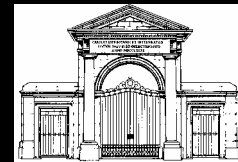


GENERACIÓN DE "AUSENCIAS" Y PSEUDOAUSENCIAS. VALIDAZACIÓN DE ALGORITMOS

Rubén G. Mateo, Real Jardín Botánico



INTRODUCCCIÓN

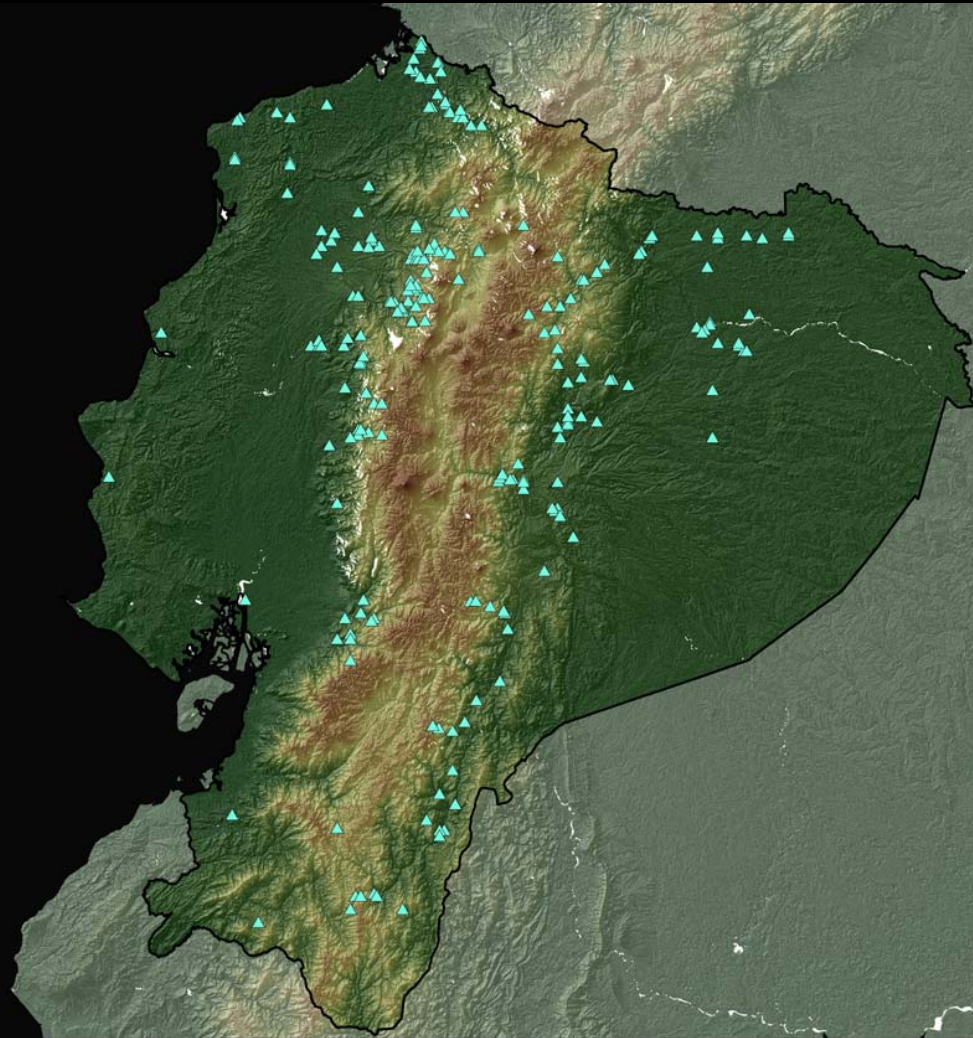
“Ausencias” y pseudoausencias

- ◆ Las colecciones botánicas ofrecen datos de la **presencia** de una determinada especie, no solemos disponer de ausencias.
- ◆ Los **métodos discriminativos** (MARS, CART, LMR, etc.) necesitan ausencias para poder generar modelos de idoneidad.
- ◆ Dispondremos de ausencias reales si realizamos **trabajo de campo**.
- ◆ La forma más frecuente de generar estas ausencias es realizar una distribución al azar de puntos (**pseudoausencias**).
- ◆ Si disponemos de una colección de datos como la perteneciente al género *Anthurium* (Araceae), podremos generar estas **ausencias** de una manera más fiable.
- ◆ El especialista de este género, **Tom Croat**, ha recolectado de una forma muy intensa gran parte de Ecuador, cubriendo todo los ecosistemas posibles.

GENERACIÓN DE DATOS "TRAINING"

Especialista género *Anthurium*

▲ Colecciones de T. Croat del género *Anthurium*



GENERACIÓN DE DATOS "TRAINING"

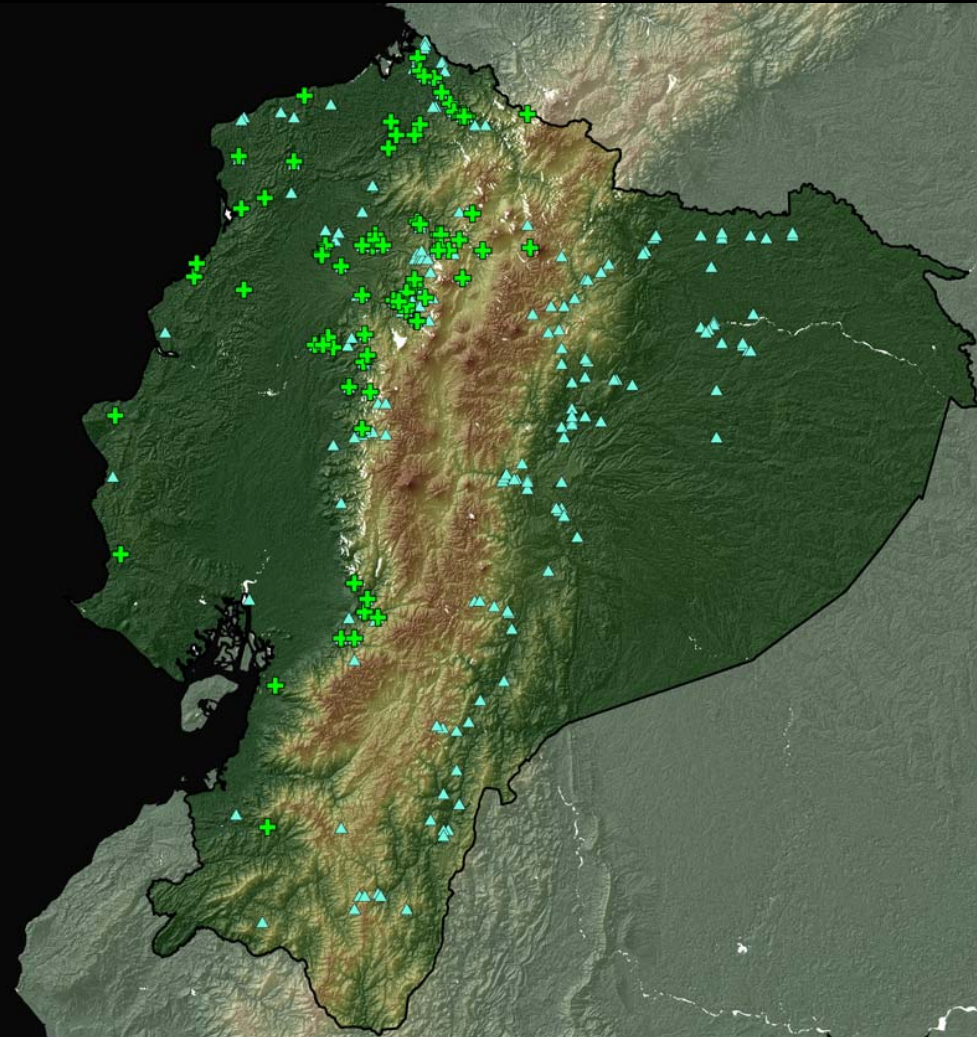
Colecciones *A. dolichostachyum*



Colecciones de T. Croat del género *Anthurium*



Colecciones de *Anthurium dolichostachyum*



GENERACIÓN DE DATOS "TRAINING"

"Ausencias"



