

# Agro Minería de Datos



Yanina Bellini Saibene

**OpenData**

**Inteligencia Artificial**

**BigData**

**DataMining**

**DataVis**

**Machine Learning**



Instituto Nacional de  
Tecnología Agropecuaria

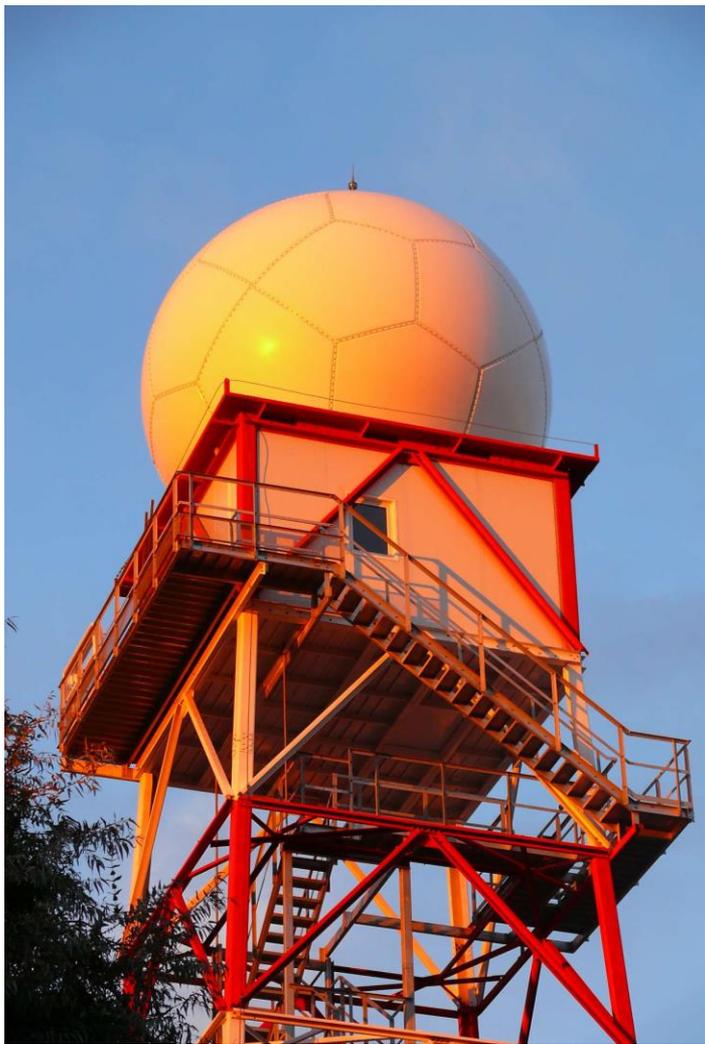


Ministerio de Agroindustria  
Presidencia de la Nación

# Contexto

- **Nuevas fuentes** de datos de **complejidad** y **volumen** crecientes.
- **Disponibilidad** de datos **diversos** y en **gran escala**.
- Creación de un cúmulo de **métodos útiles** para **almacenarlos** y **extraer** información de los mismos.
- Uso **intensivo** en la **toma de decisiones** a un nivel nunca antes imaginado.

