Categorización ambiental para la evaluación de la dinámica espacio-temporal de criaderos artificiales de *Aedes aegypti* en Clorinda, Argentina (2012-2013)

Guillermo Albrieu Llinás

Presentado ante la Facultad de Matemática, Astronomía y Física y el Instituto de Altos Estudios Espaciales Mario Gulich, como parte de los requerimientos para la obtención del grado de

MAGÍSTER EN APLICACIONES ESPACIALES DE ALERTA Y RESPUESTA TEMPRANA A EMERGENCIAS

Córdoba, 2020

©IG - CONAE/UNC 2020 © FaMAF - UNC 2020

Directora: Dra. Marta Contigiani Co–Director: Dr. Carlos Marcelo Scavuzzo

ÍNDICE DE CONTENIDOS

RESUMEN	1
ABSTRACT	2
1 - INTRODUCCIÓN	3
1.1 - Introducción de Aedes aegypti en América	4
1.2 - Características biológicas de Aedes aegypti	5
1.3 - Estrategias de control	9
1.4 - Sensado remoto y aplicaciones en ecología de mosquitos vectores	10
1.4.1 - Generalidades sobre teledetección	10
1.4.2 - Sensado remoto y epidemiología espacial	
1.5 – Construcción de la hipótesis y objetivos	
2 - MATERIALES Y MÉTODOS	19
2.1 - Área de estudio	19
2.2 - Descripción de los datos	20
2.3 - Clasificación no-supervisada de imagen SPOT	23
2.4 - Generación de capas de variables macro-ambientales	25
2.5 - Agrupamiento ambiental de barrios	27
2.6 - Modelo lineal generalizado	28
3 – RESULTADOS	32
4 - DISCUSIÓN	45
BIBLIOGRAFÍA	52
ANEXO I – Tabla de diseño de muestreos	59
ANEXO II – Tabla de datos usados en el modelo lineal generalizado	60
ANEXO III – Tabla de valores medios de variables ambientales por barrio	66
ANEXO IV – Matriz de correlación de variables ambientales por barrio	68
ANEXO V – Matriz de correlación de variables ambientales por barrio, sel con r<0.7	eccionadas 69
ANEXO VI – Tabla de variables ambientales seleccionadas por barrio, cer estandarizadas	ntradas y 70
ANEXO VII - Script de R para los diferentes pasos de la metodología	72

RESUMEN

La identificación de puntos criadero claves de Aedes aegypi en áreas urbanas es crucial para el diseño racional de estrategias de control contra este vector de enfermedades. El sensado remoto y los sistemas de información geográfica ofrecen herramientas de gran valor para el mapeo de hábitats favorables en un área determinada. De todos modos, la predicción de la presencia de una especie mediante mapas de probabilidad de distribución aprovechando muestreos entomológicos transversales es de utilidad limitada para las autoridades locales. El objetivo del presente estudio fue examinar la evolución temporal del número de casas infestadas con estados inmaduros de Ae. aegypti en barrios de la ciudad de Clorinda (Formosa) y explorar la utilidad del agrupamiento ambiental de estos barrios, generado con información provista por variables de sensado remoto, a la hora de explicar el comportamiento temporal diferencial observado. Se realizaron muestreos entomológicos entre los años 2012 y 2013, registrando el número de casas con recipientes conteniendo larvas o pupas de Ae. aegypti. Luego se utilizó una imagen tomada por el satélite SPOT 5 para obtener variables de cobertura de tierra, las cuales fueron usadas para un agrupamiento tipo k-means que asignó los barrios estudiados a categorías o clusters ambientales. Finalmente, se ajustó un modelo lineal generalizado que mostró que el número de casas positivas estuvo conjuntamente afectado por la interacción entre uno de los clusters ambientales y el año de muestreo. El número de casas positivas en ese cluster fue 9.5 veces mayor en 2013 que en 2012 (P<0.005, SE=0.29), si bien no se observaron otras diferencias estadísticamente significativas. Lejos de resolver de manera acabada un problema tan complejo como la circulación de enfermedades de transmisión vectorial, este flujo de trabajo permite sintetizar la información contenida en variables macro-ambientales, en predictores más sencillos que agrupan objetivamente las unidades administrativas de una ciudad. Este agrupamiento demuestra cierta utilidad a la hora de identificar áreas problemáticas al momento de tomar decisiones operativas.

ABSTRACT

The identification of *Aedes aegypti* breeding hotspots in urban areas is crucial for the rational design of control strategies against this disease vector. Remote sensing and geographic information systems offer valuable tools for mapping habitat suitability of a given area. However, predicting species occurrences by means of probability distribution maps based on transversal entomological surveys has limited utility for local authorities. The aim of the present study was to carefully examine the temporal evolution of the number of houses infested with immature stages of Ae. aegypti in neighbourhoods from Clorinda City (Formosa) and to explore the value of producing environmental clusters generated with information provided by remotely sensed variables to explain the observed differential temporal behaviour. Entomological surveys were conducted between 2011 and 2013 throughout the town registering the number of houses with containers harbouring immature stages of Ae. aegypti. A SPOT 5 satellite image was used to obtain land cover variables, which were subsequently submitted to k-means partitioning for grouping neighbourhoods into four environmental clusters. Finally, a generalized linear model was fitted showing that the number of houses found to be positive for Ae. aegypti was jointly affected by the interaction between environmental clusters and the year of sampling. Moreover, the number of positive houses in one of the clusters was 9.5 times higher (P<0.005, SE=0.29) in 2013 than in 2012, but we did not observe any other statistically significant differences. Far from fully solving a problem as complex as the circulation of vector-borne diseases, this workflow allowed us to synthesize the information contained in macro-environmental variables, in simpler predictors that objectively group the administrative units of a city. This approach proved to be useful in identifying problem areas when operational decisions are needed.

1 - INTRODUCCIÓN

Aedes (Stegomyia) aegypti (L.) (Diptera, Culicidae) es el principal vector de los virus Dengue, Chikungunya y Zika (Liang et al., 2015). Según la Organización Mundial de la Salud (OMS), estos arbovirus (virus transmitidos por artrópodos) infectan a más de 300 millones de personas por año en todo el mundo, causando más de 20000 muertes (OMS, 2016). Durante 2009, un brote de Dengue (género Flavivirus) de amplia distribución en Argentina causó 5 casos fatales y más de 27000 infecciones no letales en 13 provincias. Desde ese episodio siguieron reportándose casos eventuales hasta la ocurrencia del mayor brote registrado en la historia del país, durante el año 2016, con casi 77000 casos confirmados. Durante la temporada 2019-2020 ya se registraron casi 45000 infecciones (datos provistos por el Ministerio de Salud de la Nación). Además, Ae. aegypti está involucrado en las epidemias explosivas de Chikungunya (género Alphavirus) (Higgs et al., 2015) y Zika (género Flavivirus) (Gatherer y Kohl, 2016), eventos que refuerzan la importancia de su rol como vector de enfermedades de creciente circulación en las Américas y el resto del mundo. Lamentablemente, hasta el momento no contamos con estrategias antivirales o vacunas eficientes para la prevención de estos virus. Desde abril de 2016, una vacuna contra Dengue producida por Sanofi Pasteur (Dengvaxia) fue implementada en 19 países, incluso con recomendación de la Organización Mundial de la Salud, a pesar de reiterados esfuerzos de la comunidad científica en comunicar los riesgos implicados a esta estrategia de inmunización asistida. Tal como habían alertado estos estudios, la aplicación de la vacuna tuvo efectos adversos (mayor virulencia durante las infecciones) en personas que no habían estado nunca expuestas a ningún subtipo de virus Dengue (Aguiar y Stollenwerk, 2017). De hecho, la utilización de la vacuna de Sanofi estaría relacionada a los más de 800 muertos durante la epidemia de Dengue que azota a Filipinas durante 2019 (datos brindados por el Boletín Epidemiológico de Córdoba). En este escenario, la prevención de este tipo de enfermedades infecciosas debe contar con bases sólidas en la reducción o eliminación del contacto entre mosquitos vectores y humanos.

1.1 - Introducción de Aedes aegypti en América

Esta especie de mosquito habría sido introducida desde África occidental hacia el "Nuevo Mundo" entre los siglos XV y XVII, a bordo de barcos destinados al tráfico de esclavos (Christophers, 1960). Como hipótesis adicional, se propuso una primera invasión a Portugal y España en naves europeas, antes de alcanzar el continente americano (Tabachnick, 1991). Gracias a la previa retención evolutiva de rasgos domésticos a partir de una forma originalmente silvestre, este mosquito pudo adaptarse a completar el desarrollo de sus estados inmaduros (larvas y pupas) en los recipientes de almacenamiento de agua transportados a larga distancia en esos barcos. En la actualidad, *Ae. aegypti* está presente en todas las zonas templadas, tropicales y subtropicales del mundo, incluyendo casi todos los países del continente americano, dentro de los límites aproximados de 35° de latitud norte y sur (**Figura 1**). Una barrera importante para la dispersión y establecimiento de *Ae. aegypti* en determinadas regiones es la altitud, ya que no se lo encuentra por encima de los 2.400 m sobre el nivel del mar (Nelson, 1986).



Figura 1. Distribución potencial de *Aedes aegypti* basada en condiciones climáticas actuales. 0: condiciones desfavorables; 1: condiciones favorables. Tomado de Kamal et al. (2018).

En términos generales, las especies de mosquitos que utilizan recipientes con agua para su reproducción son responsables de drásticas invasiones y expansiones de rango geográfico (Lounibos, 2002). En nuestro país, como en casi toda América, Ae. aegypti se consideró erradicado luego de una campaña de control, masiva y militarizada, llevada adelante entre las décadas de 1950 y 1960. Gracias a un estudio genético-poblacional de reconstrucción del patrón de colonización en Argentina a partir de países limítrofes, Albrieu-Llinás et al. (2012) observaron historias evolutivas diferentes entre 22 poblaciones analizadas. Por un lado se detectaron linajes ancestrales hacia el Noroeste del país y Bolivia y el Noreste y Paraguay, probablemente representantes de relictos sobrevivientes a la campaña de control, a partir de los cuales las poblaciones se habrían reestablecido manteniendo una diversidad genética relativamente alta, dando cuenta de la ineficiencia a largo plazo de la campaña de erradicación. En otras palabras, sería reflejo de la introducción original de linajes independientes, que luego fueron fragmentados y finalmente restaurados a partir de relictos genéticamente polimórficos. En cambio, la campaña de control habría sido más eficiente hacia el Este y el Litoral, donde la diversidad genética original fue prácticamente erradicada. La ocupación actual de la especie en esta región está representada por una única variante genética que muestra un patrón de expansión de rango geográfico reciente, colonizando una vasta región en un corto periodo. En este sentido, la emergencia y reemergencia de enfermedades transmitidas por Ae. aegypti en nuestro país pueden atribuirse, al menos en parte, a la pérdida de interés en el monitoreo y control de mosquitos después de la campaña masiva de erradicación. Esta situación habría permitido el mantenimiento de una alta variabilidad genética en algunas provincias y una rápida recolonización a larga distancia de nuevos linajes, en otras.

1.2 - Características biológicas de Aedes aegypti

Los ciclos de vida y hábitos ecológicos de los insectos son sumamente complejos y desconocerlos a la hora de abordar su monitoreo y control puede llevar a cometer graves errores. Al igual que el resto del orden Diptera, los mosquitos sufren metamorfosis completa durante su desarrollo (**Figura 2**). Los huevos, con una longitud menor a 1 mm, son inicialmente de color blanco pero se tornan negros durante el desarrollo del embrión, el cual evoluciona en óptimas condiciones de temperatura y humedad en un lapso de 2 a 3 días. Las larvas emergentes

atraviesan cuatro estadios, seguidos del estado de pupa. En este último estado inmaduro, los individuos no necesitan alimentarse activamente y, entre los 28°C y los 32°C, completan su desarrollo hasta la emergencia del adulto en 1 a 3 días. El desarrollo de *Ae. aegypti* se completa, en óptimas condiciones de temperatura y alimentación, en 7 a 10 días (Nelson, 1986).



Figura 2. Infografía del ciclo de vida de *Aedes aegypti* tomada del material libre de divulgación de la Organización Panamericana de la Salud (<u>www.paho.org</u>).