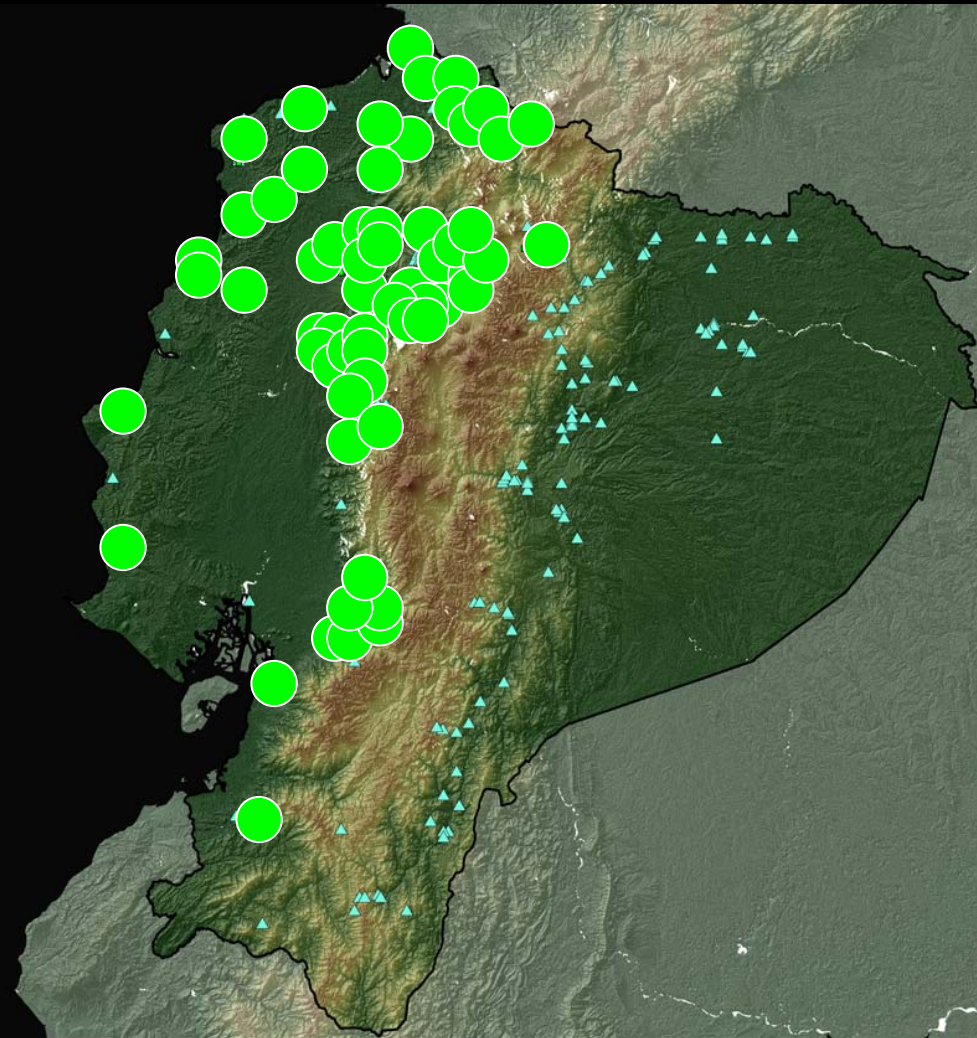
Colecciones de T. Croat del género *Anthurium*



Colecciones de *Anthurium dolichostachyum*



Zona tampón de 30 Km de radio

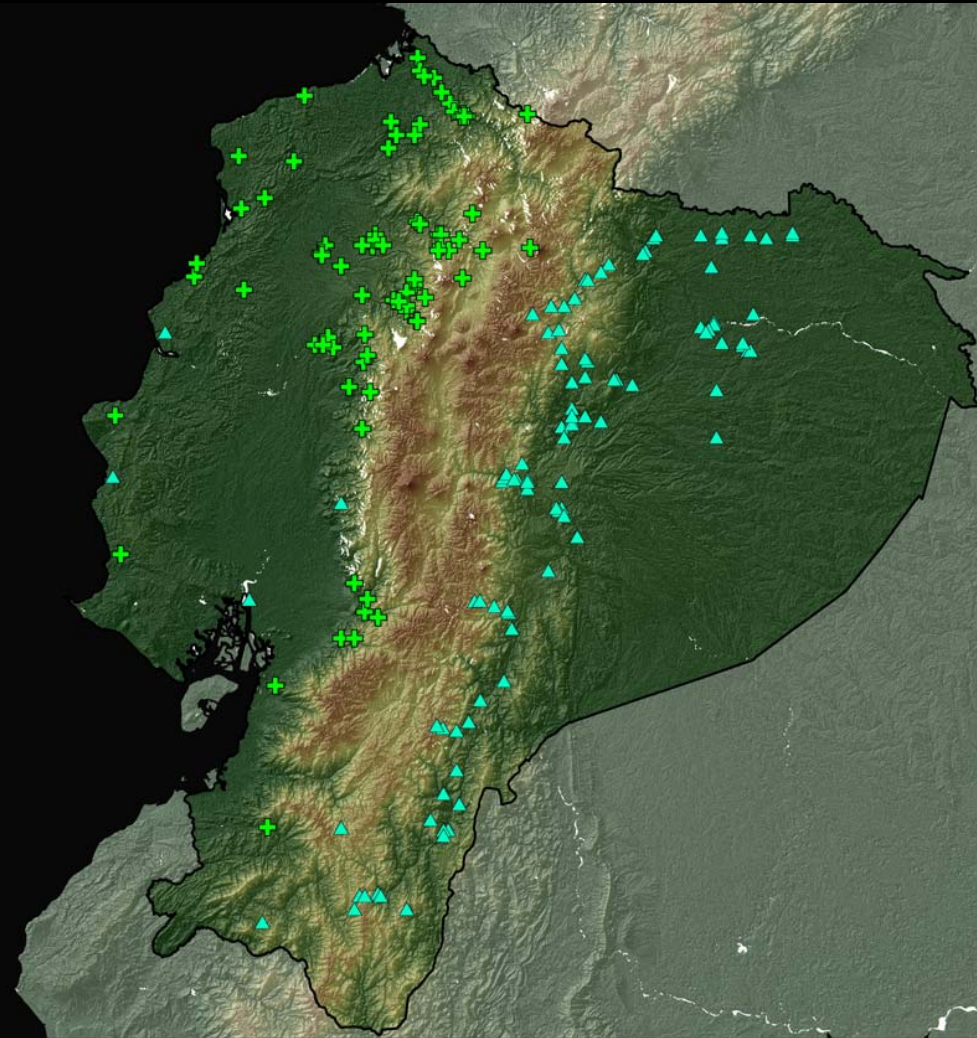


GENERACIÓN DE DATOS "TRAINING"

"Ausencias"

▲ Colecciones de T. Croat del género *Anthurium* donde no encontró *A. dolichostachyum*

✚ Colecciones de *Anthurium dolichostachyum*



GENERACIÓN DE DATOS "TRAINING"

"Ausencias"