# Aplicaciones concretas: RADAR Meteorológico

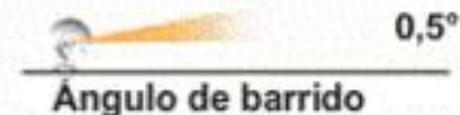
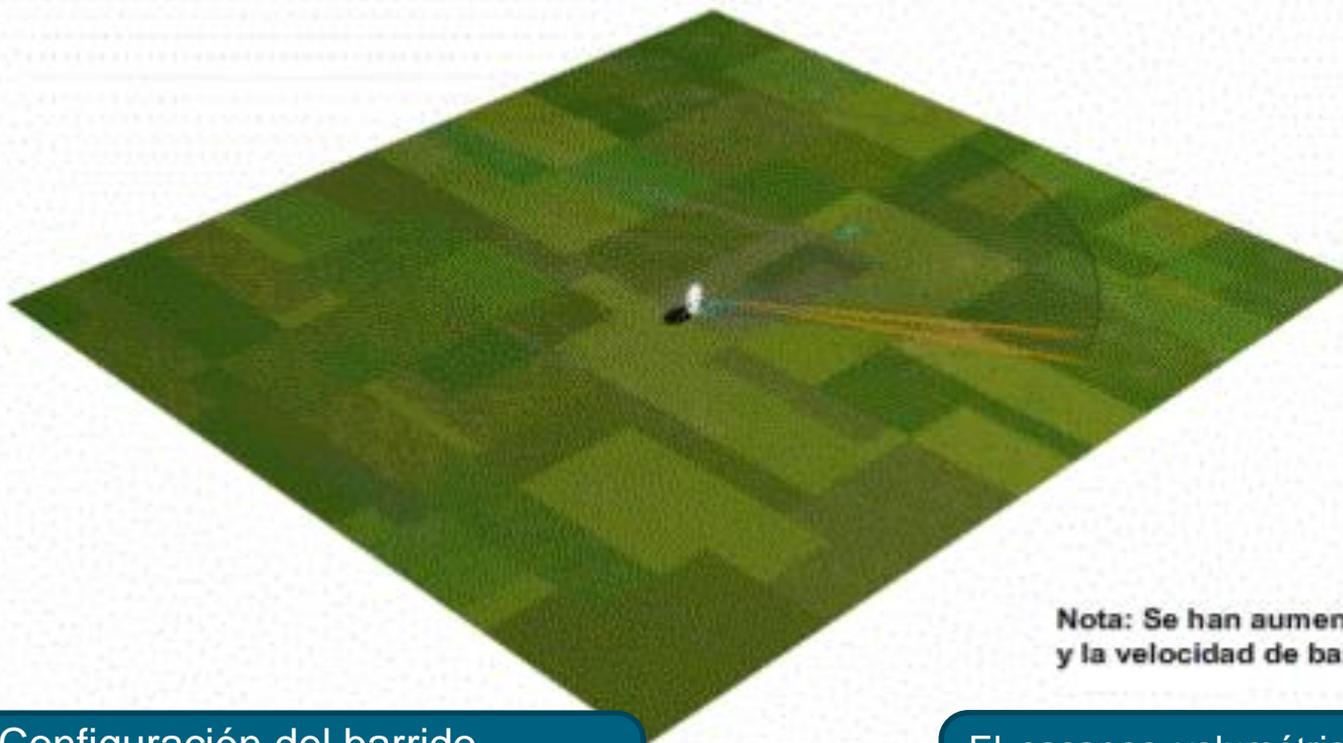


- Avisos de tormenta.
- Mejora en la capacidad de pronóstico de corto plazo y diagnóstico del tiempo en cuanto a la previsión de lluvias
- Mejora en el conocimiento de la evolución del balance hídrico.
- Aumento de la capacidad de generar alertas tempranas de tormentas severas.
- Mejora en la estimación de la precipitación y su distribución espacial.
- Identificación de áreas afectadas por granizo.

[radar.inta.gob.ar](http://radar.inta.gob.ar)

# Funcionamiento y Configuración

## Patrón de barrido del radar



**Nota:** Se han aumentado el ángulo de elevación y la velocidad de barrido para mostrar mayor detalle.

Configuración del barrido  
10 minutos. 144 adquisiciones  
por día

El escaneo volumétrico de RM, fue configurado con giros a 360° horizontal, iniciando con elevación a 0,5° hasta 15,1° en un total de 12 ángulos

## ¿Por qué granizo?

Fenómeno meteorológico capaz de infligir cuantiosos daños en edificios, cosechas.

Reducida extensión espacial y temporal de las tormentas de granizo, detectar su ocurrencia en superficie es una tarea difícil y costosa.

## ¿Por qué Data Mining?

Exploración y análisis de grandes volúmenes de datos

Para descubrir reglas y patrones que resulten útiles y comprensibles al usuario

Funciona con datos observacionales.

## ¿Por qué RADAR?

Sensor remoto.

Abarca una gran superficie.

Con una resolución uniforme en tiempo y espacio.