El individuo adulto es de color negro o castaño, con características que lo distinguen claramente: los pedicelos y ápices de los palpos tienen escamas blanco-plateadas, con escamas similares sobre el clípeo que parecen formar dos manchas sub-medianas o una sola; la superficie anterior del fémur medio se caracteriza por la presencia de una delgada línea ventro-mediana de escamas blancas, y los esternos abdominales III y V están también cubiertos con escamas blancas (Savage y Smith, 1995). En otras palabras, la manera más simple de identificar la especie en un contexto doméstico es mediante el reconocimiento de sus manchas blancas en patas y dorso.

Normalmente tanto machos como hembras requieren carbohidratos para su nutrición, generalmente disponibles en forma de néctar en flores o frutos, aunque también pueden utilizar

otras fuentes de azúcar. Ambos sexos poseen un aparato bucal modificado llamado probóscide, pero las hembras cuentan con una modificación aún mayor que les permite perforar la piel de vertebrados y así penetrar vasos sanguíneos. Si bien pueden alimentarse de sangre de distintos hospedadores, tienen una marcada preferencia por la humana. Los nutrientes que aporta la sangre se utilizan principalmente para el desarrollo de los huevos; sin embargo, también pueden usarlos como única fuente de energía en circunstancias determinadas. Por ejemplo, Edman et al. (1992) analizaron el comportamiento de hembras de Ae. aegypti en Tailandia y observaron que rara vez se alimentaban de azúcar en zonas tropicales, sino que basaban su alimentación casi completamente en la sangre de vertebrados. En un estudio reciente se demostró que la composición de la microbiota intestinal está fuertemente modulada por la ingesta sanguínea y tiene efectos observables sobre la capacidad de transmitir patógenos (Muturi et al., 2018). En general, las hembras adultas prefieren alimentarse a la luz del día, con dos pulsos de actividad manifiestos: uno durante las primeras horas de la mañana y otro por la tarde antes del anochecer, aunque dentro de las viviendas humanas suelen hacerlo a cualquier hora. A lo largo de un mismo día pueden picar muchas veces y a varias personas, comportamiento que favorece la transmisión de virus y la ocurrencia de epidemias en periodos muy cortos (Gubler, 1998). Además, sólo las hembras se dispersan activamente, 50 a 400 metros de distancia durante su ciclo de vida, según distintas estimaciones (Reiter et al., 1995; Edman et al., 1998). Sin embargo, cuando no encuentran alimento y/o lugar apropiado para la oviposición, las hembras grávidas suelen volar distancias mayores, de hasta 800 metros (Honório et al., 2003). Dado que la distancia de vuelo de Ae. aegypti es un factor crucial en la dinámica de transmisión del dengue y es función de la disponibilidad de sitios de ovipostura, algunos autores sostienen que la eliminación de criaderos en áreas con circulación viral aumentada tendría el efecto no deseado de inducir la dispersión hacia nuevas áreas, aumentando así la diseminación del virus (Reiter, 1996; Edman et al., 1998; Brown et al., 2017).

Después de una alimentación sanguínea, cada hembra puede poner aproximadamente 100 huevos. Los ovarios comienzan a desarrollarse en respuesta a la estimulación de receptores de distensión en el estómago y, además, la ingesta sanguínea activa la liberación de hormonas esenciales para el desarrollo de huevos (Vogel et al., 2015). Si bien se ha observado que hembras de *Ae. aegypti* pueden alimentarse varias veces antes de una postura de huevos (Scott et al., 1993), una vez superado cierto umbral de distensión, el impulso de alimentación sanguínea se ve suprimido (Gwadz, 1969). En condiciones óptimas de temperatura y humedad el intervalo entre

la alimentación sanguínea y la postura de huevos es de tres días y, después de la postura, la hembra puede alimentarse nuevamente. Los huevos de un mismo ciclo gonadotrófico son distribuidos por cada hembra en varios recipientes naturales y/o artificiales dentro y alrededor de las viviendas humanas, aumentando así el éxito en la eclosión y maduración de las larvas frente a un cambio meteorológico desfavorable. Luego del desarrollo del embrión, los huevos son capaces de resistir condiciones extremas de temperatura y desecación con sobrevidas de siete meses a un año. La mayor parte de cada postura es de eclosión rápida, mientras un porcentaje reducido constituye los llamados huevos resistentes, inactivos o residuales, capaces de largas sobrevidas (Nelson, 1986; Reiter, 1996). Esta gran resistencia de los huevos a la sequía es uno de los principales obstáculos para el control de estos mosquitos, ya que les permite perdurar, incluso ser transportados a grandes distancias, en recipientes secos hasta que las condiciones ambientales vuelven a ser favorables. Por lo tanto, la eliminación de mosquitos adultos y de larvas en una localidad no evita el surgimiento de los adultos emergidos de los huevos que permanecieron silentes durante varios meses en receptáculos secos.

La duración de los estadios larvales depende de la temperatura, disponibilidad del alimento y densidad de larvas en un recipiente. En condiciones óptimas, el período larval desde la eclosión hasta el desarrollo de pupa es de tan solo 5 días (Marquardt, 2004). Por lo tanto, cualquier recipiente natural o artificial en los que se acumula agua en forma accidental o deliberada se convierte en un potencial criadero. Algunos recipientes artificiales creados por el hombre para el almacenamiento de agua pueden servir como criaderos para la especie: tanques elevados (para almacenamiento a largo plazo) y recipientes a nivel del suelo (para uso diario) tales como barriles, baldes, tinajas, etc. Los recipientes desechados son los criaderos más importantes y entre ellos se destacan los neumáticos usados, muy comunes en el ambiente urbano, que pueden acumular cantidades suficientes de agua en oscuridad como para constituir uno de los hábitats preferidos para la oviposición. En países latinoamericanos los neumáticos usados no se descartan, debido a su alto valor comercial; por el contrario, el frecuente intercambio nacional e internacional de este producto puede contribuir notablemente al transporte pasivo de huevos y larvas. Otros recipientes desechados muy adecuados para el desarrollo de larvas y abundantes en ambientes urbanos son las latas, los envases de bebidas y los frascos de vidrio. Dentro de las viviendas humanas, los floreros y los recipientes con plantas acuáticas son importantes para el desarrollo de las larvas; las canaletas de los tejados y bebederos para animales también son potenciales criaderos (Nelson, 1986).

La cantidad de recipientes naturales utilizados por *Ae. aegypti* para su reproducción es muy baja en relación a los criaderos artificiales. En zonas tropicales con altas precipitaciones, estos mosquitos pueden oviponer en huecos de árboles, axilas de hojas en los jardines cercanos a las viviendas, pozos y otras depresiones del terreno en las cuales puede acumularse agua (Kay et al., 2000). A partir de un estudio realizado en dos ciudades de la provincia de Chaco, Stein et al. (2002) observaron que los neumáticos y las cajas de baterías de automotores representaron, junto con otros recipientes de tamaño pequeño y mediano, los criaderos de mayor importancia en la producción de *Ae. aegypti*. En la ciudad de Córdoba se encontraron larvas y pupas en una gran variedad de recipientes, la mayoría de los cuales habían sido descartados al aire libre, permitiendo la acumulación de agua de lluvia y el consecuente desarrollo de estados inmaduros de culícidos (Almirón y Asís, 2003).

1.3 - Estrategias de control

El control poblacional de mosquitos puede estar dirigido contra los individuos adultos o sus estadios inmaduros acuáticos (larvas y pupas), con un número considerable de métodos disponibles para cada estrategia. La promoción del "descacharrado" peri-doméstico se ha implementado como práctica rutinaria desde hace décadas, normalmente como uno de muchos elementos entre múltiples intervenciones que incluyen rociado de interiores y exteriores con insecticidas, aplicación de telas mosquiteras en aberturas y materiales tratados con insecticidas, como así también el control biológico (Bowman et al., 2016). Si bien no es posible cuantificar con exactitud la contribución específica de estos programas, existen fuertes evidencias a partir de estudios realizados en Cuba (Toledo et al., 2011) que indican que el trabajo comunitario con énfasis en la promoción del manejo ambiental reduce la presencia del vector de manera significativa, incluso más que otras estrategias. Independientemente del método elegido, los abordajes no deberían enfocarse solamente en las personas infectadas, sino también en la abundancia de mosquitos vectores (Ostfeld et al., 2005), buscando una integración entre estos factores. Las herramientas que permiten realizar análisis espacio-temporales a distintas escalas geográficas y la oportunidad de obtener datos únicos a partir de instrumentos de sensado remoto pueden ser de gran valor para la caracterización de condiciones ambientales que favorecen el desarrollo de ciclos biológicos de vectores de enfermedades infecciosas (Reisen, 2010). En otras palabras, estas tecnologías permiten identificar, al menos en parte, las condiciones de hábitat más apropiadas en un área geográfica bajo estudio.

1.4 - Sensado remoto y aplicaciones en ecología de mosquitos vectores

1.4.1 - Generalidades sobre teledetección

La teledetección engloba los procesos que permiten obtener una imagen de la superficie terrestre a partir del registro de datos a gran distancia mediante instrumentos específicos llamados "sensores remotos". Además de los procesos de registro, este conjunto de técnicas también refiere al posterior tratamiento e interpretación de la información colectada. Los sensores remotos generalmente están a bordo de aviones o satélites que orbitan nuestro planeta y permiten captar la emisión energética de objetos presentes en la superficie de la Tierra, ya sea por reflexión de la energía solar o de un haz energético artificial, o por emisión propia en casos particulares (Chuvieco, 2016). La radiación reflejada o emitida desde la superficie terrestre provee información crítica sobre las propiedades del suelo, agua, vegetación, construcciones y cualquier otro tipo de cobertura. El principal objetivo del sensado remoto consiste en entender cómo la energía electromagnética interactúa con la superficie y así extraer información relevante a partir de imágenes procesables. Entonces, gracias a la gran batería de instrumentos disponibles, podemos mencionar tres maneras de obtener información a partir de un objeto: reflexión, emisión y la combinación de ambas (Figura 3). La primera es la más común, ya que utiliza la luz solar natural. El Sol ilumina la Tierra y ella refleja parte de esta energía hacia el espacio dependiendo del tipo y composición de la superficie reflectante. La energía electromagnética reflejada es detectada por el sensor montado en la plataforma satelital, desde la cual se registra y se transmite la señal obtenida hacia una estación terrestre receptora. Por otro lado, las observaciones remotas pueden basarse también en la energía emitida desde la Tierra. Como en este caso no se depende de la radiación solar, este tipo de observación puede llevarse a cabo tanto de día como de noche. Todos los elementos que se encuentran a mayor temperatura que cero absoluto (0 K) emiten energía y mientras más caliente esté un objeto (fuego, lava, agua caliente), mayor será la energía radiante que produce. Finalmente, las observaciones remotas también pueden basarse en sensores activos, llamados así porque poseen su propia fuente de energía. Tienen la capacidad de enviar pulsos hacia objetos "blanco" para caracterizar a continuación la fracción de energía que estos objetos devuelven por reflexión.



Figura 3. Captación de radiación durante el sensado remoto: (i) reflexión, (ii) emisión, (iii) reflexiónemisión. Figura tomada de Chuvieco (2016).

Las fuentes de radiación son muy diversas, por lo tanto las radiaciones electromagnéticas pueden variar desde longitudes de onda muy cortas a muy largas. Por este motivo se tiende a clasificarlas en grupos discretos de longitudes de ondas con frecuencias que se organizan en lo que se llama espectro electromagnético (**Figura 4**). El mismo incluye un rango continuo de longitudes de onda o frecuencias pero pueden identificarse regiones espectrales o bandas con propiedades particulares. Las longitudes de onda más cortas, con la energía radiante más alta, comprenden a los rayos gamma y rayos X, con longitudes que van desde los 10⁻¹⁵ a los 10⁻⁹ m, comúnmente utilizadas en observaciones astronómicas y aplicaciones médicas. Las más largas son aquellas aplicadas en la telecomunicación tradicional (radio y televisión), con un rango de longitudes que va desde los 100 a los 10000 m.