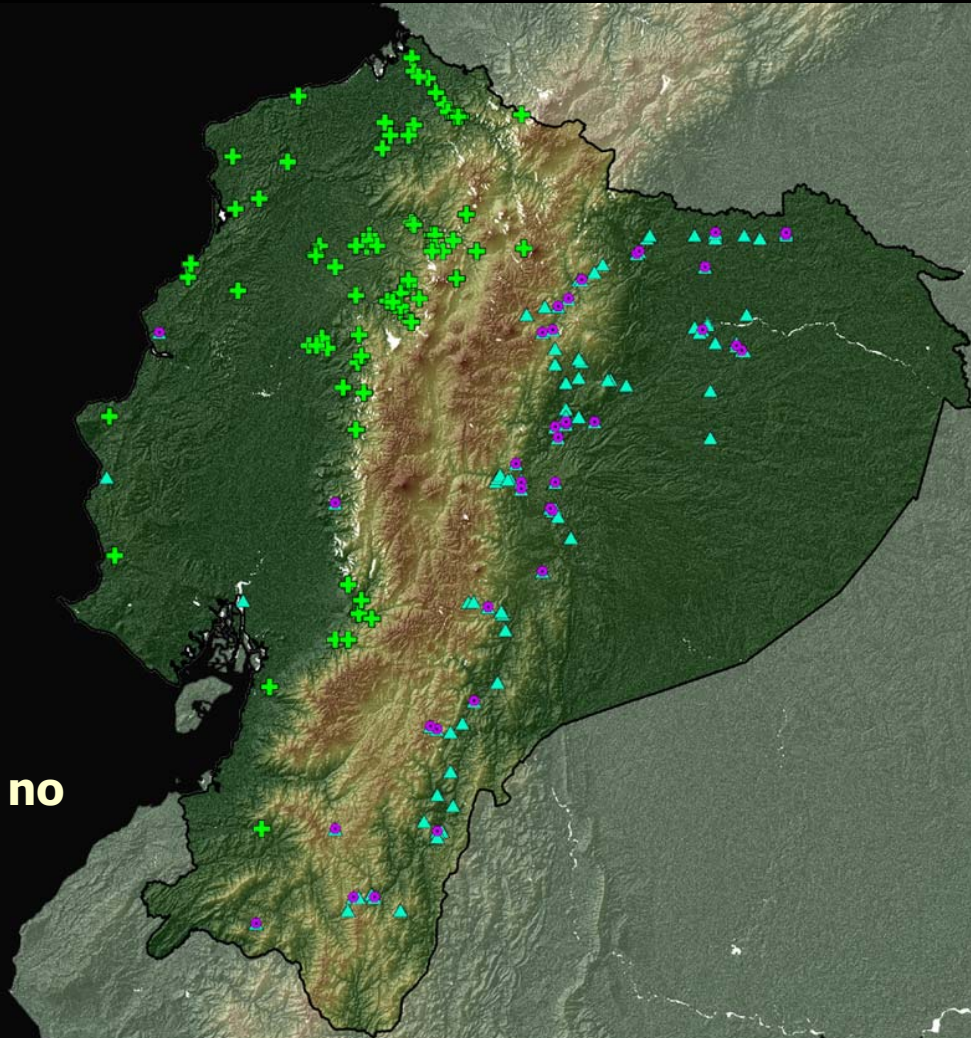
Colecciones de T. Croat del género *Anthurium* donde no encontró *A. dolichostachyum*



Colecciones de *Anthurium dolichostachyum*



Colecciones de T. Croat donde no encontró *A. dolichostachyum*
(training negativo 1)



GENERACIÓN DE DATOS DE "TRAINING"

Presencias



Colecciones de T. Croat del género *Anthurium* donde no encontró *A. dolichostachyum*



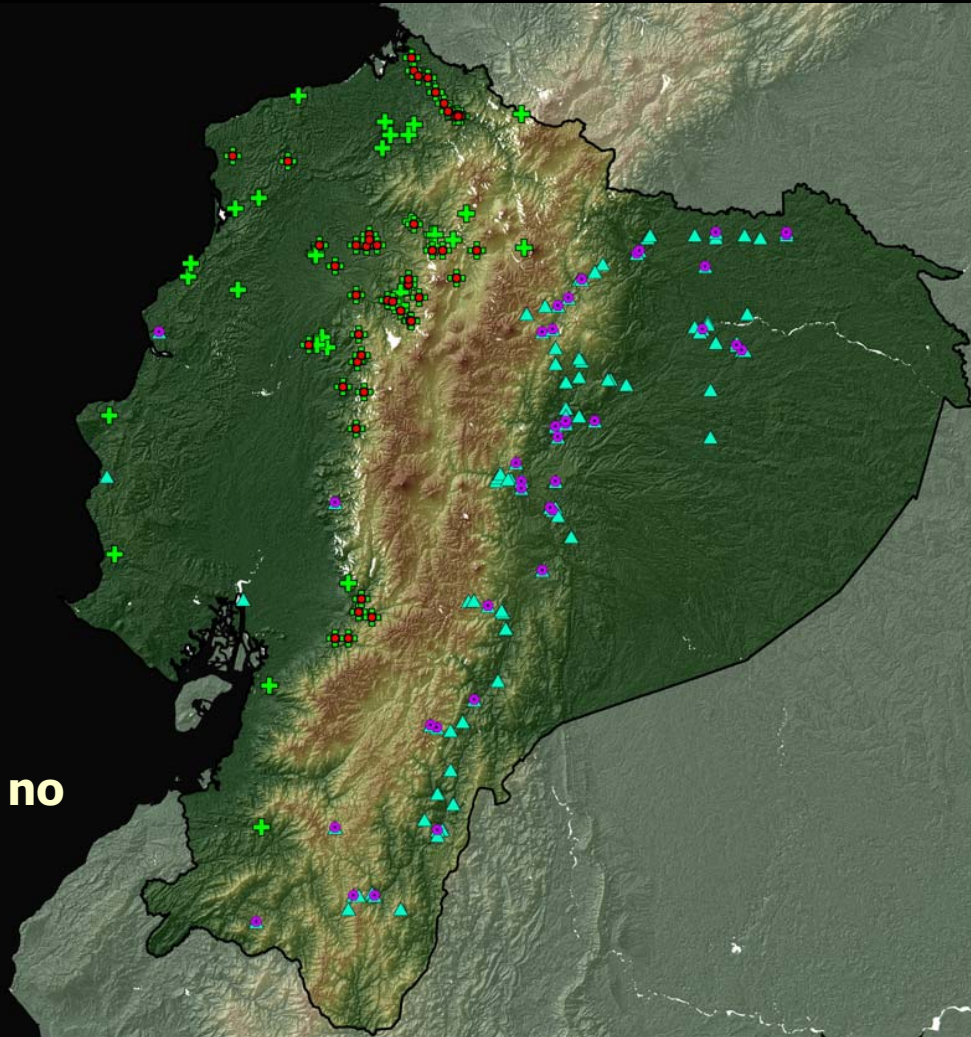
Colecciones de *Anthurium dolichostachyum*



Colecciones de T. Croat de *A. dolichostachyum*
(training positivo)



Colecciones de T. Croat donde no encontró *A. dolichostachyum*
(training negativo 1)