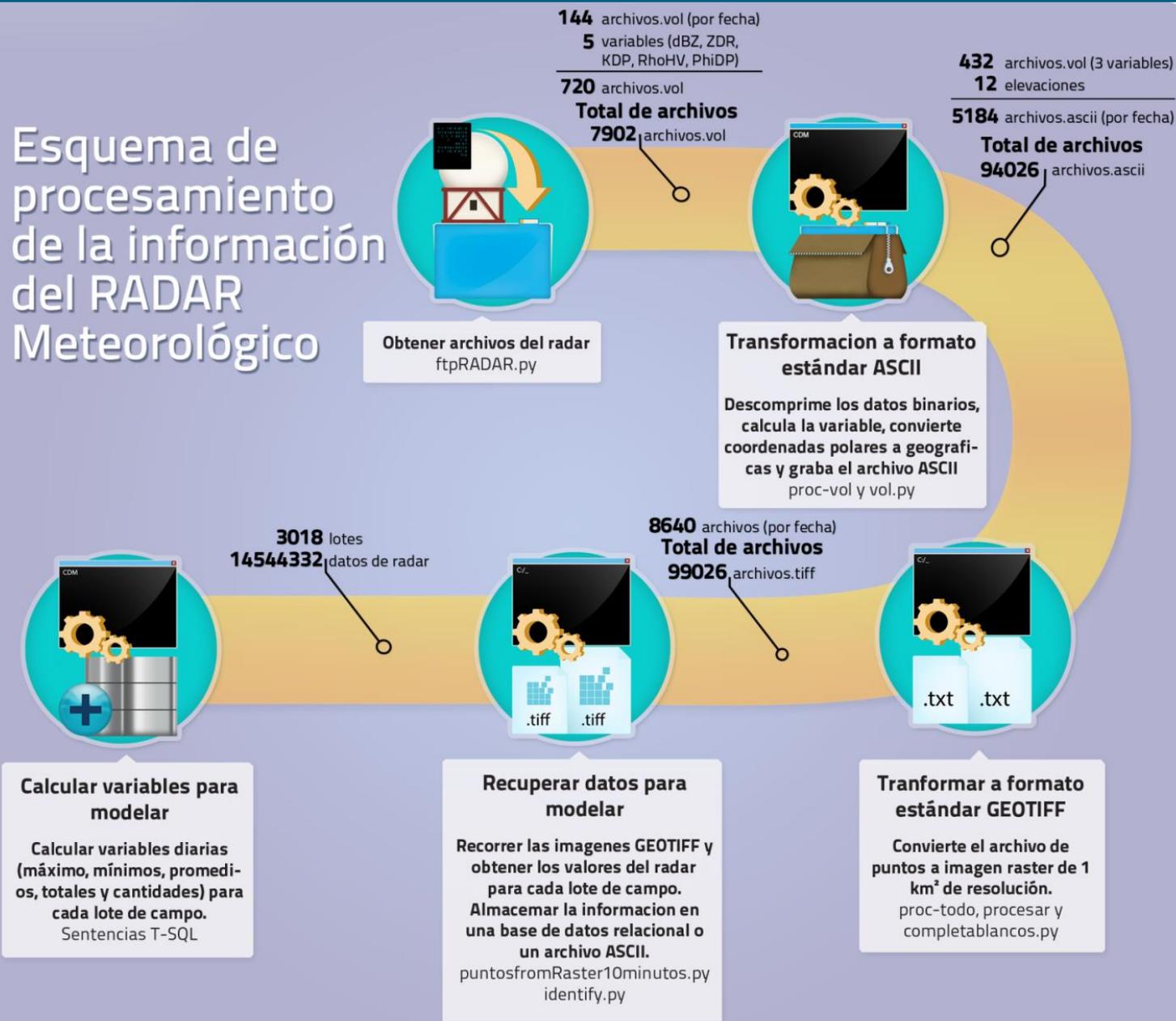
Flujo de información de 17 Gb de datos diarios aproximadamente.

## Objetivo

**Desarrollo de un modelo de estimación de ocurrencia de granizo en superficie y daño en cultivos** por medio de técnicas de **Data Mining**, tomando como base los **datos generados por el radar meteorológico de la EEA Anguil.**

# Metodología: CRISP-DM

## Esquema de procesamiento de la información del RADAR Meteorológico



Python (desde punto 1 a 4)

