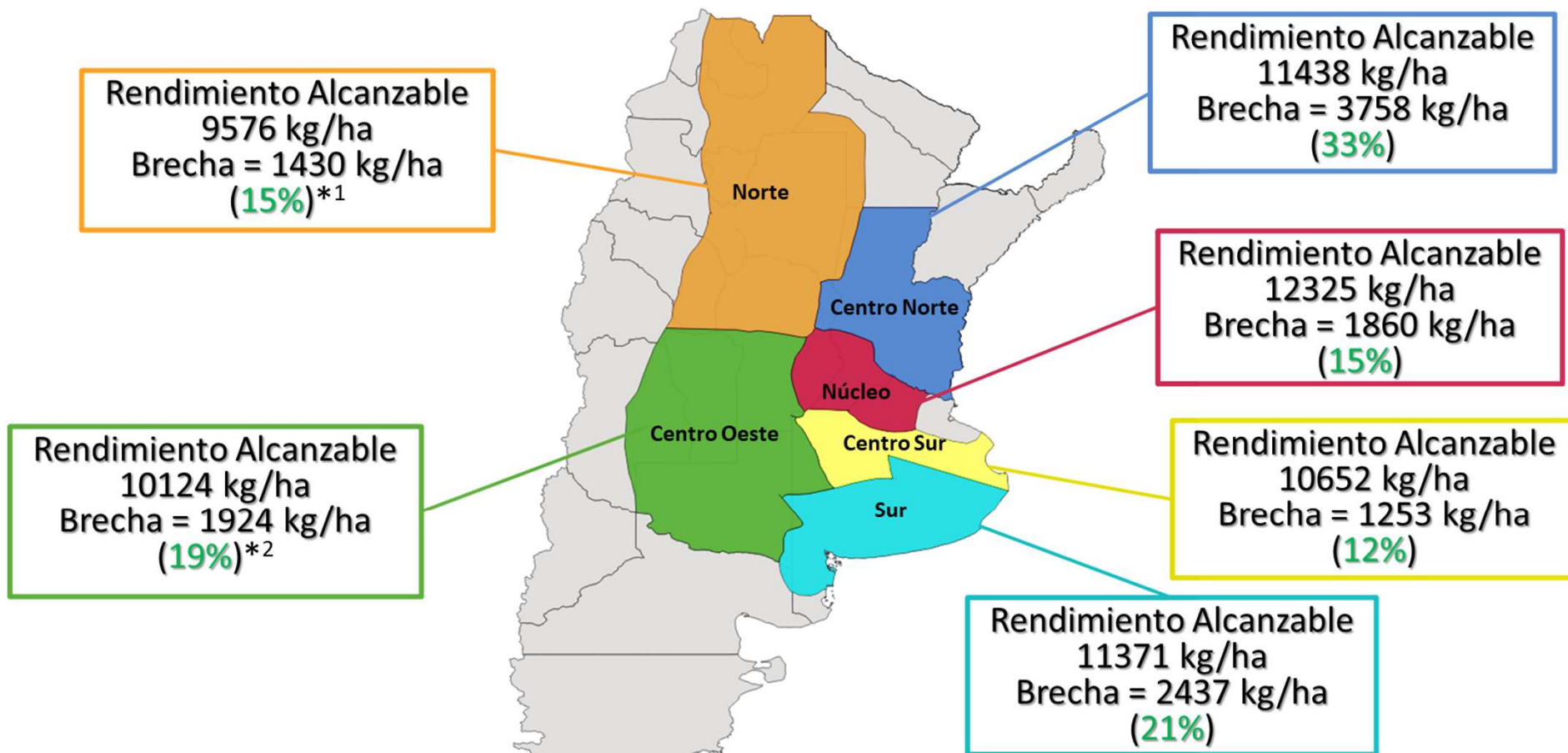


Varias interpretaciones a comunicar:

- (i) El alcanzable estuvo “alrededor” del 70% del potencial, pero esa relación disminuyó a medida que el ambiente fue más limitante en general.
- (ii) Hay bastante rendimiento que el ambiente podría “otorgar” pero no es rentable explotar.
- (iii) Hay cierto rendimiento explotable aún a través del manejo.

Llegando a resultados – Brechas en el país

Ej.: Maíz Temprano Ambientes Alto Potencial



Llegando a resultados - Brechas

Ej.: Soja 1^{ra} Ambientes Con Napa





Entonces...

SI QUEREMOS TRABAJAR CON DATOS...
TENEMOS QUE APRENDER A TRABAJAR CON DATOS...

La cantidad de datos disponible actualmente en el sector agropecuario es enorme...

Imaginemos lo que vamos a tener disponible en un tiempo...

Pero...

No se trata de “tener datos”, se trata de:

“saber qué problema queremos resolver”

“saber cómo construir una base de datos”

“saber las limitaciones de la base”

“saber qué análisis hacer para cumplir el objetivo”

“saber las limitaciones del análisis”

“saber comunicar adecuadamente”

“saber que sabemos poco y podemos aprender”

**Pero no hay atajos, debemos aprender a trabajar con datos...o...
valorar y pagarle muy bien al que sepa trabajar con datos...**

*Muchas gracias
Y ¡¡buena jornada!!*

Dr. Diego Hernán Rotili

“Proyecto Brechas”



syngenta