MODELOS CON DIFERENTES ALGORITMOS

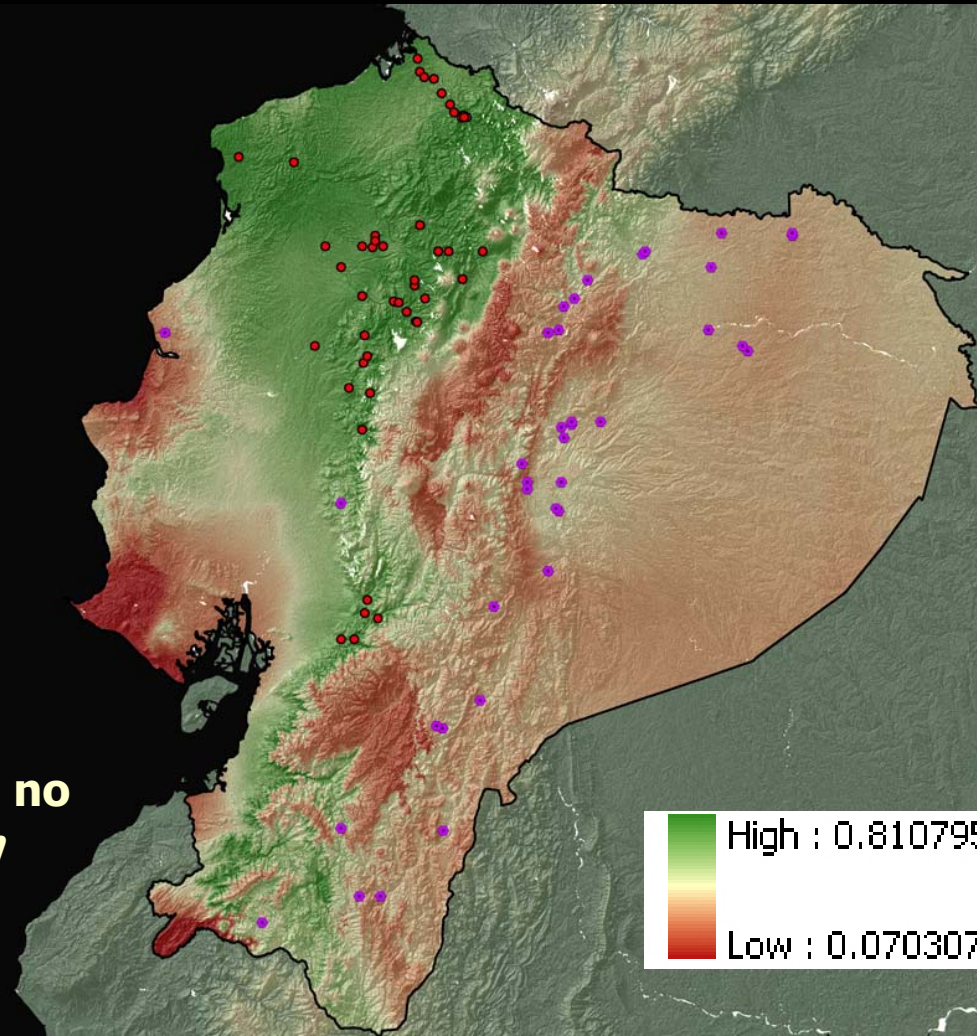
Modelo generado con MARS



**Colecciones de T. Croat de
*A. dolichostachyum***
(training positivo)

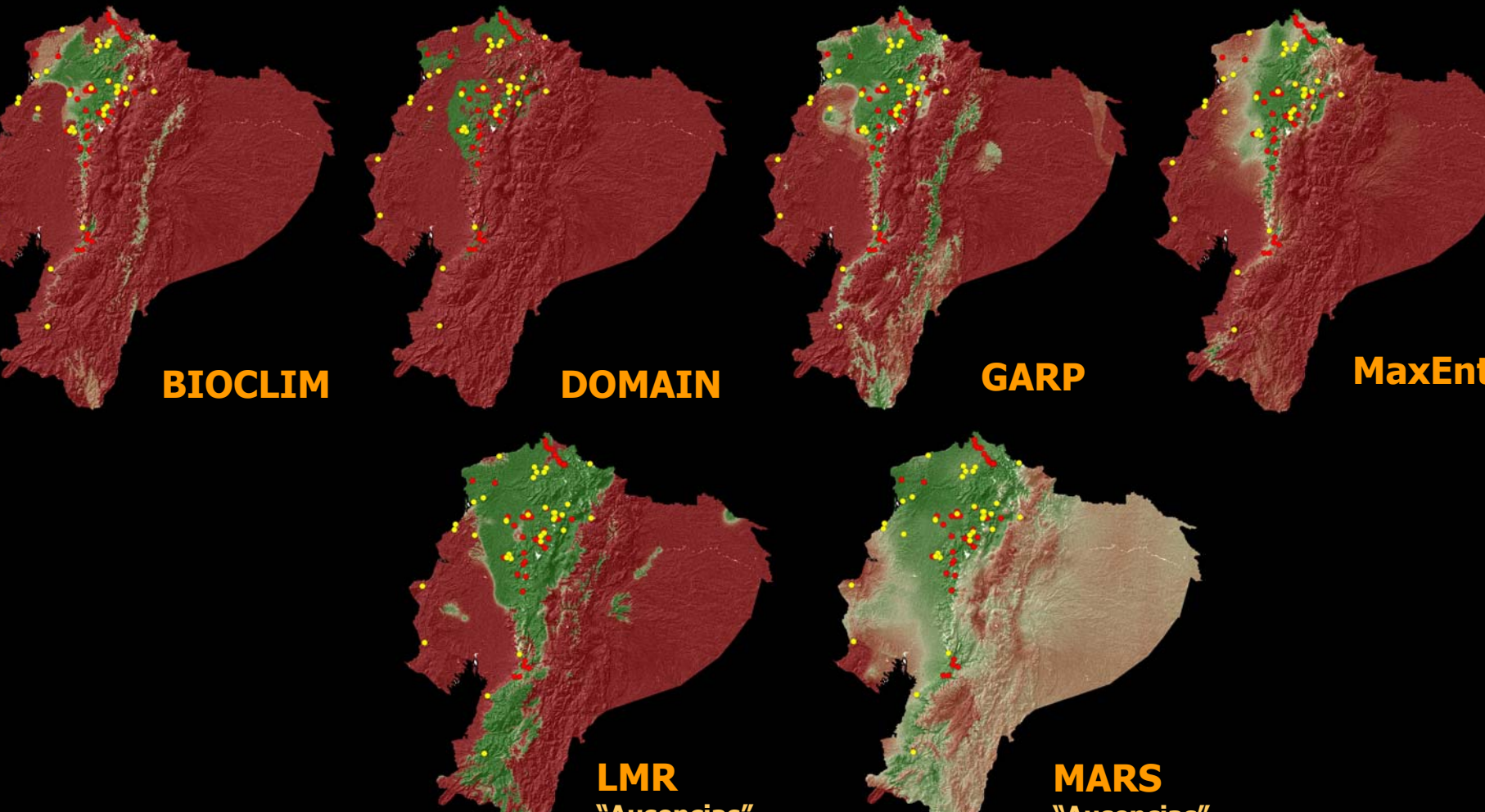


**Colecciones de T. Croat donde no
Encontró *A. dolichostachyum***
(training negativo 1)



MODELOS CON DIFERENTES ALGORITMOS

Métodos discriminativos y descriptivos



GENERACIÓN DE DATOS "TRAINING"

Pseudoausencias



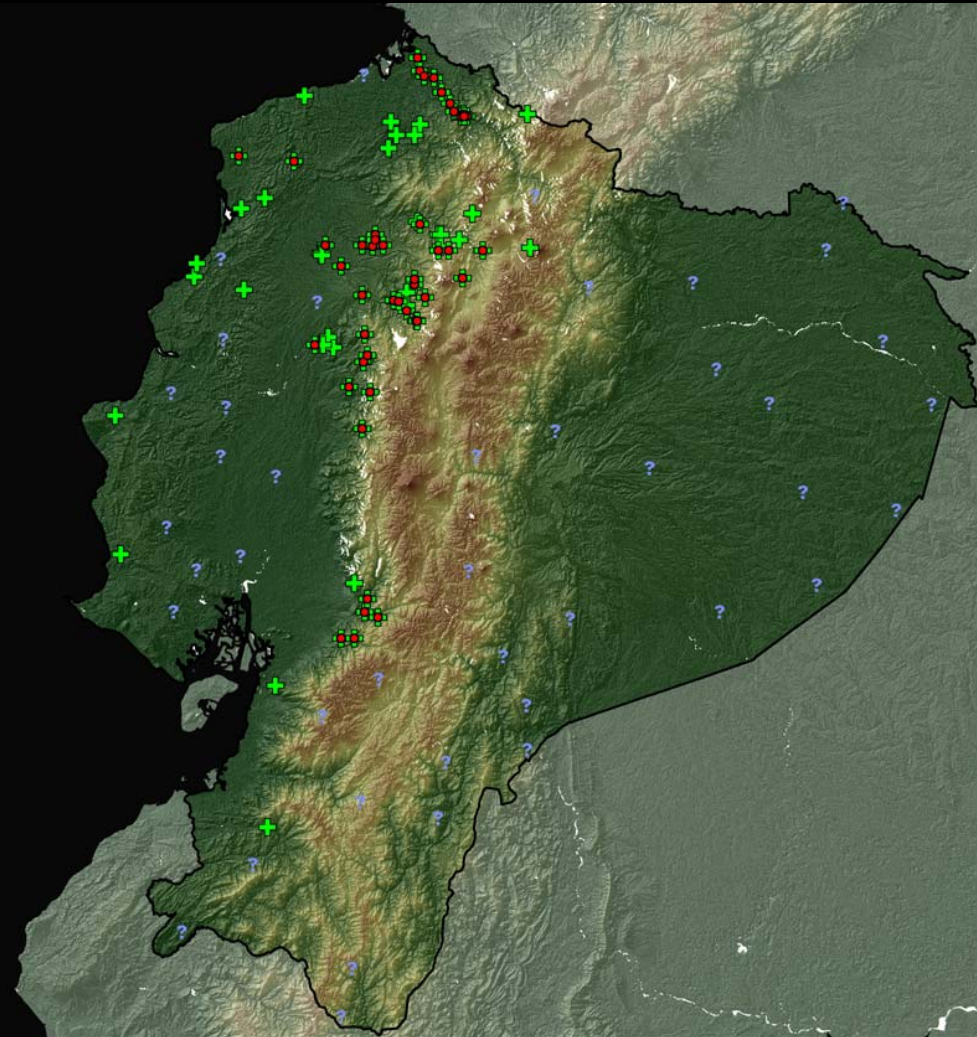
Colecciones de *Anthurium dolichostachyum*