A continuación se mencionan las regiones espectrales comúnmente usadas en la observación mediante sensores remotos:

- Región visible (0.4-0.7 μm). Abarca las longitudes de onda que el ojo humano es capaz de ver, principalmente correspondientes a la luz solar. Esta región del espectro puede ser dividida, a su vez, en los tres colores primarios: azul (0.4-0.5 μm), verde (0.5-0.6 μm) y rojo (0.6-0.7 μm).
- Infrarrojo cercano (0.7-1.2 μm). Esta porción del espectro se encuentra justo a continuación de la capacidad perceptiva del ojo humano y también es conocida como infrarrojo reflejado o fotográfico, ya que parte de esta fracción puede ser detectada con películas especiales. El infrarrojo cercano resulta de especial interés debido a su sensibilidad para determinar el estado de salud de la cobertura vegetal.
- Infrarrojo medio (1.2-8 μm). Esta región se ubica entre el infrarrojo cercano y el infrarrojo térmico, y provee las mejores estimaciones del contenido de humedad de los suelos. Entre los 3 y 8 μm, la señal consiste de una mezcla continua de energía solar reflejada y energía emitida por la superficie terrestre. La franja entre los 3-5 μm es particularmente útil para detectar fuentes de alta temperatura, como volcanes o incendios forestales.
- La región del infrarrojo térmico (8-14 μm) corresponde a la energía emitida desde la superficie de la Tierra, comúnmente usada para mapear temperaturas superficiales. Esta fracción del espectro ha sido ampliamente utilizada para detectar evapotranspiración vegetal, propiedades de masas de hielo y nubes, efectos del calentamiento urbano y discriminación de cuerpos rocosos.
- La región de microondas abarca radiaciones de longitudes de onda mayores a 1 cm de longitud. En esta región del espectro trabajan los sistemas de radar y su mayor ventaja está dada por la bajísima absorción atmosférica, lo que permite observar incluso a través de las nubes. Este tipo de radiación también puede penetrar las frondas vegetales y son muy aprovechables para el análisis de humedad de suelo y de rugosidad de la tierra.



Figura 4. Principales bandas dentro del espectro electromagnético. Figura tomada de Chuvieco (2010).

La energía electromagnética recibida por un sensor a lo largo de diferentes regiones espectrales varía de acuerdo al tipo de cobertura terrestre, en base a sus propiedades biofísicas y bioquímicas. Los procesos naturales, la intervención humana y la interacción entre ambos factores generan cambios constantes en la superficie terrestre (composición, tipo y distribución de coberturas). Una observación sistemática de la superficie de la Tierra resulta crucial a la hora de profundizar el conocimiento sobre sus variaciones en el tiempo, ya que los datos colectados pueden complementar la medición *in situ* de ciertas variables ambientales y permiten ensayar predicciones sobre escenarios futuros.

1.4.2 - Sensado remoto y epidemiología espacial

El abordaje de problemas sanitarios causados por enfermedades infecciosas, que involucra el monitoreo epidemiológico y el desarrollo de tratamientos, vacunas y estrategias de control apropiadas, representa una preocupación creciente en un contexto de rápidos cambios ambientales. Los ecosistemas atraviesan cambios violentos debido a la actividad humana, tales como cambio climático, deforestación, urbanización e intensificación de la agricultura. Todos estos factores, asociados a la intensificación del tráfico globalizado, pueden favorecer la dispersión de enfermedades tradicionales y/o emergentes (Patz et al., 2003; Gottdenker et al., 2014).

Varias disciplinas (ecología, geografía, antropología, informática) se han asociado para producir grandes volúmenes de investigaciones en el campo de la salud humana y animal, unidas a través de varias iniciativas que evolucionaron hacia el concepto de "One Health" (Una Salud; Zinsstag et al., 2011). Si consideramos un sistema hospedador-vector-patógeno en un ambiente físico y humano, los estudios integrales buscan avanzar en el conocimiento de mecanismos y en la identificación de factores de riesgo para la emergencia, dispersión y persistencia de enfermedades. La transmisión de un agente patógeno no depende solamente de factores intrínsecos relacionados a su propia biología, sino también de los de su hospedador, vector y de otros factores extrínsecos relacionados al ambiente (Weaver y Barrett, 2004).

La epidemiología espacial tiene un objetivo más específico: conectar datos de interés epidemiológico con la variabilidad espacio-temporal de factores ambientales y sociales que pueden tener efectos sobre la ocurrencia de brotes de enfermedad. Estos análisis permiten el uso de una gran variedad de datos espaciales, muchas veces disponibles a partir de censos poblacionales, unidades administrativas, servicios meteorológicos, y en general relacionados con indicadores demográficos, económicos y sociales (Kirby et al., 2017). Sin embargo, el conocimiento ambiental relacionado con las enfermedades suele ser muy limitado, sobre todo por la cantidad de información disponible o su accesibilidad. Por este motivo, el sensado remoto se ha convertido en una herramienta crucial de la epidemiología espacial proveyendo, a partir de los datos disponibles en función de limitantes económicos y computacionales, capas de información con cierta adaptación a los requerimientos de cada estudio (nivel de resolución, cobertura espacial, fechas, etc.). Otra gran ventaja del sensado remoto es que no está restringido por fronteras administrativas, lo que facilita la comparación entre áreas alejadas (Tran et al., 2016; Palaniyandi et al., 2017).

La creación de sistemas de observación de la Tierra ha contribuido enormemente al desarrollo de la epidemiología espacial y, más específicamente, a la ecología de la transmisión. Desde la década de 1970, las técnicas de sensado remoto han sido consideradas como "nuevos ojos para los epidemiólogos" (Cline, 1970) y desde entonces se han desarrollado numerosas aplicaciones de estas técnicas, haciendo posible evaluar su potencial (Beck et al, 2000; Thomson y Connor, 2000) y limitaciones (Herbreteau et al, 2007). Desde la década del 2000 el acceso a

datos obtenidos mediante satélites observacionales y meteorológicos ha mejorado notablemente, con una gran cantidad de información de libre disponibilidad, particularmente imágenes Landsat (https://landsat.gsfc.nasa.gov/), MODIS (*Moderate Resolution Imaging Spectroradiometer*, https://modis.gsfc.nasa.gov/) y Sentinel (www.sentinel-hub.com). La descarga de productos se ha facilitado mediante portales en Internet o gracias a la posibilidad de acceso desde paquetes de *software* de estadística o sistemas de información geográfica. De esta manera se han eliminado importantes obstáculos, resultando en la aplicación cada vez más frecuente de enfoques espaciales en la resolución de problemas relacionados con el control de enfermedades infecciosas (Roiz et al., 2011; Tran et al., 2016; He et al., 2019).

Con el objetivo de entender la ecología de estas enfermedades, la investigación suele enfocarse en las condiciones ambientales que favorecen la presencia de organismos vectores u hospedadores y reservorios del agente infeccioso. Estas condiciones son las que determinan qué áreas tienen mayor probabilidad de contener la/s especie/s animal/es que participan del ciclo de transmisión de patógenos. En otras palabras, el rango de distribución geográfica de vectores, hospedadores y reservorios está fuertemente conectado con condiciones bioclimáticas determinantes. Además, las especies no están presentes en todos los tipos de hábitat encontrados dentro de su rango geográfico de distribución, sino que suelen exhibir preferencias que determinan su distribución diferencial incluso a nivel local. El estudio de hábitats de especies requiere de una escala relativamente fina, ya que estos ambientes pueden ser de dimensiones reducidas y variar de manera considerable en forma y estructura. A estos fines, el sensado remoto puede valerse de imágenes satelitales de resolución media (Landsat, Spot) o de muy alta resolución (Quickbird, IKONOS, Worldview, GeoEye).

En términos generales, podemos considerar las siguientes etapas en el proceso de análisis de hábitat (Tran et al., 2016):

- Uso de técnicas de sensado remoto para mapeo de coberturas terrestres. La elección de imágenes y de clases de cobertura de tierra debe corresponderse con los requerimientos del estudio, es decir, con las cantidades y tipos de clases que interesa diferenciar. Esa elección estará restringida por las cualidades intrínsecas de la imagen, sobre todo en relación a sus resoluciones espectral (número y ancho de bandas espectrales registradas) y espacial (área terrestre representada por cada pixel de imagen).

- Cálculo de indicadores o métricas de paisaje, para definir las características del ambiente en los sitios de captura u observación del organismo bajo estudio, por ejemplo, una especie de mosquito vector. Un indicador muy simple es el que representa la distancia entre una observación/captura y una clase de cobertura específica. Si bien son de fácil cálculo, estas variables resultan muy útiles a la hora de interpretar, por ejemplo, la proximidad de una especia a los humanos o a cualquier cobertura terrestre como cuerpos de agua o vegetación. Este análisis también da una idea más clara sobre ciertas características ambientales (como porcentajes de coberturas de cada clase) dentro de un radio determinado (zonas *buffer* alrededor de puntos de captura/observación).
- Análisis estadístico tendiente a buscar un modelo o ecuación que permita predecir la distribución y/o abundancia de una especie en relación a un *set* de variables ambientales, o bien evaluar el efecto relativo de estas variables sobre esta distribución y/o abundancia.

Por lo anteriormente expuesto, el sensado remoto ofrece herramientas potentes a la hora de identificar características del paisaje y factores climáticos asociados al riesgo de circulación de enfermedades de transmisión vectorial (Hayes et al., 1985; Dale et al., 1998; Stefani et al., 2013). Por ejemplo, estas metodologías en articulación con los SIG (sistemas de información geográfica), se han vuelto esenciales para el mapeo de la distribución espacial de enfermedades como malaria, encefalitis Japonesa, filariasis, dengue, chikungunya y leishmaniasis visceral, haciendo hincapié en los aspectos ecológicos de la transmisión epidémica (Barnes et al., 1979; Beck et al., 1997; Bavia et al., 2005; Zou et al., 2006; Tran et al., 2008; Palaniyandi et al., 2014). La información sobre abundancia y distribución de vectores, como así también aquella sobre la actividad de los agentes patógenos transmitidos por ellos, puede ser integrada a datos de mapas temáticos de variables geo-climáticas o de uso o cobertura de suelo, con el fin de identificar zonas problemáticas o bajo mayor riesgo de epidemia (Palaniyandi et al., 2017).

Desde 1998, el Instituto de Altos Estudios Espaciales "Mario Gulich" (CONAE) ha realizado diferentes aportes con investigaciones dentro del campo de la epidemiología panorámica, fundamentalmente enfocadas en estudios y desarrollos sobre Chagas (Porcasi et al., 2018; Weinberg et al., 2018), dengue (Rotela et al., 2017), leishmaniasis (Gómez-Bravo et al., 2017), fiebre hemorrágica Argentina y hantavirus (Andreo et al., 2014; Andreo et al., 2019), entre muchos otros trabajos. Estas tareas de investigación se realizan en vinculación sostenida con

organismos ministeriales y fundaciones en demanda de productos y conocimiento para optimizar gestiones en Salud Pública (http://ig.conae.unc.edu.ar/lineas-de-investigacion/). En relación a mosquitos vectores de enfermedades infecciosas, Rotela et al. (2016) realizaron un análisis espacial de la distribución y abundancia de sitios criaderos de Ae. aegypti en la ciudad de Tartagal (Salta) identificando, mediante análisis de agrupamientos espaciales, diferentes hotspots o zonas de mayor densidad en áreas periféricas de la ciudad. Además realizaron un modelo predictivo basado en un enfoque de máxima entropía, con el fin de explicar la abundancia y distribución de Ae. aegypti mediante variables de cobertura terrestre. En otro estudio similar, Espinosa et al. (2016 b) realizaron un análisis de dinámica espacio-temporal de sitios criaderos de Ae. aegypti en la ciudad de Clorinda, Formosa, también utilizando la aproximación de máxima entropía. Al tratarse de una ciudad pequeña analizada a una escala temporal reducida, se esperaba un comportamiento uniforme en la variación de la abundancia de criaderos entre años diferentes. Sin embargo, los autores observaron que este comportamiento mostraba claras diferencias entre 5 áreas definidas de acuerdo a los patrones de infestación observados. En base a estos resultados, plantearon la necesidad de realizar nuevos estudios que permitan identificar posibles variables reguladoras de la dinámica espacio-temporal.

En términos generales, el uso de herramientas de sensado remoto y SIG ha demostrado un gran potencial durante las últimas décadas. Los diferentes enfoques aplicados a los análisis realizados hasta el momento en diferentes partes del mundo nos acercan cada vez más a la posibilidad de generar productos de mayor poder predictivo, aplicables por los organismos efectores de Salud Pública.