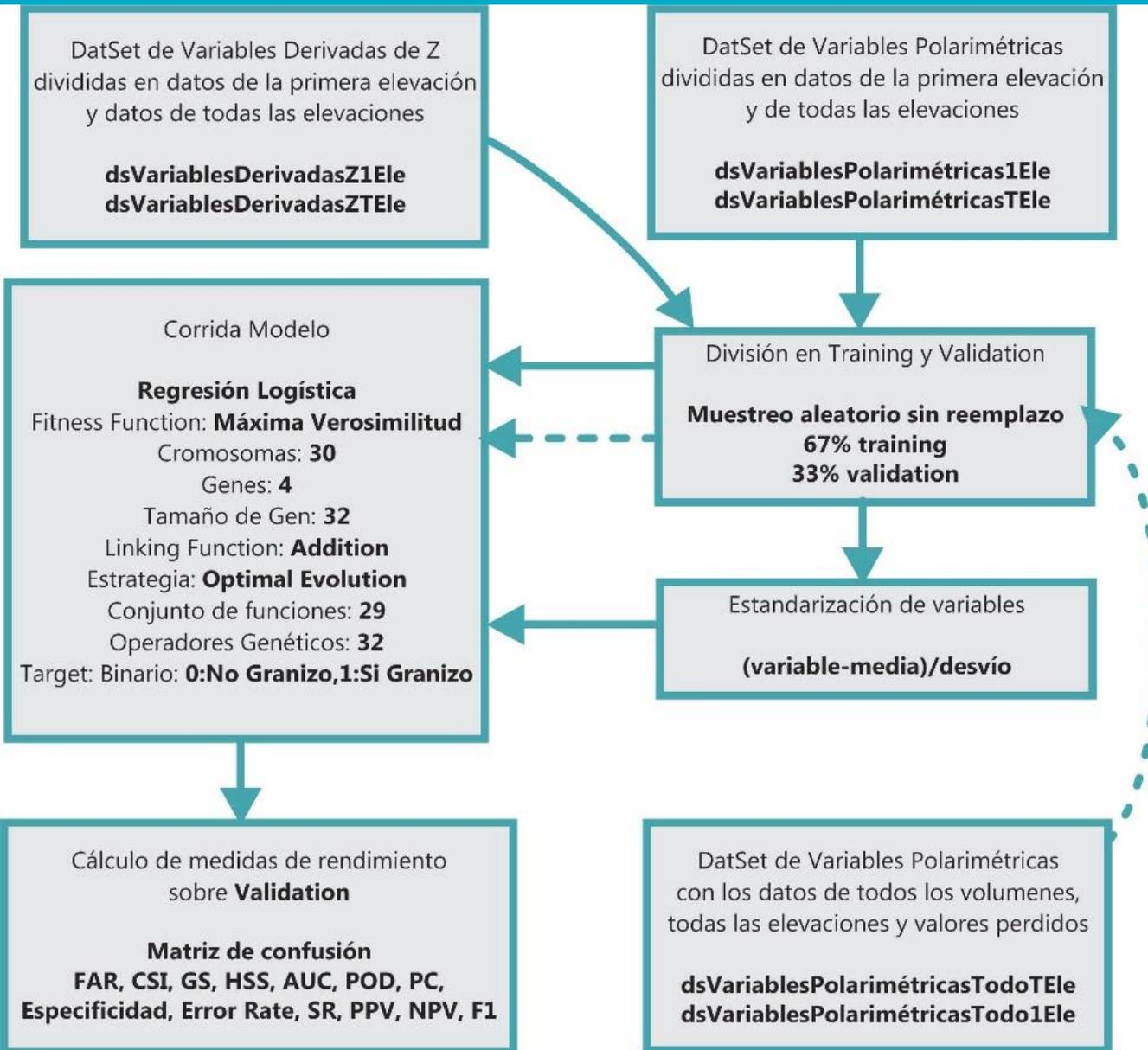
SQL (punto 5)

R (análisis estadístico de los datos y modelado)

Código Fuente disponible en:

<https://github.com/INTA-Radar>

# Metodología: esquema de modelado



**Gene Expression Programming**  
(evolutivo) con  
**Regresión Logística.**

Conjunto de datos que **unifica** los datos del radar con los datos de campo (+**3000** casos).