Colecciones de T. Croat de
A. dolichostachyum
(training positivo)




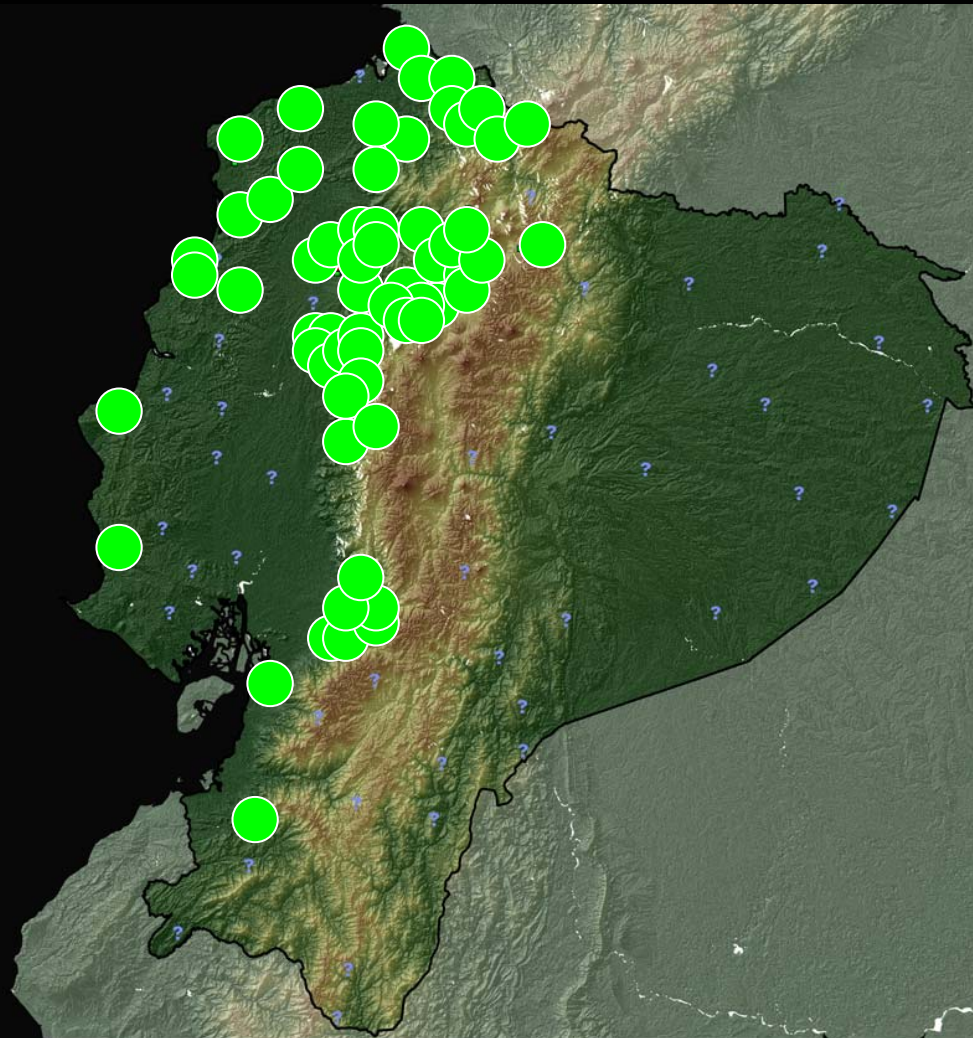
Pseudoausencias
(training negativo 2)



GENERACIÓN DE DATOS "TRAINING"

Pseudoausencias

- ◆ Se realizan modelos con $8n$, $4n$, $2n$ y n pseudoausencias al azar. Se eliminan las situadas a una distancia inferior de **30 km** a cualquier presencia. 
- ◆ Los mejores modelos de idoneidad se obtienen con **$n=1$** , siendo n el número de presencias para esa especie.
- ◆ Los resultados de la mayoría de técnicas estadísticas están **influenciados por la prevalencia** (proporción de la clase diana en la muestra). Tienden a predecir mejor la clase más abundante.



GENERACIÓN DE DATOS "TRAINING"

Pseudoausencias



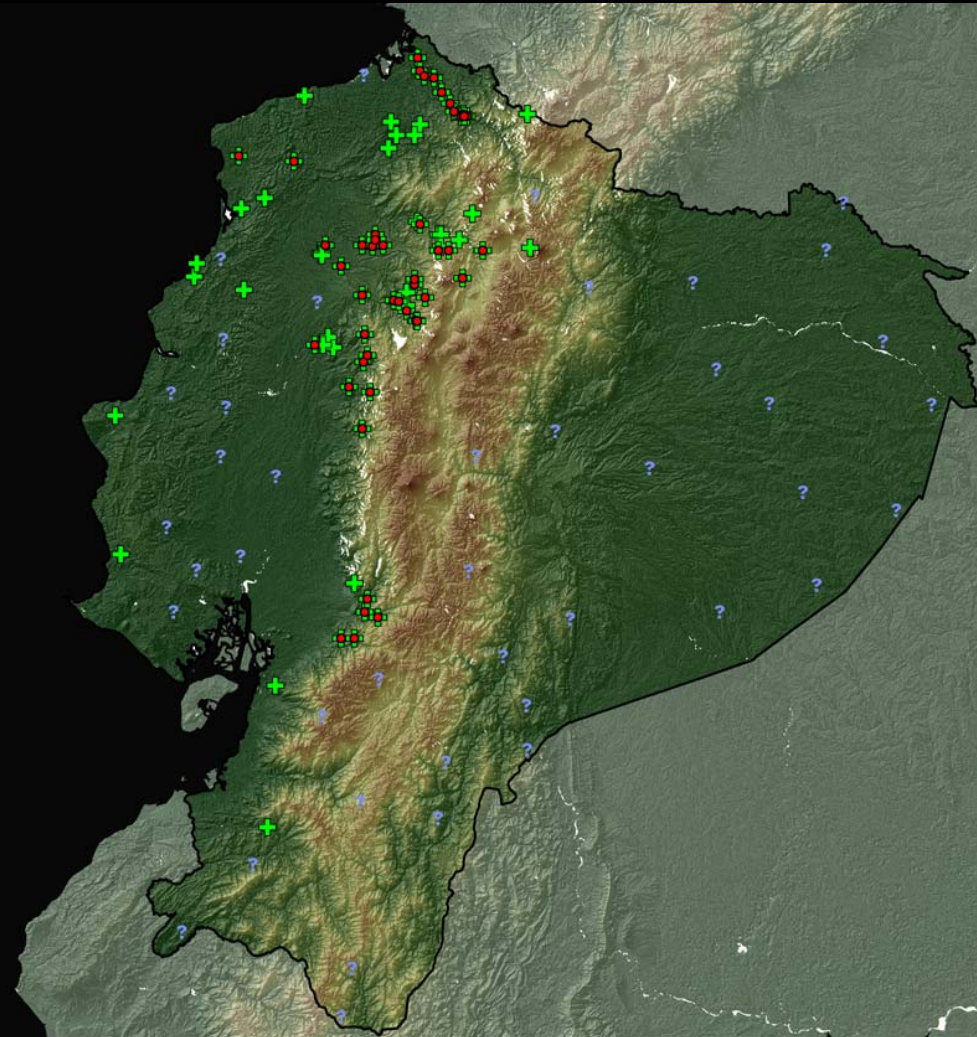
Colecciones de *Anthurium dolichostachyum*



Colecciones de T. Croat de
A. dolichostachyum
(training positivo)



Pseudoausencias
(training negativo 2)