1.5 – Construcción de la hipótesis y objetivos

Desde el año 2008, la Fundación Mundo Sano (www.mundosano.org) implementó un programa de monitoreo extensivo de criaderos domésticos de *Ae. aegypti* en Clorinda, localidad situada en la frontera entre la provincia de Formosa y Paraguay. En un estudio previo mencionado anteriormente, se analizaron los cambios en abundancia de sitios infestados con estadios inmaduros de *Ae. aegypti* y se reconocieron patrones temporales diferentes dentro de cada zona (Espinosa et al., 2016 b). El hallazgo de esta variabilidad espacial cobra particular importancia, ya que se trata de una ciudad pequeña que además no estuvo sometida a presiones selectivas fuertes como actividades de control durante los años previos al estudio.

El presente estudio se basa en la hipótesis de que diferentes variables estimadas a partir de información provista por sensores remotos pueden utilizarse para agrupar un número de unidades espaciales (barrios) en un conjunto más reducido de categorías o *clusters* ambientales y que esta estructuración permite el modelado de la evolución temporal de la abundancia de criaderos positivos de *Ae. aegypti* en Clorinda. De esta manera, pueden definirse zonas de la ciudad, ya no de manera arbitraria sino en base a sus características ambientales, y luego evaluar su potencial explicativo a la hora de modelar las variaciones en abundancia de criaderos de *Ae. aegypti*.

Nuestro principal objetivo fue examinar en mayor profundidad la evolución temporal del número de casas infestadas con estadios inmaduros de *Ae. aegypti* en cada barrio en particular, para finalmente explorar si las características de la dinámica diferencial pueden atribuirse a áreas que comparten condiciones ambientales similares. Con este fin, se utilizó información provista por variables estimadas a partir de sensores remotos para agrupar los barrios muestreados en un conjunto más reducido de categorías o *clusters* ambientales y evaluar su utilidad para el modelado de la evolución temporal de esta población local de *Ae. aegypti*.

2 - MATERIALES Y MÉTODOS

2.1 - Área de estudio

Clorinda (25°17'S, 57°43'O) se ubica en el límite occidental de la provincia de Formosa, a 4 km de la frontera con Paraguay (**Figura 4**). La ciudad está asentada sobre la costa derecha del río Pilcomayo, 10 km antes de su confluencia con el río Paraguay, en una región caracterizada por un clima subtropical sin estación seca. La temperatura mediana anual es de alrededor 23°C y la precipitación media anual de 1300 mm, principalmente aportada durante la temporada lluviosa entre Octubre y Mayo (Gürtler et al., 2009). De acuerdo al último censo en Argentina llevado adelante durante 2010, la ciudad cuenta con 16600 hogares y una población de aproximadamente 23300 habitantes.



Figura 4. Ubicación geográfica de Clorinda (punto amarillo), en la provincia de Formosa (área naranja).

2.2 - Descripción de los datos

La Fundación Mundo Sano diseñó un monitoreo extensivo de mosquitos Ae. aegypti en viviendas de prácticamente toda la ciudad de Clorinda. Se registraron, mediante visitas y observaciones directas en cada casa, la presencia y la abundancia de recipientes criaderos infestados con estados inmaduros (larvas o pupas) de Ae. aegypti. La Figura 5 muestra la distribución de los barrios censados en las encuestas. De un total de 12 muestreos realizados en recorridos por la ciudad (M1-M12), se seleccionaron 7 encuestas entomológicas realizadas entre Octubre de 2011 y Noviembre de 2013, que representan un total de 11.022 visitas a domicilios particulares. Esta selección se basó en que las fechas de las observaciones correspondientes a estos muestreos permitieron agruparlos aproximadamente por estaciones (primavera, verano, otoño) correspondientes a dos periodos consecutivos comparables en el modelado posterior (2012 y 2013). Si bien Clorinda cuenta con un clima sub-tropical sin la estacionalidad marcada de los climas templados, se reconocen diferencias inter-estacionales en relación al régimen de precipitaciones y temperaturas ambientales registradas. Las fechas de visita a cada barrio y el agrupamiento apareado de los muestreos por estación y por año se muestran en la tabla del Anexo I; dicho agrupamiento se esquematiza además en la Figura 6A y su distribución geográfica parcial se muestra en la Figura 6B. En cada oportunidad, aproximadamente el 20% de los hogares de la ciudad fue seleccionado de manera arbitraria para su monitoreo, previo consentimiento de sus habitantes. Las larvas y pupas colectadas durante las tareas fueron transportadas al laboratorio de entomología de la Fundación Mundo Sano para su identificación taxonómica usando una clave morfológica específica (Rossi y Almirón, 2004). Las casas en las que se encontró al menos un contenedor infestado fueron catalogadas como sitios "positivos".

Se utilizó el SIG de acceso libre y código abierto "Quantum GIS (QGIS), versión 2.10.1 PISA (www.qgis.org) para construir las capas vectoriales (archivos *shape*) a fin de representar los puntos de muestreo y colectar los datos espacializados de las distintas fechas de muestreo. Los sitios en los que se encontraron estadios inmaduros de *Ae. aegypti* fueron registrados como positivos, mientras que aquellos sin infestar se consideraron negativos. Mediante el uso de la herramienta "mapa de calor" de QGIS se estimó la densidad de casas positivas por año y, con el fin de evaluar visualmente la variabilidad espacial de densidades, se calculó la diferencia entre capas de diferentes años y se presentó en nuevas imágenes *raster* por temporada. Este producto se obtuvo realizando una resta simple entre cada par de *rasters* interanuales, mostrados por estación, en la herramienta "calculadora raster" de QGIS. Finalmente, si bien se registraron conteos casa por casa, para los objetivos del presente estudio se calculó el número de viviendas positivas por barrio para cada año y este valor fue utilizado en el modelado espacio-temporal. Se tomó el barrio como unidad de observación, ya que representa la principal unidad administrativa sobre la cual se diseñan e implementan acciones de Salud Pública.



Figura 5. Distribución de los barrios de Clorinda censados durante las encuestas entomológicas incluidas en el presente estudio. Cada número (1-32) corresponde a una designación arbitraria para cada barrio.

Α



В



Figura 6. A) Representación gráfica del diseño de muestreos (M1-M9) realizados en la ciudad de Clorinda. B) Imagen parcial de la ciudad de Clorinda, con algunos de los puntos de muestreo (viviendas) distribuídos por barrios.

2.3 - Clasificación no-supervisada de imagen SPOT

Se utilizó una imagen obtenida mediante el satélite SPOT 5, del día 9 de Abril de 2013, con el fin de generar productos clasificados de cobertura terrestre y subsecuentes variables macro-ambientales. Este sistema provee imágenes con resolución multiespectral de 10 m en su producto SPOT 5 – J, con 4 bandas espectrales: verde (0.50-0.59 µm), rojo (0.61-0.68 µm), infrarrojo cercano (0.78-0.89 µm) e infrarrojo medio (1.58-1.75 µm). La imagen utilizada fue provista por la Comisión Nacional de Actividades Espaciales (CONAE; https://catalogos.conae.gov.ar/SPOT/). La clasificación del área bajo estudio se realizó por la metodología no-supervisada "k-means" implementada en el programa ENVI (versión 4.8; Exelis Visual Information Solutions). Este método nos permite partir un conjunto de n observaciones en k agrupamientos, cada uno de los cuales está representado por el promedio de los puntos que lo componen (centroide). La cantidad de grupos k debe fijarse como parámetro a priori y luego el método comienza con k centroides de ubicación aleatoria y asigna cada observación al centroide más cercano. Una vez terminadas estas asignaciones, los centroides se desplazan a la ubicación promedio de todos los datos asociados a él, y se vuelven a asignar los puntos de acuerdo a estas nuevas posiciones. De esta forma se agrupan las observaciones de manera tal que todas las que compartan el mismo grupo sean lo más semejantes entre sí y las pertenecientes a grupos distintos sean lo más distintas entre sí. Estas semejanzas y desemejanzas se miden en términos de distancias euclídeas en el plano. El algoritmo es muy sensible a la selección aleatoria inicial de centroides, pero este efecto puede reducirse realizando varias repeticiones o iteraciones. Una medida para indicar cuán bien los centroides representan a los miembros de su grupo es la suma de los errores al cuadrado, es decir, la sumatoria de las distancias al cuadrado de cada observación al centroide de su grupo. En cada iteración, el algoritmo intenta reducir el valor de esta sumatoria (MacQueen, 1967). Este es un método efectivo para agrupar datos contenidos en imágenes multiespectrales y extraer información sobre tipos de cobertura de suelo. Se aplica para agrupar píxeles con características espectrales similares, es decir que estas operaciones numéricas se llevan a cabo para buscar agrupamientos naturales de las propiedades o atributos del conjunto de píxeles de la imagen. Una vez que los datos han sido clasificados, se realiza un etiquetado a posteriori para asignar o transformar estos grupos espectrales naturales en clases temáticas de interés (vegetación, agua, construcciones urbanas, etc.). Como se mencionó más arriba, los algoritmos de agrupamiento utilizan cálculos iterativos para encontrar un set óptimo de bordes entre los tipos de clases obtenidas; en este caso se trabajó con 10 iteraciones del proceso de clasificación. Se obtuvieron 7 clases de coberturas terrestres: suelo desnudo, agua superficial, humedales, baja/arbustos. vegetación vegetación alta/árboles. urbanización V pastizales/cultivos. Se definieron 40 puntos de control ("verdad de campo") verificados en Google Earth para cada cobertura ambiental con el fin de corroborar la eficacia de la clasificación. Para realizar dicha corroboración se utilizó una matriz de confusión realizada con la herramienta "confusion matrix" del programa ENVI, en la cual los resultados de la clasificación fueron comparados con la información de "verdad de campo" para así identificar la naturaleza y cantidad de los errores de asignación a clases. En una matriz de este tipo, las columnas representan clases verdaderas (tomadas a partir de la "verdad de campo") y las filas corresponden a las predicciones del clasificador. Se trata de una matriz cuadrada con todas las clasificaciones correctas en la diagonal. A partir de esta matriz, se extrajo la precisión general de clasificación (overall accuracy) y el coeficiente Kappa. El primer indicador se calcula sumando el número de valores correctamente clasificados y dividiendo esta suma por el número total de valores. El coeficiente kappa (Cohen, 1960) mide el acuerdo entre la clasificación y los datos de "verdad de campo". Un valor de kappa igual a 1 representa una concordancia perfecta, mientras que un valor igual a 0 significa que no hay tal concordancia. Este coeficiente se computa de la siguiente manera:

$$\kappa = \frac{N\sum_{i=1}^{n} m_{i,i} - \sum_{i=1}^{n} (G_i C_i)}{N^2 - \sum_{i=1}^{n} (G_i C_i)}$$

Donde...

i es el número de clases

N es el número total de valores clasificados comparados con los de "verdad de campo"

m_{i,i} es el número de valores que pertenecen a la clase "verdadera" i que ha sido clasificada comoi (los valores que se encuentran en la diagonal de la matriz de confusión)

 C_i es el número total de valores predichos que pertenecen a la clase i

 G_i es el número total de valores de "verdad de campo" que pertenecen a la clase i

2.4 - Generación de capas de variables macro-ambientales

A continuación se generaron dos tipos de variables macro-ambientales para cada clase, expresadas como mapas de distancia a cada clase (buffer) y como porcentajes de cada clase por área de estudio (tipo de suelo, vegetación, agua, etc.). Por un lado, la distancia de cada pixel hasta cada clase de interés fue calculada utilizando la herramienta "buffer zone image" en ENVI, con la que se producen imágenes en las cuales cada pixel tiene un valor definido como la distancia euclidiana desde dicho pixel hasta el más próximo perteneciente a la clase seleccionada. Una vez obtenidos los mapas, se utilizó un archivo shape con polígonos representativos de los barrios de Clorinda incluidos en nuestro estudio. Esta capa vectorial fue diseñada manualmente en QGIS a partir de un mapa catastral compartido por la Fundación Mundo Sano. Una vez trazados los polígonos correspondientes a todos los barrios analizados, se unificaron en una capa vectorial simple mediante la herramienta "multiparts to singleparts" para poder hacer extracciones de valores que se agrupen en la tabla de atributos del objeto creado, de manera tal que cada columna corresponda a un barrio distinto. Esa capa de polígonos fue superpuesta a cada mapa de distancias buffer y se extrajo, con la herramienta "zonal statistics" del mismo programa, la media de la distancia a la clase de interés representada en el mapa. Por otro lado, para la extracción de porcentajes de clases de cobertura de suelo, se utilizó la herramienta "Landscape Vector Overlay" del plugin LecoS-Landscape Ecology, implementado en QGIS (Jung, 2013). Este recurso permite utilizar una capa vectorial para extraer métricas precisas a partir de un producto clasificado, en formato raster. En este caso, se utilizó como capa vectorial el mismo archivo shape de polígonos mencionado anteriormente y se extrajo la proporción de cada clase para cada barrio. Además, como medida de la heterogeneidad del paisaje se calculó el índice de diversidad de Shannon, una estimación cuantitativa del número de categorías diferentes. Este índice es uno de los más usados en la historia de la Ecología y se basa en que la diversidad en un sistema natural puede medirse de manera similar a la información contenida en un código o mensaje. Asume que los individuos son muestreados al azar a partir de una comunidad de tamaño infinito y que todas las especies se encuentran representadas en la muestra. Se calcula a partir de la siguiente ecuación:

$$H' = -\sum_{i=1}^{S} p_i \ln p_i$$

Donde:

S es el número de especies.

pi es la proporción de individuos (píxeles) encontrados de la i-ésima especie (clase).