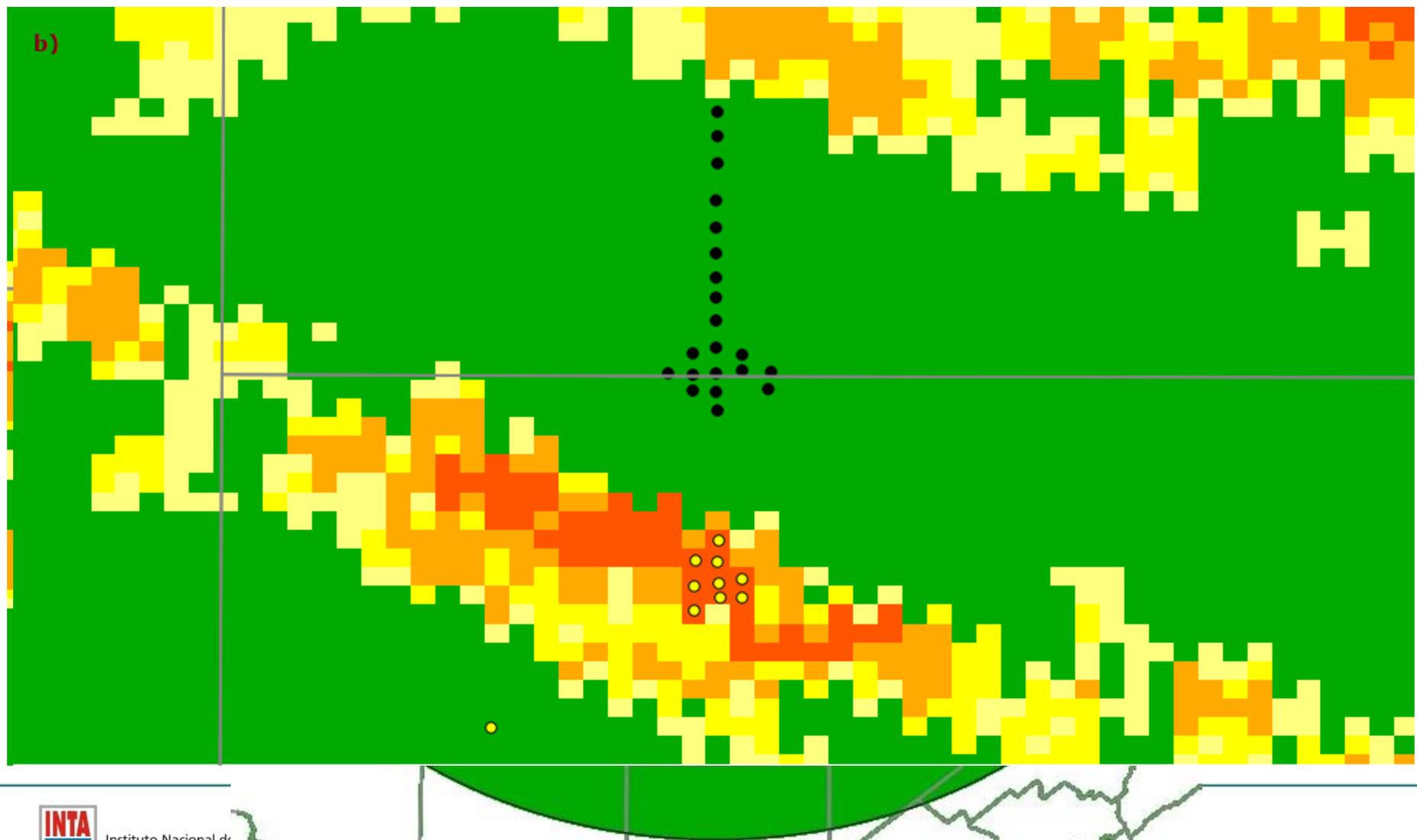
# Resultados

Para **ocurrencia de granizo** se generaron los modelos **Kurá** y **Pire** con alta performance:

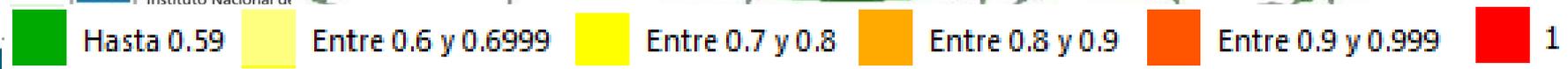
- $POD = VP / (VP + FN) \rightarrow 84\%$ ,
- $FAR = FP / (VP + FP) \rightarrow < 23\%$ ,
- $PC = (VP + VN) / (VP + VN + FP + FN) \rightarrow > 84\%$

quedando **terceros** entre **30** modelos internacionales.

		Valores Observados	
		Si	No
Valores predichos	Si	<b>VP</b> (Verdaderos Positivos) Identificados Correctamente	<b>FP</b> (Falso Positivo) Identificados incorrectamente
	No	<b>FN</b> (Falso Negativo) Incorrectamente rechazados	<b>VN</b> (Verdadero Negativo) Correctamente rechazados



**INTA** Instituto Nacional de



# ¡Muchas gracias!

[bellini.yanina@inta.gob.ar](mailto:bellini.yanina@inta.gob.ar)  
@yabellini

 rian.inta.gob.ar  
geointa.inta.Gob.ar  
inta.gob.ar/anguil

 @geointa @intaanguil

 /intaanguil /geointa