MODELOS CON DIFERENTES ALGORITMOS

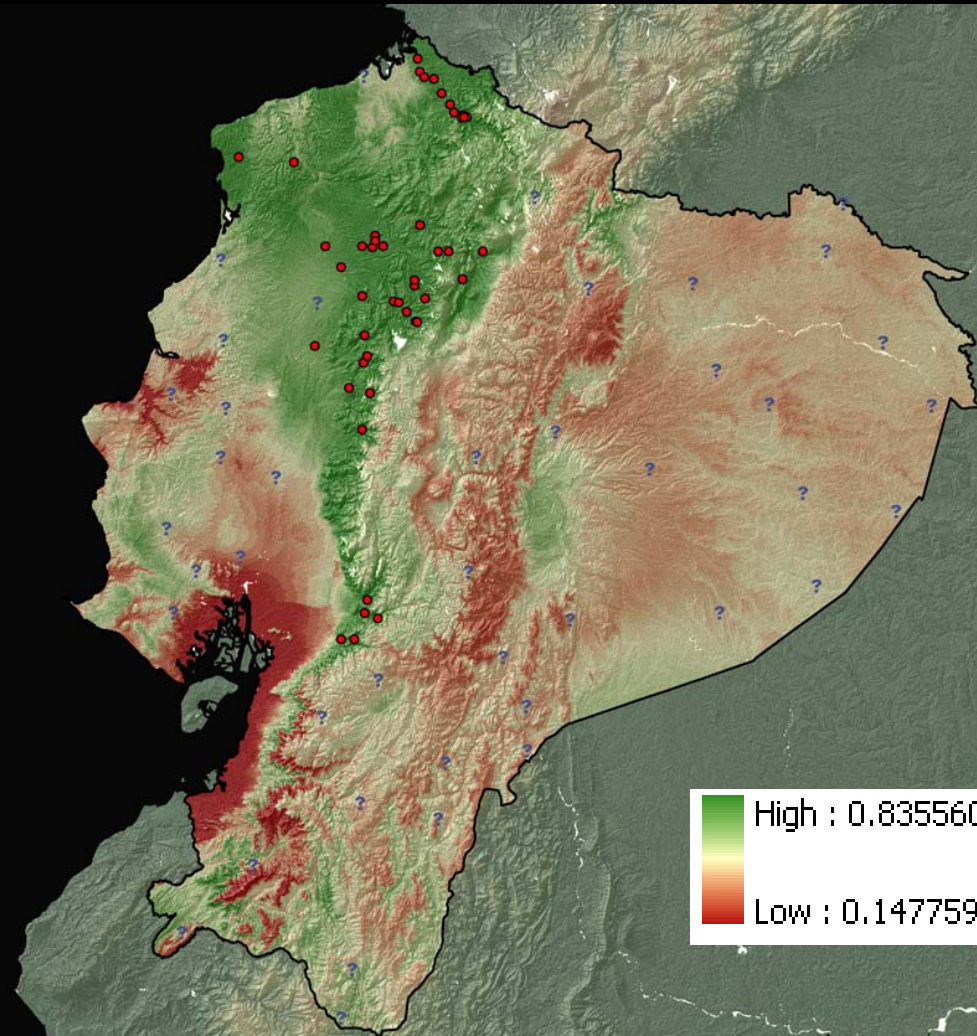
Modelo generado con MARS



**Colecciones de T. Croat de
*A. dolichostachyum***
(training positivo)

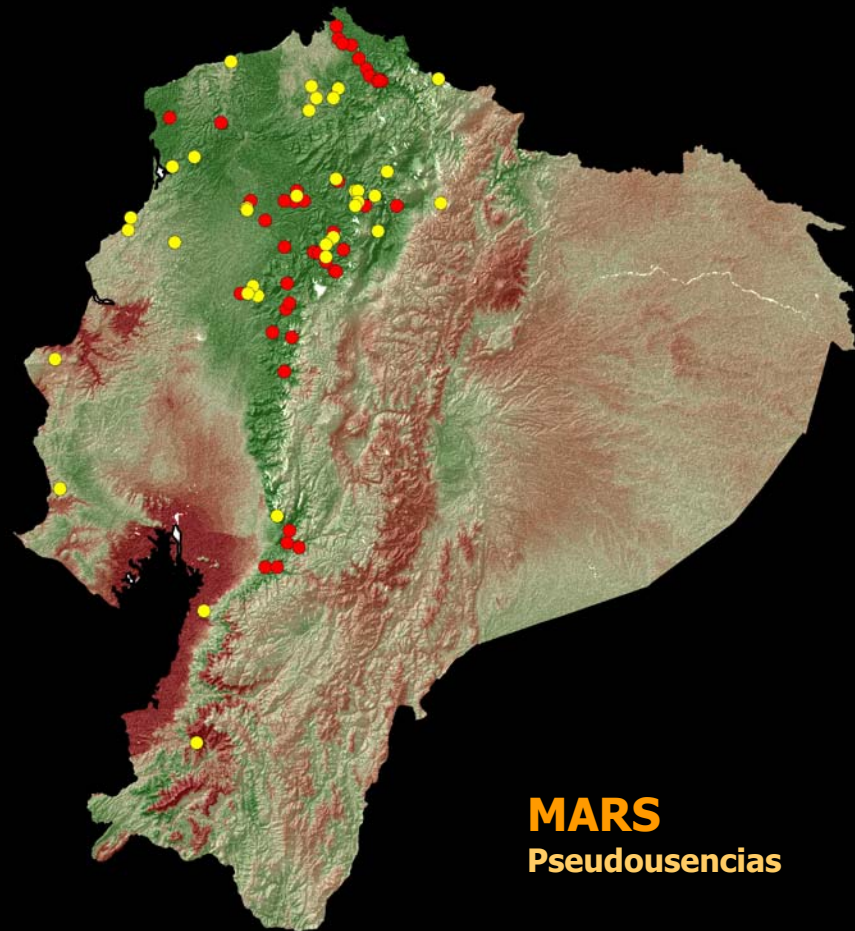
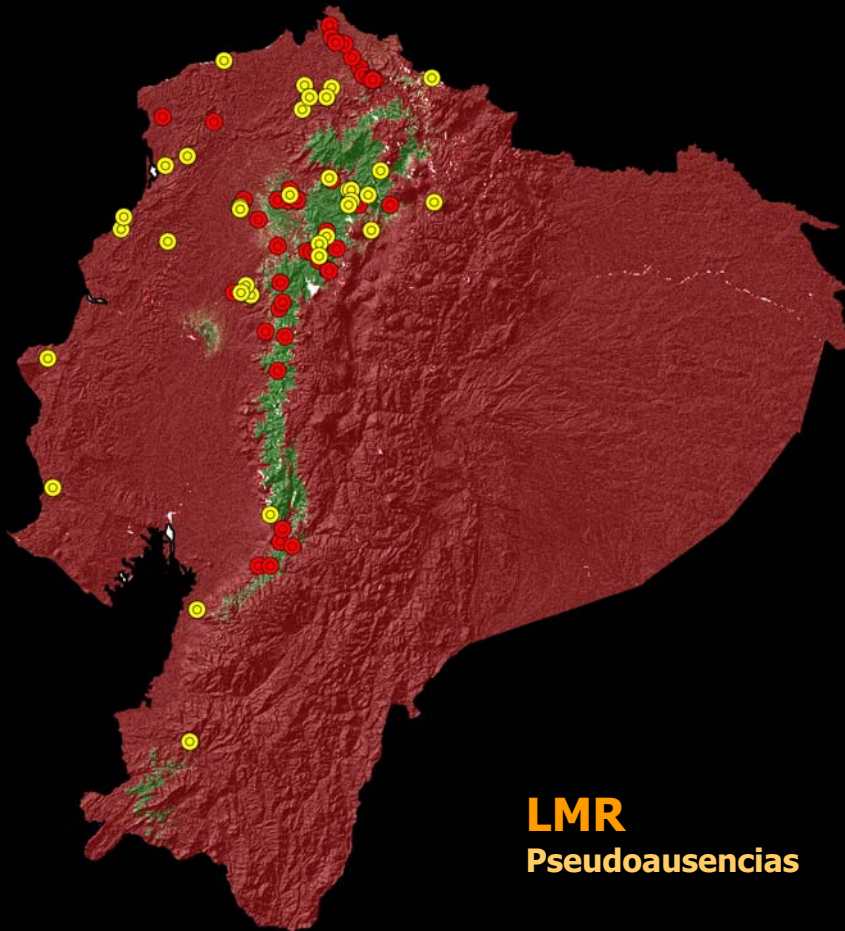


Pseudoausencias
(training negativo 2)



MODELOS CON DIFERENTES ALGORITMOS

Métodos discriminativos



GENERACIÓN DATOS DE "TESTING"

Datos testing



Colecciones de T. Croat del género *Anthurium* donde no encontró *A. dolichostachyum*



