

Estratificación temporal de *Aedes Aegypti* basada en herramientas geoespaciales y aprendizaje automático

Juan M. Scavuzzo

Facultad de Matemática, Astronomía, Física y Computación
Universidad Nacional de Córdoba

Diciembre de 2018

Motivación de este trabajo: problemática epidemiológica

- *Aedes aegypti* es el principal vector de Dengue, Chikungunya, Zika y Fiebre Amarilla urbana

Motivación de este trabajo: problemática epidemiológica

- *Aedes aegypti* es el principal vector de Dengue, Chikungunya, Zika y Fiebre Amarilla urbana
- Datos de la Organización Mundial de la Salud:

Motivación de este trabajo: problemática epidemiológica

- Aedes aegypti es el principal vector de Dengue, Chikungunya, Zika y Fiebre Amarilla urbana
- Datos de la Organización Mundial de la Salud:
 - 80 millones de personas se infectan de Dengue anualmente

Motivación de este trabajo: problemática epidemiológica

- Aedes aegypti es el principal vector de Dengue, Chikungunya, Zika y Fiebre Amarilla urbana
- Datos de la Organización Mundial de la Salud:
 - 80 millones de personas se infectan de Dengue anualmente
 - 550 mil enfermos requieren hospitalización

Motivación de este trabajo: problemática epidemiológica

- Aedes aegypti es el principal vector de Dengue, Chikungunya, Zika y Fiebre Amarilla urbana
- Datos de la Organización Mundial de la Salud:
 - 80 millones de personas se infectan de Dengue anualmente
 - 550 mil enfermos requieren hospitalización
 - 20 mil personas mueren

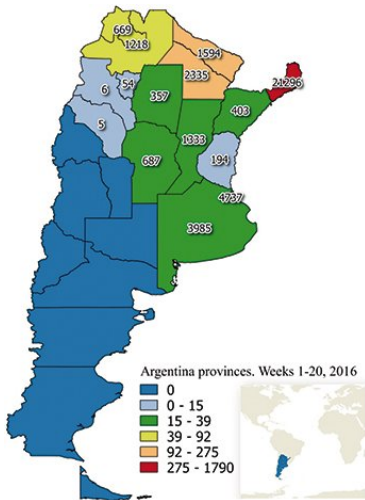
Motivación de este trabajo: problemática epidemiológica

- Aedes aegypti es el principal vector de Dengue, Chikungunya, Zika y Fiebre Amarilla urbana
- Datos de la Organización Mundial de la Salud:
 - 80 millones de personas se infectan de Dengue anualmente
 - 550 mil enfermos requieren hospitalización
 - 20 mil personas mueren
 - 2.500 millones de personas corren riesgo de contraer la enfermedad

Motivación de este trabajo: problemática epidemiológica

- Aedes aegypti es el principal vector de Dengue, Chikungunya, Zika y Fiebre Amarilla urbana
- Datos de la Organización Mundial de la Salud:
 - 80 millones de personas se infectan de Dengue anualmente
 - 550 mil enfermos requieren hospitalización
 - 20 mil personas mueren
 - 2.500 millones de personas corren riesgo de contraer la enfermedad
 - Más de 100 países con transmisión endémica

Motivación de este trabajo: problemática epidemiológica



Motivación de este trabajo: problemática epidemiológica

Características del *Aedes aegypti*

Motivación de este trabajo: problemática epidemiológica

Características del *Aedes aegypti*

- Gran capacidad adaptativa

Motivación de este trabajo: problemática epidemiológica

Características del *Aedes aegypti*

- Gran capacidad adaptativa
- Resistencia a insecticidas

Motivación de este trabajo: problemática epidemiológica

Características del *Aedes aegypti*

- Gran capacidad adaptativa
- Resistencia a insecticidas
- Resistencia de huevos a la desecación

Motivación de este trabajo: problemática epidemiológica

Características del *Aedes aegypti*

- Gran capacidad adaptativa
- Resistencia a insecticidas
- Resistencia de huevos a la desecación
- Presencia en el medio urbano

Motivación de este trabajo: problemática epidemiológica

Características del *Aedes aegypti*

- Gran capacidad adaptativa
- Resistencia a insecticidas
- Resistencia de huevos a la desecación
- Presencia en el medio urbano
- Preferencia de cría en contenedores artificiales