De esta forma, el índice contempla la cantidad de especies presentes en el área de estudio (riqueza de especies) y la cantidad relativa de individuos de cada una de esas especies (abundancia). Suele tomar valores entre 1.5 y 3.5 cuando se obtiene a partir de datos empíricos y muy rara vez toma valores por encima de 4, lo cual suele ocurrir cuando se trabaja con un número excesivamente alto de especies (10^5 o más para H'>5). Su interpretación puede ser dificultosa ya que el rango de valores posibles es muy estrecho; sin embargo, las mediciones en ambientes heterogéneos (como un *raster* de cobertura urbana o los bosques tropicales) suelen aportar a la diferenciación de áreas de interés. El índice de Shannon fue abordado con mayor profundidad y perspectiva ecológica en Magurran (2013).

A continuación se muestra la lista de variables calculadas:

- Distancia a cuerpos de agua superficial.
- Distancia a humedales.
- Distancia a cultivos.
- Distancia a suelo desnudo.
- Distancia a construcción urbana.
- Distancia a vegetación alta.
- Distancia a vegetación baja.
- Porcentaje de agua superficial.
- Porcentaje de humedales.
- Porcentaje de cultivos.
- Porcentaje de suelo desnudo.
- Porcentaje de construcción urbana.
- Porcentaje de vegetación alta.
- Porcentaje de vegetación baja.
- índice de diversidad Shannon.

Este conjunto de variables fue sometido a un análisis de correlación lineal utilizando el entorno de herramientas estadísticas R (<u>www.r-project.org</u>), implementado en RStudio y sólo se retuvieron para posteriores análisis aquellas variables con índice de correlación menor a 0.7 (<u>Anexo VII</u>-2). En el caso de variables ambientales como capas de cobertura de suelo, los efectos de algunas de ellas en la explicación del comportamiento de una variable respuesta pueden estar

contenidos en otras. Incluir estas variables redundantes puede generar modelos con indicadores de ajuste irrealmente buenos, como en muchos trabajos en los que se modela presencia y abundancia de especies utilizando grandes *sets* de variables sin depurar (Zuur et al., 2010). Un punto de corte en el valor absoluto 0.7 en índice de Pearson es una forma muy simple y frecuentemente sugerida en la literatura durante la exploración de colinearidad (Dormann et al., 2011). Este procedimiento se realizó agregando un argumento "cutoff = 0.7" a la función "findcorrelation" de R (librería "Caret"), aplicada sobre la tabla de valores estandarizados de variables ambientales.

2.5 - Agrupamiento ambiental de barrios

Todos los análisis estadísticos se llevaron a cabo en el entorno R. Debido a que las variables a analizar toman valores a escalas muy diferentes (<u>Anexo III</u>), como paso preliminar las mismas fueron estandarizadas tal como sugieren Zuur et al. (2007). Para esto, los valores fueron escalados y centrados mediante la función "*scale*", basada en:

$$z_i = \frac{x_i - \bar{x}}{s}$$

Donde x_i es el valor de la i-ésima observación, \overline{X} es la media de la muestra y S su desviación estándar.

Luego de este paso de estandarización (<u>Anexo VII</u>-1), la magnitud de las fluctuaciones de las variables analizadas se vuelve similar y desaparecen las grandes diferencias en valores absolutos entre series.

A continuación, con el fin de atender el principal objetivo de este estudio, se buscó resumir toda la información ambiental colectada por sensado remoto en grupos de barrios con características similares, o *clusters* ambientales. A tal fin, la matriz de valores de variables ambientales asignados a los diferentes barrios fue sometida al método de agrupamiento *k-means* (Hartigan y Wong, 1979). De manera sintética, este método se utiliza para agrupar un conjunto de *n* observaciones en *k* grupos de forma tal que cada observación pertenece al grupo cuyo valor medio es más cercano. Los pasos del algoritmo son los que se describieron anteriormente para la implementación de este método en la clasificación de la imagen satelital. La suma de cuadrados intragrupos disminuye a medida que aumenta *k* y esto puede observarse mediante un "gráfico de sedimentación" (suma de cuadrados en función de número de grupos/k). En el eje y de este gráfico se representa la suma de las distancias euclídeas al cuadrado entre cada observación (x_i) y el centroide (μ) de su *cluster* (C_k). Esto equivale a la suma de cuadrados internos del *cluster*:

$$W(C_k) = \sum_{x_i,\in C_k} (x_i - \mu_k)^2$$

Una forma simple de estimar el número k óptimo de *clusters* cuando no tenemos otra información adicional, consiste en aplicar el algoritmo de *k-means* para un rango de valores de k e identificar aquel valor a partir del cual la reducción en la suma total de la varianza intra-*cluster* deja de ser sustancial. Esta estrategia se la conoce como el "método del codo", ya que se basa en la identificación del punto a partir del cual la curva de sedimentación forma un "codo", es decir, donde la reducción en la suma total de cuadrados internos parece estabilizarse. Este gráfico se realizó utilizando la librería "*vegan*" y expresando las sumas de cuadrados intra-grupos en función del número de *clusters* ambientales (k). Una vez definido el número de grupos a discriminar, la agrupación del conjunto de datos puede graficarse como una partición del espacio de datos en celdas de Voronoi. Es decir, se subdivide el conjunto de datos en subconjuntos correspondientes a los k *clusters* definidos, de manera tal que cada punto en el plano euclidiano pertenece al subconjunto que lo agrupa de manera más próxima al centroide de la celda k_a . Este proceso se realizó utilizando el paquete "*cluster*" (Maechler et al., 2017) para graficar el arreglo de barrios obtenido a partir de este análisis (Anexo VII-3).

2.6 - Modelo lineal generalizado

Con el fin de evaluar la contribución de la interacción entre los factores "*cluster* ambiental" y "periodo de muestreo" (variable explicativa) a las variaciones en abundancia de sitios criaderos de *Ae. aegypti* (variable respuesta), se ajustó un modelo linear generalizado (GLM, <u>Anexo VII</u>-4). En términos generales, un GLM tiene tres componentes básicos:

Componente aleatoria: se refiere a la variable aleatoria respuesta Y, con observaciones independientes $(y_1, ..., y_n)$, y su distribución de probabilidad. En este caso cada observación es un recuento (número de casas positivas), por lo que podemos asignar a Y una distribución de Poisson o una binomial negativa.

Componente sistemática: se refiere a las variables explicativas (independientes o predictoras) usadas en la función predictora lineal. La combinación lineal de variables explicativas

 $\alpha + \beta_1 x_1 + \cdots + \beta_k x_k$

se conoce como predictor lineal.

Función link: se trata de una función del valor esperado de Y, E(Y), como una combinación lineal de las variables predictoras. El valor esperado de Y se denota como μ =E(Y), entonces la función link especifica una función g(.) que transforma ese valor esperado en un predictor lineal:

 $g(\mu) = \alpha + \beta_1 x_1 + \cdots + \beta_k x_k$

La función link es invertible, por lo tanto el GLM puede pensarse como un modelo lineal para la transformada de la respuesta esperadao como una regresión no-lineal de esa respuesta.

Inicialmente se planteó un GLM asumiendo que el componente aleatorio "número de casas positivas" sigue una distribución de Poisson. En esta distribución unimodal la media y la varianza coinciden, es decir, cuando el número de recuentos es mayor en media, también tiene mayor variabilidad. En el GLM se utiliza frecuentemente el logaritmo de la media para la función link, entonces el modelo log-lineal con una variable explicativa x se puede expresar de esta manera:

Y_i = Número de casas con criaderos infestados por larvas de *Ae. aegypti*

$$\begin{split} Y_i &\sim \text{Poisson}(\pmb{\mu}i) \\ E(Y_i) &= \pmb{\mu}i \\ \log(\pmb{\mu}) &= \pmb{\alpha} + \pmb{\beta}x \\ \text{entonces}... \end{split}$$

 $\mu = \exp \left[\alpha + \beta x\right]$

Los *clusters* diferentes y los dos años de muestreo fueron asignados como factores y usados como variables explicativas del número de casas positivas (variables respuesta). La *deviance* para

un modelo que incluye sólo la constante de regresión α , es decir para un modelo que no depende de ninguna variable, se conoce como *deviance* nula (D_0) , mientras que la *deviance* para el modelo en cuestión, que depende de las variables explicativas, se denomina deviance residual (D1). Si el GLM se ajusta a los datos de manera razonable cuando asumimos distribución Poisson, la deviance residual es muy parecida a los grados de libertad de los residuos, o la varianza es igual a la media. Pero cuando trabajamos con datos reales esta hipótesis no suele cumplirse. Con frecuencia la varianza es mayor que la media (se observa una deviance residual muy alta), es decir, la variación condicional del valor esperado para la variable respuesta excede la variación explicable por una variable con distribución Poisson. Este fenómeno se denomina sobredispersión y, normalmente, esta situación es consecuencia de la heterogeneidad entre las diferentes observaciones y puede interpretarse como una mezcla de distribuciones Poisson. Una forma de detectar este efecto consiste en calcular un parámetro de dispersión, definido como el cociente $D_1/grados$ de libertad. Si este cociente es sustancialmente mayor a 1, estamos ante un problema de sobredispersión. Si bien es difícil fijar un umbral a partir del cual la sobredispersión se vuelve considerable, en estudios de sistemas biológicos bajo condiciones no-controladas suelen fijarse valores entre 1.2 y 1.5. En nuestro caso tomamos como valor límite de tolerancia 1.2.

Como suele ocurrir con los conteos, nuestros datos mostraron sobredispersión al ajustar el GLM asumiendo una distribución del tipo Poisson. Para evitar las distorsiones que pueden emerger a partir de este tipo de situación, a continuación se siguieron las recomendaciones de Fox (2015) y se implementó un GLM asumiendo distribución binomial negativa en el que:

 $E(Y) = \mu$

 $Var(Y) = \mu + \mu^2/k$

El parámetro 1/k es un parámetro de dispersión, de modo que si $1/k \rightarrow 0$ entonces $Var(Y) \rightarrow \mu y$ la distribución binomial negativa converge a una distribución tipo Poisson. Además, para un valor fijo de k la distribución pertenece a la familia exponencial natural, por lo que se puede definir un modelo GLM binomial negativo. En general se utiliza una función link de tipo logaritmo. La regresión binomial negativa se puede utilizar para datos sobredispersos de recuentos, es decir cuando la varianza condicional es mayor que la media condicional. Se puede considerar como una generalización de la regresión de Poisson, ya que tiene su misma estructura de medias y además un parámetro adicional para el modelo de sobredispersión.

El GLM se ajustó utilizando la librería MASS (Venables et al., 1999), tomando los datos de la tabla del <u>Anexo II</u>, y luego se hizo una exploración gráfica de la distribución de los residuos (Zuur et al., 2010). Recordemos que la variable respuesta "abundancia de casas positivas" proviene de muestreos realizados durante encuestas entomológicas de la Fundación Mundo Sano. Estas encuestas se realizaron con esfuerzos heterogéneos; es decir, la cantidad de casas visitadas difiere de barrio a barrio. Por este motivo, los conteos en valores absolutos se corrigieron según el número total de casas muestreadas por barrio incluyendo el argumento "offset" en la función "glm" en el modelado estadístico. De esta manera se evitaron desbalances relacionados a esfuerzos de muestreo diferentes, ya que se trabajó con proporciones de casas "positivas" en relación al total de casas muestreadas en cada barrio.