Colecciones de *Anthurium dolichostachyum*



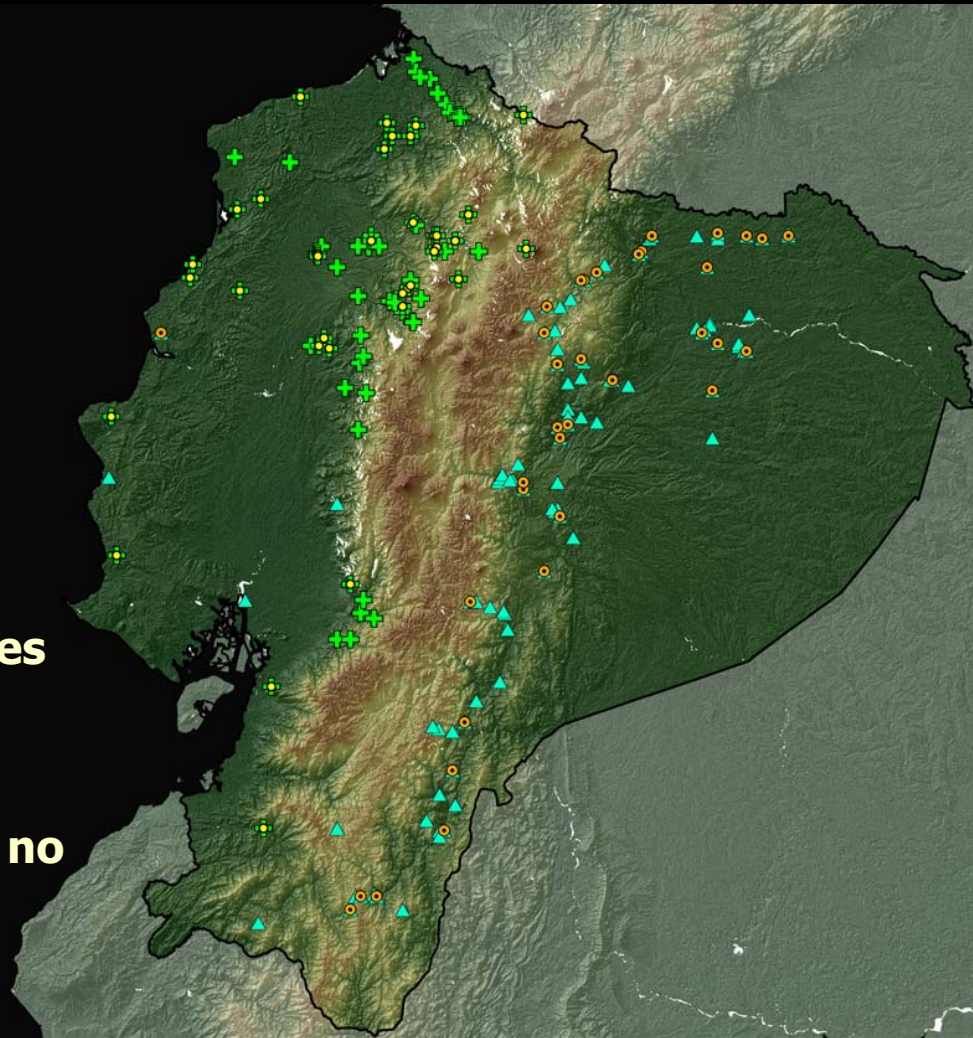
Colecciones de *A. dolichostachyum* de colectores diferentes a T. Croat

(testing positivo)



Colecciones de T. Croat donde no encontró *A. dolichostachyum*

(testing negativo)



DATOS "TRAINIG" Y "TESTING"

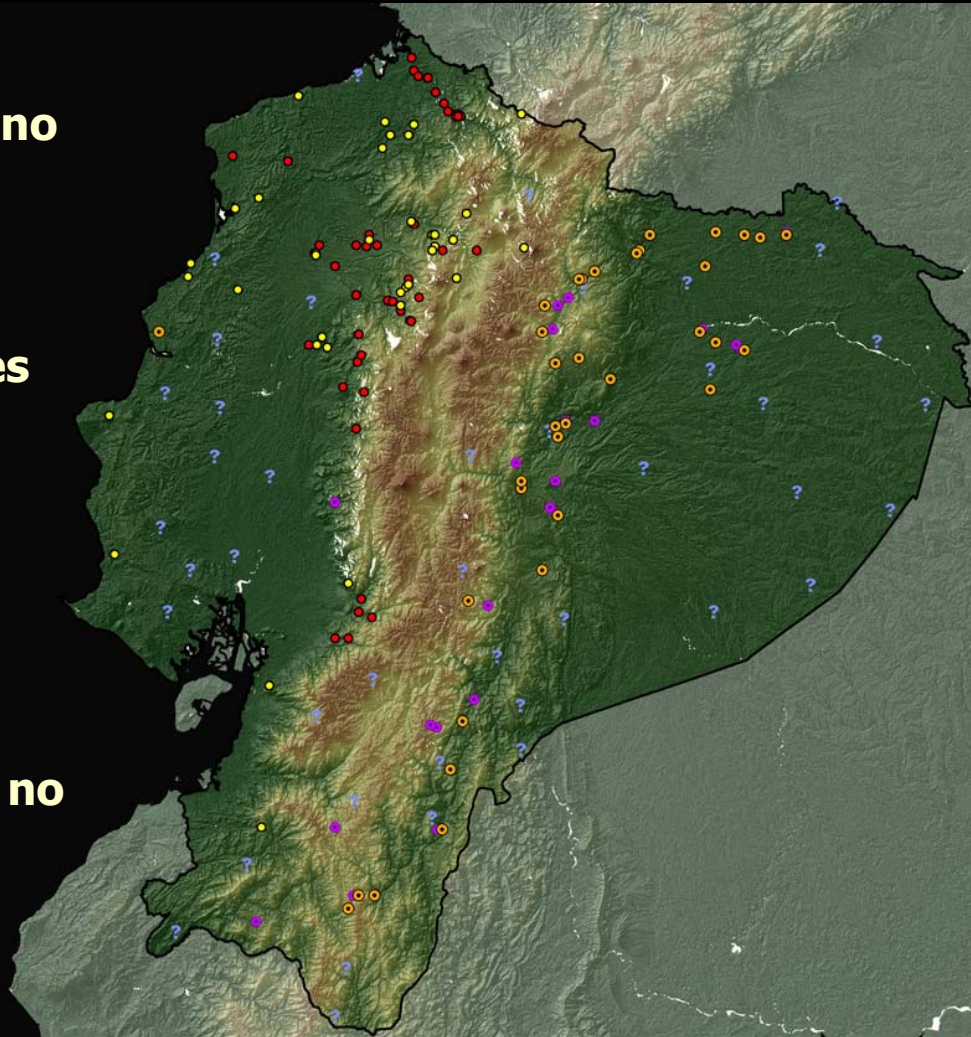
○ Colecciones de T. Croat donde no encontró *A. dolichostachyum* (testing negativo)

● Colecciones de *A. dolichostachyum* de colectores diferentes a T. Croat (testing positivo)

● Colecciones de T. Croat de *A. dolichostachyum* (training positivo 1 y 2)