Motivación de este trabajo: sistemas de modelado actuales

Modelar utilizando información satelital

Motivación de este trabajo: sistemas de modelado actuales

Modelar utilizando información satelital

- Información ambiental con alcance regional

Motivación de este trabajo: sistemas de modelado actuales

Modelar utilizando información satelital

- Información ambiental con alcance regional
- Información espacio-temporal

Motivación de este trabajo: sistemas de modelado actuales

Modelar utilizando información satelital

- Información ambiental con alcance regional
- Información espacio-temporal
- Grandes avances en los últimos años

Motivación de este trabajo: sistemas de modelado actuales

Modelar utilizando información satelital

- Información ambiental con alcance regional
- Información espacio-temporal
- Grandes avances en los últimos años

Motivación de este trabajo: sistemas de modelado actuales

Modelar utilizando información satelital

- Información ambiental con alcance regional
- Información espacio-temporal
- Grandes avances en los últimos años

Pero actualmente se utilizan modelos lineales para relacionar las distintas variables!

Motivación de este trabajo

Algunas cuestiones a tener en cuenta

Motivación de este trabajo

Algunas cuestiones a tener en cuenta

- La prevención de las enfermedades en cuestión debe ser a través de control de vectores

Motivación de este trabajo

Algunas cuestiones a tener en cuenta

- La prevención de las enfermedades en cuestión debe ser a través de control de vectores
 - Modelos predictivos!

Motivación de este trabajo

Algunas cuestiones a tener en cuenta

- La prevención de las enfermedades en cuestión debe ser a través de control de vectores
 - Modelos predictivos!
- Será correcto asumir relaciones lineales entre las variables ambientales y la abundancia del vector?

Motivación de este trabajo

Algunas cuestiones a tener en cuenta

- La prevención de las enfermedades en cuestión debe ser a través de control de vectores
 - Modelos predictivos!
- Será correcto asumir relaciones lineales entre las variables ambientales y la abundancia del vector?
 - Modelos no-lineales con... aprendizaje automático!

Motivación de este trabajo

Algunas cuestiones a tener en cuenta

- La prevención de las enfermedades en cuestión debe ser a través de control de vectores
 - Modelos predictivos!
- Será correcto asumir relaciones lineales entre las variables ambientales y la abundancia del vector?
 - Modelos no-lineales con... aprendizaje automático!
- Si es un sistema regional, cómo extrapolo los modelos?

Motivación de este trabajo

Algunas cuestiones a tener en cuenta

- La prevención de las enfermedades en cuestión debe ser a través de control de vectores
 - Modelos predictivos!
- Será correcto asumir relaciones lineales entre las variables ambientales y la abundancia del vector?
 - Modelos no-lineales con... aprendizaje automático!
- Si es un sistema regional, cómo extrapolo los modelos?
 - A través de relaciones entre características ambientales!

Objetivos

Objetivos

- Implementar una herramienta, sencilla, para generar modelos predictivos

Objetivos

- Implementar una herramienta, sencilla, para generar modelos predictivos
- Validar la hipótesis de que "modelos no-lineales son mejores para predecir la oviposición que los lineales"

Objetivos

- Implementar una herramienta, sencilla, para generar modelos predictivos
- Validar la hipótesis de que "modelos no-lineales son mejores para predecir la oviposición que los lineales"
- Proponer una solución a la problemática de escasos de datos que se evidencia al pensar en sistemas regionales de estimación de riesgo

Algunos conceptos importantes

Algunos Conceptos: *Epidemiología Panorámica*

Algunos Conceptos: *Epidemiología Panorámica*

- La teledetección y su capacidad de adquirir información

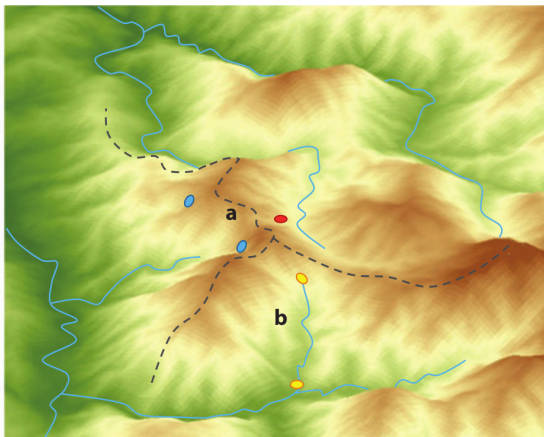
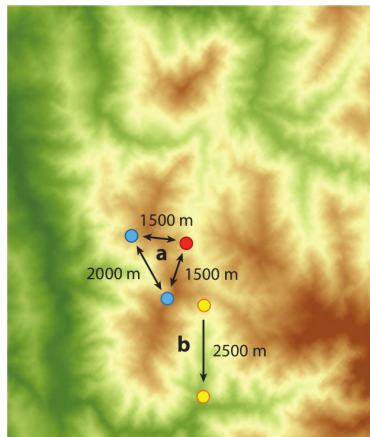
Algunos Conceptos: *Epidemiología Panorámica*