Finalmente, se realizaron comparaciones múltiples para confirmar la existencia de diferencias significativas entre los efectos de las interacciones (Anexo VII-5). Este tipo de técnicas se engloba bajo la denominación de "contrastes para comparaciones múltiples" y nos permite identificar qué efectos son diferentes y en cuánto oscila el valor de esas diferencias. Se utilizaron los paquetes *multicomp* (Hothorn et al., 2008) y *factorplot* (Armstrong, 2016) para realizar las comparaciones a posteriori de a pares, con un ajuste de Bonferroni incorporado como argumento, y representar gráficamente las magnitudes de efectos de interacción y sus niveles de significancia. Muy simplemente, la corrección de Bonferroni consiste en dividir el nivel de significancia de un efecto (α) entre el número de comparaciones realizadas de a pares.

$\alpha_{ajustado} = \alpha/n$ úmero de grupos

Este método garantizaría que la probabilidad de obtener al menos un falso positivo entre todas las comparaciones sea $\leq \alpha$ y permite contrastar una hipótesis nula general de forma simultánea. Se considera un método excesivamente conservador, sobre todo a medida que aumenta el número de comparaciones. Como alternativa se probó el ajuste Holm-Bonferroni, con el cual el valor de α se corrige secuencialmente haciéndolo menos conservador que el Bonferroni tradicional. Si bien estos métodos son usados de manera muy extensiva, en Cabin y Mitchell (2000) y en Nakagawa (2004) se hace un análisis crítico donde se discuten posibles implicancias ligadas a su bajo poder estadístico.

3 – RESULTADOS

Mediante una exploración preliminar de los datos se evaluó la variación en las densidades de casas positivas en los barrios muestreados durante 6 temporadas de 2 años consecutivos (2012 y 2013), observándose importantes cambios temporales en los patrones de abundancia entre barrios distintos y entre periodos diferentes para los mismos barrios (**Figura 7**). Los "mapas de calor" de la **Figura 8** representan superficies interpoladas que muestran los incrementos positivos o negativos de densidad de casas positivas en Clorinda entre los dos años separados también por temporada. Puede observarse un aumento generalizado entre una primavera en la siguiente, e incrementos más agudos y localizados entre los veranos y los otoños.

La clasificación no supervisada de la escena de interés permitió discriminar 7 capas de coberturas de suelo diferentes: suelo desnudo, cuerpos de agua superficial, humedales, vegetación alta, vegetación baja, áreas urbanizadas, pastizales/cultivos. La matriz de confusión con puntos "controles" o de validación arrojó una precisión general de clasificación del 77.6% y un coeficiente *kappa* de 0.74. Algunas clases (agua superficial, pastizales/cultivos, humedales, vegetación alta y baja) alcanzaron valores por encima del 80% de precisión, mientras que otras (urbanización y suelo desnudo) presentaron precisiones del 50% y 76%, respectivamente.

En la **Figura 9** se muestra, a modo de ejemplo, un fragmento del mapa de clasificación de coberturas terrestres obtenido. Del *set* completo de variables (<u>ANEXO III</u>), las siguientes 8 fueron elegidas para agrupar mediante *k-means* los barrios muestreados, por mostrar un valor absoluto de coeficiente de correlación de Pearson r<0.7:

- distancia a cuerpos de agua.
- distancia a humedales
- distancia a cultivos.
- distancia a vegetación alta
- distancia a vegetación baja
- porcentaje de cuerpos de agua
- porcentaje de cultivos.
- porcentaje de vegetación baja.

Las matrices de correlación para el set completo de variables ambientales y para el subconjunto de variables seleccionadas para los análisis siguientes (r<0.7) se muestran en los Anexos IV y V, respectivamente.



Figura 7. Distribución temporal de casas positivas por barrio, 2012-2013. Los muestreos entomológicos fueron realizados en primavera, verano y otoño de 2012 y 2013.


Figura 8. Variación relativa inter-anual de las densidades estimadas de casas positivas en Clorinda. Azul: incrementos negativos. Rojo: incrementos positivos. A) primavera; B) Verano; C) otoño.



Figura 9. Fragmento del mapa de clasificación obtenido por la metodología no-supervisada *k-means* para la ciudad de Clorinda (inferior), en relación a una captura de Google Earth de la misma zona (superior).

La Figura 10 muestra un ejemplo de mapa *buffer* que, en este caso, representa la distancia de cada píxel a la clase "cuerpo de agua superficial". Como puede observarse, las menores distancias se muestran hacia el color azul, mientras que las mayores hacia el color blanco. De la misma manera se obtuvieron los otros mapas *buffer* de distancia a las otras clases de interés.

Luego de la observación del gráfico de sedimentación (**Figura 11**) se determinó que el valor *k* que incorpora una alta proporción de la variabilidad entre grupos fuera ajustado a 4 para la subsecuente asignación de barrios a clases ambientales. Por encima de ese valor de *k*, no se observaron caídas importantes en las sumas de cuadrados intragrupos a medida que se agregaban números de *clusters* al agrupamiento. Repitiendo el procedimiento aumentando el número de *clusters* se observó que valores mayores a 4 llevaban a arreglos menos parsimoniosos con superposición de *clusters* en el espacio. Es decir, la superposición espacial de las celdas de Voronoi dio cuenta de que un número mayor de grupos no discriminaba correctamente los subconjuntos de barrios entre sí.

A continuación, la matriz de valores de las 8 variables ambientales por barrio fue clasificada mediante *k-means*, obteniéndose el agrupamiento de los 32 barrios muestreados en 4 *clusters* ambientales diferenciables (**Figuras 12 y 13A**). De acuerdo a información brindada por personal de la Fundación Mundo Sano, el *cluster k-*I estuvo representado por áreas residenciales y una zona comercial densamente construida, principalmente ocupada por galpones y talleres de reparación de vehículos, como así también por el cementerio de la ciudad (en el barrio 11). El *cluster k-*II incluye barrios con actividad comercial intensa, viviendas precarias y zonas inundables cerca de ríos, mientras que el *cluster k-*III está mayormente compuesto por barrios de planes sociales densamente poblados (barrios 6, 7, 24 y 31), como así también un asentamiento del pueblo originario Toba (barrio 26). Finalmente, el *cluster k-*IV alberga zonas de cultivos y barrios periféricos que deben ser evacuados con frecuencia debido a la proximidad con el río Pilcomayo.



Figura 10. Fragmento de mapa *buffer* de distancia a la clase "cuerpo de agua superficial" (inferior) en relación al mismo fragmento clasificado por *k-means* (medio) y una captura de Google Earth de la misma zona (superior).



Figura 11. Gráfico de sedimentación en el que se muestra la variación de la suma de cuadrados intragrupos a medida que aumenta el número de grupos (*clusters*, k) a los que se asigna el conjunto de barrios estudiados. El círculo rojo y la línea punteada marcan el límite a partir del cual no se observa una caída significativa en la suma de cuadrados (k=4).



Figura 12. Agrupamiento por *k*-means de los 32 barrios analizados en Clorinda. El agrupamiento se realizó incluyendo el set de variables ambientales seleccionadas y el valor de *k* fue fijado en 4 en base a la observación del gráfico de sedimentación.



Figura 13. A) Agrupamiento por k-means de los 32 barrios analizados en Clorinda. El agrupamiento se realizó incluyendo el set de variables ambientales seleccionadas con k = 4. B) Variación en el número de casas positivas entre 2012 y 2013 por *cluster* ambiental.

k4 k1

clusters ambientales

k2

k3

k3

k2

En la **Tabla 2** se muestra la salida (*summary*) del GLM con familia de distribución Poisson, respetando el formato crudo arrojado en la consola de R Studio. El cociente entre la *deviance* residual y sus grados de libertad (parámetro de dispersión) da 8.92, por lo que estamos ante una clara situación de sobredispersión, de acuerdo a los criterios ya mencionados en Materiales y Métodos.

Estimate	Std. Error	z value	Pr(> z)	
-2.10171	0.07881	-26.668	<2 e-16	***
1.47717	0.08682	17.015	< 2e-16	***
0.20785	0.09576	2.171	0.030	*
0.48673	0.12438	3.913	9.11e-05	*
-0.14204	0.12812	-1.109	0.268	
-1.49441	0.11386	-13.124	< 2e-16	.***
-1.90794	0.16984	-11.234	< 2e-16	***
-1.28341	0.15554	-8.251	< 2e-16	***
	Estimate -2.10171 1.47717 0.20785 0.48673 -0.14204 -1.49441 -1.90794 -1.28341	EstimateStd. Error-2.101710.078811.477170.086820.207850.095760.486730.12438-0.142040.12812-1.494410.11386-1.907940.16984-1.283410.15554	EstimateStd. Errorz value-2.101710.07881-26.6681.477170.0868217.0150.207850.095762.1710.486730.124383.913-0.142040.12812-1.109-1.494410.11386-13.124-1.907940.16984-11.234-1.283410.15554-8.251	EstimateStd. Errorz value $Pr(> z)$ -2.101710.07881-26.668 $<2 e-16$ 1.477170.0868217.015 $<2e-16$ 0.207850.095762.1710.0300.486730.124383.9139.11e-05-0.142040.12812-1.1090.268-1.494410.11386-13.124 $<2e-16$ -1.907940.16984-11.234 $<2e-16$ -1.283410.15554-8.251 $<2e-16$

Tabla 2. Salida cruda (*summary*) del modelo lineal generalizado corrido en R, utilizando familia de distribución binomial negativa.

Signif. codes: 0 **** 0.001 *** 0.01 ** 0.05 ·? 0.1 ** 1 Null deviance: 2114.2 on 161 degrees of freedom Residual deviance: 1358.6 on 154 degrees of freedom

A continuación se ajustó el GLM asumiendo una distribución binomial negativa. La salida de este GLM se muestra en la **Tabla 3**. La *deviance* del modelo propuesto (*Residual deviance* = 182.33) es menor que la del modelo nulo (*Null deviance* = 307.47), lo que indica un mejor ajuste del primero en relación al segundo.

	Estimate	Std Error	z value	$Pr(\geq z)$	
(Intercept)	-2.4902	0.2187	-11.385	-2e-16	***
año 2013:k1	2.2541	0.2856	7.893	2.95e-15	***
año 2012: <i>k</i> 2	0.5045	0.2605	1.936	0.0528	
año 2012: <i>k</i> 3	0.7216	0.3355	0.151	0.0315	*
año 2012: <i>k</i> 4	0.2333	0.3045	0.766	0.4436	
año 2013: <i>k</i> 2	-2.0529	0.3470	-5.916	3.30e-09	***
Año 2013: <i>k</i> 3	-2.5642	0.4616	-5.555	2.77e-08	***
Año 2013: <i>k</i> 4	-2.0836	0.4080	-5.107	3.28e-07	

Tabla 3. Salida cruda (*summary*) del modelo lineal generalizado corrido en R, utilizando familia de distribución binomial negativa.

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Null deviance: 307.47 on 161 degrees of freedom Residual deviance: 182.33 on 154 degrees of freedom

En este caso, el factor de dispersión no superó a 1.2 (182.33/154 = 1.18), por lo que el nivel de sobredispersión es marcadamente menor al del modelo que asumía distribución Poisson.

Así como en los modelos de regresión lineal el valor de R^2 nos permite evaluar la proporción de la variación explicada por el modelo, en el caso de los GLMs el equivalente indicador del nivel de ajuste global es el D^2 , que refleja la cantidad de *deviance* explicada. A diferencia del R^2 , el valor de D^2 no viene incluido en el *summary* del GLM en R. En cambio, se obtiene realizando la siguiente operación: (*Null deviance* - *Residual deviance*)/*Null deviance*, donde "*null deviance*" es la *deviance* del modelo sólo con el intercepto, mientras que la "*residual deviance*" es la *deviance* ne modelo luego de que todas las variables fueran incluidas. En nuestro caso sería:

 $D^2 = (307.47 - 182.33)/307.47 = 0.41$

Es decir que el modelo ajustado explica un 41% de la variabilidad, mientras que un modelo ideal directamente no tiene *deviance* residual y su D^2 toma el valor 1. Este indicador toma más relevancia cuando se realizan comparaciones entre múltiples modelos que difieren, por ejemplo, en la cantidad y/o tipo de variables explicativas que se incluyen. En Guisan y Zimmermann (2000) se comparte bibliografía especializada y se discute con claridad su utilización en un marco de aplicación de GLMs para la predicción de distribución de hábitat.