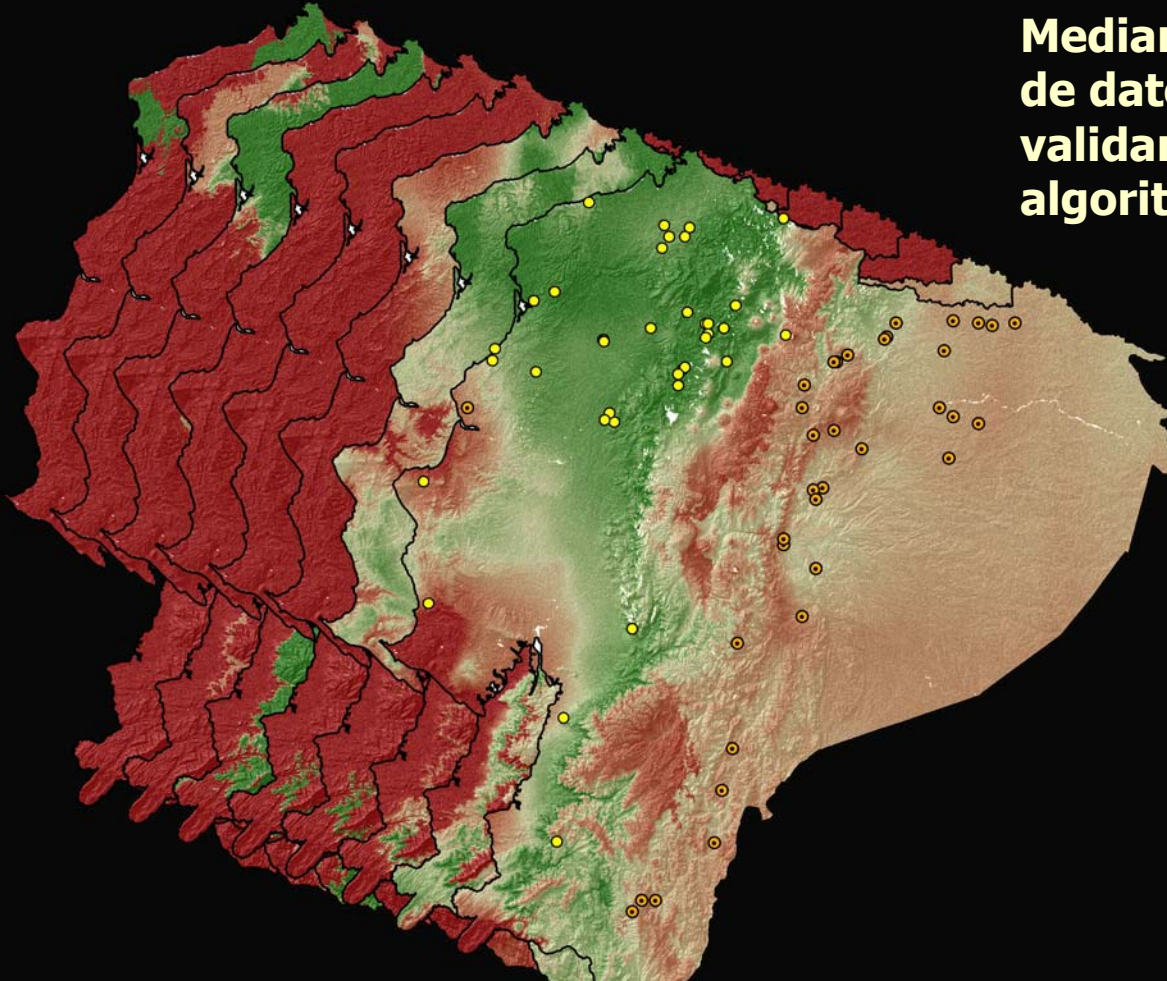
⬡ Colecciones de T. Croat donde no encontró *A. dolichostachyum* (training negativo 1)

? Pseudoausencias (training negativo 2)



VALIDACIÓN DIFERENTES ALGORITMOS

Orden SAMPLE



Mediante el conjunto de datos "testing" validamos todos los algoritmos



VALIDACIÓN DIFERENTES ALGORITMOS

Resultados AUC

ESPECIE	BIOCLIM	BIOCLIM/pa	DOMAIN	DOMAIN/pa	GARP	MaxEnt
<i>A. dolichostachyum</i>	0.799	0.708	0.830	0.722	0.846	0.901
<i>A. harlingianum</i>	0.786	0.750	0.932	0.536	0.813	0.951
<i>A. propinquum</i>	0.786	0.714	0.833	0.714	0.831	0.980
<i>A. truncicolum</i>	0.727	0.694	0.810	0.641	0.627	0.817
<i>A. versicolor</i>	0.587	0.564	0.652	0.628	0.593	0.687

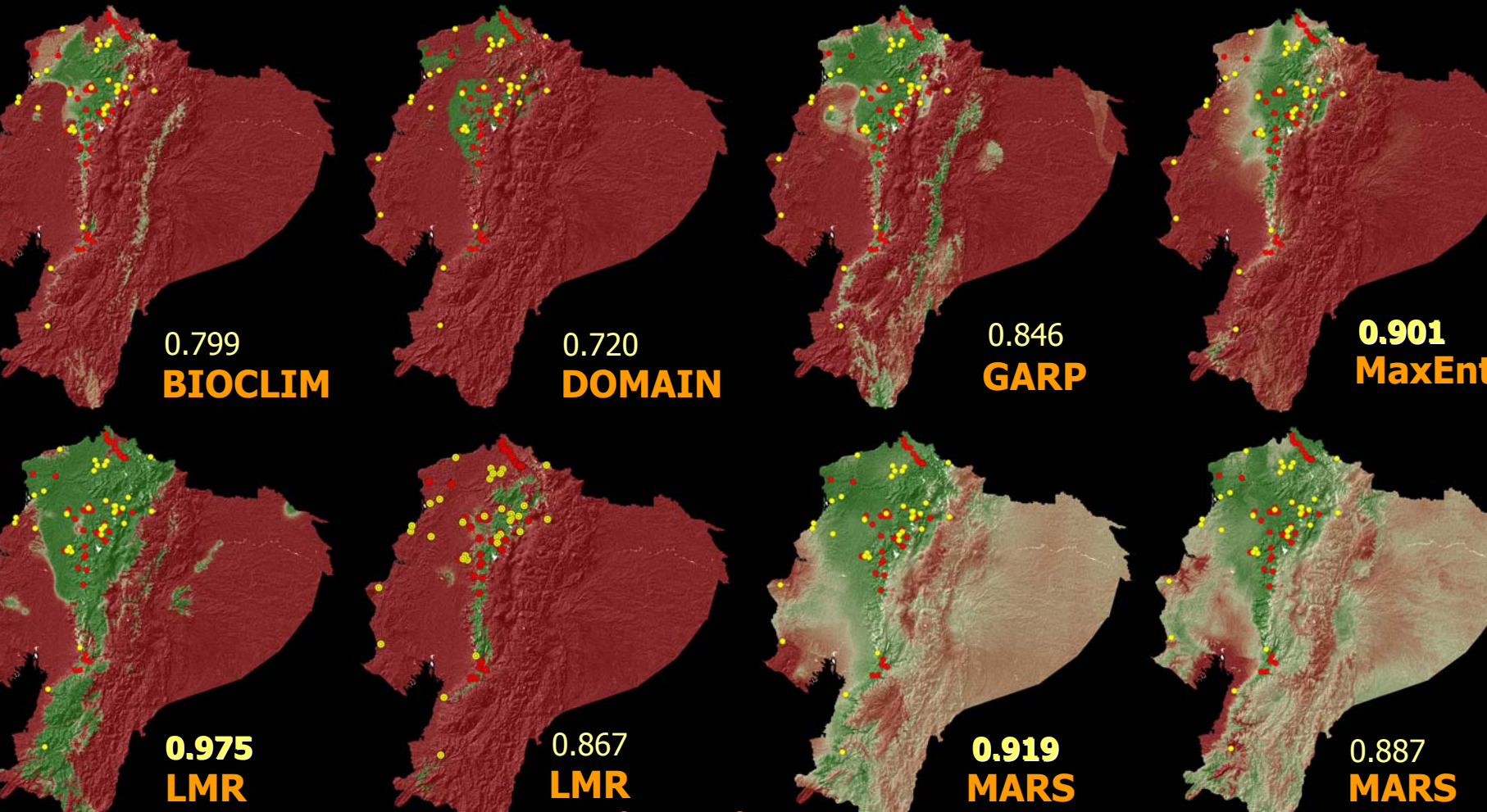
ESPECIE	LMR ausencias	LMR pseudoausencias	MARS ausencias	MARS/pa ausencias	MARS pseudoausencias	MARS/pa pseudoausencias
<i>A. dolichostachyum</i>	0.975	0.867	0.919	0.903	0.887	0.887
<i>A. harlingianum</i>	0.929	0.864	0.968	0.964	0.877	0.782
<i>A. propinquum</i>	0.867	0.844	0.980	0.980	0.969	0.840
<i>A. truncicolum</i>	0.846	0.644	0.732	0.735	0.560	0.565
<i>A. versicolor</i>	0.709	0.740	0.897	0.856	0.631	0.585

VALIDACIÓN DIFERENTES ALGORITMOS

CONCLUSIONES

● Training positivo

● Testing positivo



Agradecimientos

Agradecemos la financiación de la Fundación Banco Bilbao-Vizcaya-Argenteria al proyecto (BIOCON-26). Missouri Botanical Garden proveyó datos sobre las plantas a partir de su base de datos TROPICOS.



<http://www.unex.es/eweb/kraken/index.html>