- La teledetección y su capacidad de adquirir información
- Información sobre hábitat de insectos y artrópodos

Algunos Conceptos: *Epidemiología Panorámica*

- La teledetección y su capacidad de adquirir información
- Información sobre hábitat de insectos y artrópodos
- Fuente de datos sobre la distribución espacio-temporal de enfermedades transmitidas por vectores (Pavlovsky)

Algunos Conceptos: *Epidemiología Panorámica*



Algunos Conceptos: *Epidemiología Panorámica*

- Ecología panorámica

Algunos Conceptos: *Epidemiología Panorámica*

- Ecología panorámica
- Focalidad

Algunos Conceptos: *Epidemiología Panorámica*

- Ecología panorámica
- Focalidad
 - Vectores con capacidad de transmisión de la infección

Algunos Conceptos: *Epidemiología Panorámica*

- Ecología panorámica
- Focalidad
 - Vectores con capacidad de transmisión de la infección
 - Vertebrados capaces de funcionar como reservorio de la infección

Algunos Conceptos: *Epidemiología Panorámica*

- Ecología panorámica
- Focalidad
 - Vectores con capacidad de transmisión de la infección
 - Vertebrados capaces de funcionar como reservorio de la infección
 - Huéspedes susceptibles, como humanos o animales domésticos

Algunos Conceptos: *Epidemiología Panorámica*

- Ecología panorámica
- Focalidad
 - Vectores con capacidad de transmisión de la infección
 - Vertebrados capaces de funcionar como reservorio de la infección
 - Huéspedes susceptibles, como humanos o animales domésticos

Algunos Conceptos: *Epidemiología Panorámica*

- Ecología panorámica
- Focalidad
 - Vectores con capacidad de transmisión de la infección
 - Vertebrados capaces de funcionar como reservorio de la infección
 - Huéspedes susceptibles, como humanos o animales domésticos

Epidemiología Panorámica

Algunos Conceptos: *Aprendizaje automático*

*Se dice que un programa de computadora **aprende** de experiencia E con respecto a alguna tarea T y una métrica de rendimiento M , si con la experiencia E se incrementa su rendimiento en la tarea T , medida por M .*

Tom Mitchell, 1997 [?]

Algunos Conceptos: *Aprendizaje automático (ML)*

- Enfoque empírico efectivo para *regresiones y clasificaciones*
- Distintos métodos:
 - Supervisados: Regresiones Lineales, SVMs, ANNs, DTRs...
 - No-supervisados: K-NNs, K-means, PCA...
 - Semi-supervisados
- Usado en muchos ámbitos:
 - Académico
 - Industrial
 - Gubernamental

Algunos Conceptos: *Métodos Supervisados*

- Aprenden a través de pares de ejemplos (X, Y_{verd})
- Conjuntos de entrenamiento y validación
- Evitar *overfitting*
- Ajuste de hiperparámetros...

Algunos Conceptos: *Parámetros* y *Hiperparámetros*

- Los algoritmos de ML poseen *parámetros* e *hiperparámetros*
 - Los hiperparámetros definen el comportamiento durante el proceso de entrenamiento
 - Los parámetros se ajustan para definir el modelo luego del entrenamiento

Algunos Conceptos: *Parámetros* y *Hiperparámetros*

- Los algoritmos de ML poseen *parámetros* e *hiperparámetros*
 - Los hiperparámetros definen el comportamiento durante el proceso de entrenamiento
 - Los parámetros se ajustan para definir el modelo luego del entrenamiento

Algunos Conceptos: *Ajuste de hiperparámetros*

Supongamos una **Ridge Regression**:

$$\hat{y}(w, x) = w_0 + w_1x_1 + \dots + w_px_p$$

$$\min_w ||Xw - y||_2^2 + \alpha ||w||_2^2$$