Por otro lado, al evaluar el gráfico de dispersión y el "Q-Q plot" de los residuos observados en relación a los teóricos para una distribución binomial negativa, no se consideró

que la dispersión estadística reflejada en ellos impactara sustancialmente en la inferencia biológica del proceso en estudio (Figura 14 A y B).

En nuestro caso nos interesa particularmente el aporte de cada *cluster* o grupo ambiental (k_1-k_4) , es decir la intensidad de efecto de este factor, sobre la variación en el número de casas positivas por barrio entre períodos. Recordemos que el modelo lineal se relaciona con la variable respuesta a través de una función de enlace (*link*), permitiendo que la magnitud de la varianza de cada medición sea función de su valor predicho. En el caso de la familia de distribución binomial negativa, la función de enlace utilizada es *logit*, por lo que los coeficientes observados en la tabla corresponden a logaritmos de la razón de probabilidad. Podemos quitar el logaritmo y convertir el coeficiente en una intensidad de efecto más intuitiva aplicando en R la función exp(). En este caso, el coeficiente estimado para el "intercepto" corresponde al *cluster k*-I del año 2012, mientras que el siguiente es el del cluster k-I del año 2013, y así sucesivamente. Los resultados del GLM mostraron que la interacción global entre años y *clusters* ambientales fue estadísticamente significativa, tal como muestra la tabla de análisis de deviance arrojada (Tabla 4; LR Chisq: 45.094; Gl: 3; p < 0.001), lo que significa que el número de casas positivas está conjuntamente afectado por la interacción entre al menos un *cluster* ambiental y el año de muestreo. Para identificar cuál/es clusters era/n responsable/s de esa interacción, se realizaron comparaciones múltiples con corrección Bonferroni entre los téminos de interacción (cada año con cada *cluster* ambiental), mediante las cuales se pudo corroborar que sólo el *cluster k-1* albergó un incremento significativo, de gran magnitud, en el número de casas positivas entre los dos años (**Tabla 5**). Es decir, la media de ese contraste es significativamente diferente al resto, y el número de casas positivas en k-1 fue 9.52 veces mayor (p<0.005, ES: 0.29) en 2013 en relación a 2012 (Figura **13 B**). Como se dijo antes, la función de transformación del GLM es reversible y el valor de 9.52 proviene de aplicar la función exp() al valor de coeficiente para la interacción k1:2013, en relación al intercepto k1:2012, en el summary de la Tabla 3. No se observaron otros incrementos significativos entre las otras comparaciones pareadas.



Figura 14. Gráficos de ajuste del modelo lineal generalizado con familia de distribución binomial negativa. A: diagrama de dispersión de residuos observados vs. valores predichos. B: QQ-plot de residuos estandarizados vs. cuantiles teóricos.

	LR Chisq	GL	р
Año	14.931	1	0.0001115 ***
Cluster (k)	39.744	3	1.207e-08 ***
Interacción año*k	45.094	3	8.838e-10 ***

Tabla 4. Tabla de análisis de deviance, obtenida mediante la aplicación de un test de ANOVA (tipo 2) sobre el modelo lineal generalizado.

LR Chisq: test Chi-cuadrado de la razón de verosimilitud (*Likelihood Ratio Chi-square*). GL: grados de libertad Tabla 5. Comparaciones múltiples *a posteriori* entre los términos de interacción (cada año con su correspondiente *cluster* ambiental).

	2013/ <i>k</i>	2012/ <i>k</i> ll	2013/ <i>k</i> ll	2012/ <i>k</i> III	2013/ <i>k</i> III	2012/ <i>k</i> IV	2013/ <i>k</i> IV
2012/ <i>k</i> l	9.52 0.29	1.66 0.26	2.03 0.26	2.06 0.34	1.51 0.34	1.26 0.30	1.50 0.30
2013/ <i>k</i> l		0.17 0.23	0.21 0.23	0.22 0.31	0.16 0.32	0.13 0.28	0.16 0.27
2012/ <i>k</i> II			1.22 0.20	1.24 0.29	0.91 0.29	0.76 0.25	0.90 0.25
2013/ <i>k</i> II				1.02 0.29	0.75 0.29	0.62 0.25	0.74 0.24
2012/ <i>k</i> III					0.73 0.36	0.61 0.33	0.73 0.32
2013/ <i>k</i> III	Significativamente > 0 0.8					0.84 0.33	0.9 9 0.33
2012/ <i>k</i> IV		lo significativ	0				1.19
		Significativamente < 0					0.29

Negrita: magnitud del efecto; itálica: SE

Se toma como significativo a aquel contraste con p<0.005

4 - DISCUSIÓN

Está ampliamente aceptado que la ecología de sistemas de transmisión de enfermedades infecciosas que involucran organismos vectores es compleja y responde a cambios relacionados con las alteraciones del ambiente, como en el uso del suelo, los patrones de urbanización, el transporte y el clima (Weaver, 2013; Estrada-Peña et al., 2014). Por este motivo, resulta importante profundizar los conocimientos sobre los sistemas ecológicos de vectores en su contexto ambiental, lo cual implica adoptar una posición cooperativa con especialistas de diferentes campos de estudio como la microbiología, entomología, meteorología, ecología, urbanismo, ciencias sociales y políticas, Salud Pública, etc. En el caso de mosquitos vectores de enfermedades, uno de los retos que enfrentan las autoridades sanitarias en áreas urbanas es el monitoreo de potenciales sitios criadero de estados inmaduros (hábitats favorables para larvas y pupas), necesario para el diseño e implementación de medidas de control. En el presente estudio, evaluamos la utilidad de datos tomados mediante sensores remotos con el fin de explorar el uso del agrupamiento ambiental de barrios en una ciudad pequeña, a la hora de explicar incrementos en el número de casas con sitios infestados con formas inmaduras de *Ae. aegypti*.

Si bien Clorinda es una localidad pequeña, las diferencias observadas en el número de casas con contendedores infestados con larvas de Ae. aegypti (casas positivas) entre años, evidencian cierta estructura en los patrones de variación a lo largo de la ciudad. En un estudio previo, Espinosa et al. (2016 b) analizaron los números absolutos de casas positivas por año, comparando 5 áreas geográficas definidas arbitrariamente a priori por los autores. Para ensayar una mirada alternativa, en el presente estudio analizamos las diferencias entre los barrios muestreados, teniendo en cuenta los valores observados por estación. La variación en las densidades de casas positivas mostró no sólo que diferentes áreas pueden tener comportamientos divergentes (Figura 7), sino también que el aumento de densidad de casas positivas se extendió sobre casi toda la ciudad durante la primavera (Figura 8). Estos cambios se mostraron más intensos y localizados en zonas específicas durante el verano y el otoño, con valores máximos en aumento a lo largo del tiempo. De acuerdo con datos provistos por el Servicio Meteorológico Nacional, las temperaturas máximas durante el verano y el otoño fueron mayores en 2013 que en 2012 (Figura 15), con poca variabilidad en los valores promedio mensuales, lo cual resulta de interés ya que es sabido que las altas temperaturas favorecen y aceleran el ciclo de vida de Ae. aegypti (Brady et al., 2013). De todos modos, esta deducción no da cuenta de la estructuración espacial observada en los mapas de la **Figura 8**. Es decir, en conjunto las observaciones preliminares demuestran que cada zona en particular podría tener su propia dinámica espacio-temporal.



Figura 15. Temperaturas mínimas y máximas anuales en Clorinda (2012-2013). Datos provistos por el Servicio Meteorológico Nacional.

El cambio constante en la distribución espacial de zonas de alta densidad de *Ae. aegypti* constituye una dificultad significativa a la hora de identificar blancos clave para las acciones de control. Esta situación fue considerada previamente en Iquitos, Perú, donde la identificación de

áreas con altos niveles históricos de transmisión viral resultó más efectiva que la localización de sitios con alta densidad de mosquitos vectores (LaCon et al., 2014). En base a esta idea de dinámica espacio-temporal diferencial, hemos orientado nuestros análisis a investigar la posible asociación entre la evolución temporal de la abundancia de casas positivas y un *set* de capas de variables macro-ambientales obtenidas a partir de imágenes de sensado remoto. De manera interesante, los *clusters* o grupos ambientales obtenidos por "*k*-means" (**Figura 13 A**) fueron consistentes con las descripciones a campo provistas por la sede local de la Fundación Mundo Sano, según las cuales estas áreas tienen características bien distintivas. Tal como se expresó en Resultados, uno de los *clusters* ambientales (*k*-I) mostró un aumento significativo en el número de casas positivas entre los dos años bajo estudio.

Es bien sabido que, a excepción de su forma selvática ancestral, Ae. aegypti es una especie de mosquitos extremadamente antropófila. Es decir que estos insectos encuentran las condiciones favorables para el desarrollo de su ciclo de vida en los asentamientos humanos, especialmente en las viviendas (Powell y Tabachnick, 2013). En este sentido, se espera que variables ambientales relacionadas con áreas más urbanizadas sean buenas predictoras de la abundancia y distribución de Ae. aegypti, como se ha demostrado en estudios previos utilizando también sensores remotos. Fuller et al. (2010) usaron imágenes satelitales de media y alta resolución (ASTER y QuickBird) para mapear cobertura de árboles y áreas urbanizadas en una localidad de Costa Rica. Los autores generaron diferentes variables mediante algoritmos de clasificación y las mismas fueron analizadas en relación a la distribución de hábitats favorables. Mediante un modelo basado en regresión lineal múltiple revelaron que las áreas construidas explicaban un 63.6% de la variabilidad en el número de hábitats larvales, sugiriendo que las estructuras urbanas pueden ser usadas para predecir la presencia de este mosquito vector. Por otro lado, Landau y van Leeuwen (2012) realizaron un estudio similar en Arizona-EEUU, utilizando imágenes de un sensor multiespectral y de radar, gracias a las cuales pudieron observar que la presencia de Ae. aegypti estaba asociada de manera positiva con estructuras urbanas. A nivel local, Espinosa et al. (2016 a) realizaron muestreos sistemáticos en la ciudad de Tartagal (Salta) entre 2009 y 2014 y se basaron en un modelo predictivo de MaxEnt para asociar la abundancia de sitios criaderos de Ae. aegypti con diferentes variables ambientales. Los autores encontraron que la cobertura urbana estaba entre las tres principales variables explicativas de la distribución de Ae. aegypti. En nuestro estudio, el cluster k-1 se caracteriza por la presencia de galpones y talleres mecánicos, emplazados como grandes superficies construidas, intercaladas con viviendas familiares de baja altura. Por lo anteriormente expuesto sobre los hábitos antropofílicos de esta especie y la importancia de las construcciones urbanas como determinantes de su distribución, resulta lógico que la fisonomía urbana de este *cluster* aumente la heterogeneidad ambiental y la disponibilidad de sitios criaderos. Estas zonas urbanas están alejadas de las grandes superficies cubiertas por vegetación alta, típicas de las áreas cercanas a los grandes ríos. Es sabido que factores ligados a la vegetación afectan los patrones de distribución de Ae. aegypti de manera dependiente de las características del área analizada, y de manera independiente de la región geográfica que rodea el área de estudio. En otras palabras, son variables que dependen mucho del contexto geográfico de observación. Por ejemplo, Hayden et al. (2010) usaron trampas de oviposición para evaluar la importancia de factores micro-climáticos y sociales en la predicción de la distribución de Ae. aegypti en un ambiente árido. En ese caso, encontraron que la presencia de la especie estaba asociada de manera positiva con áreas de mucha vegetación alta dentro de cada barrio estudiado. En nuestro estudio, en cambio, el grupo k-1 no está asociado con superficies cubiertas por árboles. Si bien Clorinda está ubicada en la región chaqueña subtropical y húmeda, donde las condiciones físicas son óptimas para el desarrollo de ciclos de vida de diferentes especies de mosquitos, la vegetación típicamente espesa característica de esta región ofrece un nicho ecológico ideal para especies selváticas, pero no necesariamente para otras de hábitos domésticos que encuentran sus requerimientos de cría dentro y alrededor de estructuras urbanas. De manera similar, los cuerpos de agua naturales, como los ríos y humedales alrededor y dentro de Clorinda, no explicaron la distribución de criaderos de Ae. aegypti. Esto es debido a que no son zonas favorables para esta especie doméstica que elige aguas claras y estancas de recipientes artificiales para su oviposición. El almacenamiento de agua en tanques improvisados es una práctica recurrente en Clorinda, debida a la provisión irregular de agua potable. De hecho, esta práctica es mucho más evidente en los barrios 4, 30 y 5, incluidos en el grupo k-1. En muestreos anteriores, el barrio 4 (Primero de Mayo) mostró altos niveles de infestación en relación al resto de Clorinda (Garelli et al., 2009), y esos valores observados estuvieron asociados con la presencia de tanques de agua a nivel del suelo (Garelli et al., 2011). Como ya se mencionó, el cementerio de la ciudad se ubica también dentro de este *cluster*, ofreciendo gran disponibilidad de criaderos, ya que todavía es una práctica recurrente utilizar floreros con agua.

Nuestros resultados podrían indicar que la síntesis lograda con variables macroambientales es un buen predictor de una parte de los eventos que ocurren a un nivel microambiental y que un agrupamiento simple y robusto de los barrios en *clusters* ambientales permite una rápida y fácil identificación de las áreas más problemáticas dentro de la ciudad al momento de tomar decisiones operativas. Es decir que las variables ambientales observables mediante técnicas de sensado remoto pueden realizar aportes a la hora de identificar zonas de mayor abundancia de Ae. aegypti, pero al mismo tiempo es fundamental considerar factores socioculturales en cada caso bajo estudio, ya que resultan determinantes importantes de hábitats favorables. Si bien este tipo de aproximaciones puede ser útil para explicar, al menos en parte, la dinámica temporal de poblaciones de mosquitos, en un futuro inmediato deberían planificarse estudios de campo que incluyan la participación de la comunidad y la interacción reticulada entre sectores académicos especializados en la realización de encuestas demográficas y socioeconómicas durante el monitoreo de criaderos domésticos. Obenauer et al. (2017) usaron MaxEnt para modelar la distribución de Ae. aegypti en el Sudeste de Estados Unidos. En este caso evaluaron los efectos de agregar indicadores de "pobreza" (sit) y "densidad poblacional" como variables adicionales a otros modelos basados sólo en datos meteorológicos, demostrando que los modelos que incluían las variables socio-demográficas exhibían mayor poder predictivo y producían mapas de distribución más precisos. Más aún, las variables relacionadas a las características de la población humana hicieron mayores aportes que las variables climáticas a estos modelos. Según la Organización Mundial de la Salud (OMS, 2009) debemos considerar otros varios factores además de los ambientales, como determinantes de la vulnerabilidad de una comunidad ante las epidemias de dengue. En este sentido, variables como densidad de la población humana, características de los asentamientos urbanos, tipo de viviendas, nivel de educación y condición socioeconómica, son importantes para evaluar riesgos epidémicos y planear acciones de control. Como ya se mencionó, Espinosa et al. (2016 a) asociaron un déficit en la provisión de agua potable con mayores abundancias de criaderos de mosquitos en áreas particulares de Clorinda. Sin embargo, este tipo de información es registrada sólo cada 10 años en nuestro país, a través del Censo Nacional de Población. Además, la escala geográfica a la que se informan los datos (radios censales) no es lo suficientemente fina como para modelar la distribución de especies de mosquitos a niveles de micro-hábitat.

Consideraciones finales:

En síntesis, cuando los tomadores de decisiones en Salud Pública se enfrentan a problemas relacionados a enfermedades transmitidas por vectores, normalmente el objetivo principal de los analistas geo-espaciales es construir un mapa de riesgo basado en información geográfica *in situ* y variables ambientales obtenidas por diferentes medios, incluidos los satélites de observación de la Tierra. Un supuesto típicamente implícito es que un mapa dado, incluso si fue obtenido para un tiempo específico, podría ser válido durante un periodo más prolongado, como si los valores respuesta fueran estáticos en el tiempo. Sin embargo, hemos observado que el patrón de abundancia de casas positivas es variable en el espacio y el tiempo. Estas diferencias podrían atribuirse a:

- I) Respuestas diferenciales determinadas por la interacción entre variables macroambientales, como las que analizamos en el presente estudio.
- II) El efecto de procesos sólo observables a nivel de micro-hábitat. Las aplicaciones de sensado remoto a problemas epidemiológicos todavía se encuentran en fase de investigación y fortalecimiento. Todavía se están investigando cuáles son las variables que afectan la distribución de especies vectoras y cuál es la escala relevante, determinada por la extensión geográfica y el nivel de resolución requerido por la pregunta planteada (Phillips et al., 2006).
- III) La dificultad de obtener un panorama preciso de la situación entomológica en muestreos realizados a lo largo de un asentamiento urbano durante cierto tiempo.

Estudios futuros deberían tener en cuenta la observación de características ambientales a nivel de micro-hábitat que podrían afectar la disponibilidad de criaderos de mosquitos, muestreos entomológicos longitudinales sostenidos y la incorporación de factores sociodemográficos como variables explicativas. Finalmente, es importante mencionar que los estudios transversales sobre presencia y abundancia de *Ae. aegypti* históricamente no han dado buenos resultados a la hora de encontrar buenas correlaciones entre índices entomológicos y episodios epidémicos de dengue (Cromwell et al., 2017) y ningún umbral de presencia de larvas en domicilios ha servido hasta el momento para predecir epidemias de virus transmitidos por esta especie (Bowman et al., 2014). Esto puede deberse a que el ciclo de transmisión del virus Dengue es complejo y varía a lo largo del espacio y el tiempo, entonces es muy improbable caracterizar de manera adecuada la relación entre la densidad de mosquitos y el riesgo de infección a lo largo de monitoreos entomológicos periódicos. De todos modos, las nuevas tecnologías de sensado remoto, herramientas informáticas y métodos de modelado, pueden ser de gran ayuda para mapear la distribución espacial de vectores y/o para predecir su diseminación y dinámica estacional de manera relativamente económica. Un buen entendimiento de la tipología y productividad de hábitats adecuados para el desarrollo larval de mosquitos es esencial a la hora de diseñar estrategias de control entomológico. Nuestros resultados apoyan la idea de que estos esfuerzos aportan evidencias parciales y no soluciones finales a una problemática compleja que requiere de abordajes tan diversos como la ecología, el diseño de estrategias de control, la movilización social con educación sanitaria, participación comunitaria, estrategias de comunicación, como así también del compromiso multisectorial de ministerios, municipalidades, organizaciones no-gubernamentales y el sector productivo privado.

BIBLIOGRAFÍA

Aguiar M, Stollenwerk N. 2017. Dengvaxia: age as surrogate for serostatus. The Lancet. doi.org/10.1016/S1473-3099(17)30752-1

Albrieu-Llinás G y Gardenal CN. 2012. Phylogeography of *Aedes aegypti* in Argentina: longdistance colonization and rapid restoration of fragmented relicts after a continental control campaign. Vector-Borne and Zoonotic Diseases, 12(3), 254-261.

Almirón WR, Asís R. 2003. Índice de abundancia de larvas y pupas de *Aedes aegypti* (Diptera: Culicidae) en la ciudad de Córdoba. Revista de la Facultad de Ciencias Médicas (Córdoba) 60: 37-41.

Andreo V, Neteler M, Rocchini D, Provensal C, Levis S,..., Polop J. 2014. Estimating Hantavirus Risk in Southern Argentina: A GIS-Based Approach Combining Human Cases and Host Distribution. Viruses 6: 201-222. doi: 10.3390/v6010201.

Andreo V, Belgiu M, Brito Hoyos D, Osei F, Provensal C, Stein A. 2019. Rodents and satellites: Predicting mice abundance and distribution with Sentinel-2 data. Ecological Informatics 51: doi.org/10.1016/j.ecoinf.2019.03.001

Armstrong D, 2016. Factorplot: graphical presentation of simple contrasts. R J 5:4-15. Available from: <u>https://journal.rproject.org/archive/2013/RJ.../RJ-2013-021.pdf</u>

Barnes CM, Cibula WG. 1979. Some implications of remote sensing technology in insect control programs including mosquitoes. Mosq News 39: 271-282.

Bavia ME, Carneiro DD, Gurgel Hda C, Madureira Filho C, Barbosa MG. 2005. Remote Sensing and Geographic Information Systems and risk of American visceral leishmaniasis in Bahia, Brazil. Parassitologia 47:165-169.

Beck LR, Rodriguez MH, Dister SW, Rodriguez AD, Washino RK, Roberts. 1997. Assessment of a remote sensing-based model for predicting malaria transmission risk in villages of Chiapas, Mexico. Am. J Trop. Med & Hyg 56: 99-106.

Beck LR, Lobitz BM, Wood BL. 2000. Remote sensing and human health: new sensors and new opportunities. Emerging Infectious Diseases, 6: 217-226

Bowman LR, Donegan S, McCall PJ. 2016. Is dengue vector control deficient in effectiveness or evidence? Systematic review and meta-analysis. PLoS Neglected Tropical Diseases 10:1-24.

Bowman LR, Runge-Ranzinger S, McCall PJ. 2014. Assessing the relationship between vector indices and dengue transmission: a systematic review of the evidence. PLoS Neglected Tropical Diseases 8: doi.org/10.1371/journal.pntd.0002848 PMID: 24810901

Brady OJ, Johansson MA, Guerra CA, Bhatt S, Golding N,..., Hay SI. 2013. Modelling adult Aedes aegypti and Aedes albopictus survival at different temperatures in laboratory and field settings. Parasites & Vectors 6:351.

Brown HE, Cox J, Comrie AC, Barrera R. 2017. Habitat and density of oviposition opportunity influences *Aedes aegypti* (Diptera: Culicidae) flight distance. Journal of Medical Entomology doi.org/10.1093/jme/tjx083.

Cabin RJ, Mitchell RJ. 2000. To Bonferroni or not to Bonferroni: when and how are the questions. Bulletin of the Ecological Society of America 81: 246-248.

Christophers RC. 1960. *Aedes aegypti*. The Yellow Fever Mosquito: Its life history, bionomics and structure. Cambridge: Cambridge University Press 739 pp.

Chuvieco E. 2010. Teledetección Ambiental, 3a Edición. Editorial Ariel 590 pp.

Chuvieco E. 2016. Fundamentals of satellite remote sensing: An environmental approach. CRC press 416 pp.

Cline BL. 1970. New eyes for the epidemiologists: aerial photography and other remote sensing techniques. Journal of Epidemiology, 92: 85-89.

Cohen J. 1960. A Coefficient of Agreement for Nominal Scales. Educational and Psychological Measurement 20: 37–46.

Cromwell EA, Stoddard ST, Barker CM, Van Rie A, Messer WB, Meshnick SR, et al. 2017. The relationship between entomological indicators of Aedes aegypti abundance and dengue virus infection. PLoS Neglected Tropical Diseases 11: doi.org/10.1371/journal.pntd.0005429 PMID: 28333938.

Dale PER, Ritchie SA, Territo BM, Morris CD, Muhar A, Kay BH. 1998. An overview of remote sensing and GIS for surveillance of mosquito vector habitats and risk assessment. J Soc Vector Ecol 23: 54-61.

Dormann CF, Elith J, Bacher S, Buchmann C, Carl G, et al. 2013. Collinearity: a review of methods to deal with it and a simulation study evaluating their performance. Ecography 36:27-46.

Edman JD, Strickman D, Kittayapong P, Scott TW. 1992. Female *Aedes aegypti* (Diptera: Culicidae) in Thailand rarely feed on sugar. Journal of Medical Entomology 29: 1035-1038.

Espinosa M, Weinberg D, Rotela CH, Polop F, Abril M, Scavuzz CM. 2016a. Temporal dynamics and spatial patterns of Aedes aegypti breeding sites, in the context of a Dengue control program in Tartagal (Salta province, Argentina). PLoS Neglected Tropical Diseases 10:e0004621.

Espinosa MO, Polop F, Rotela CH, Abril M, Scavuzzo CM. 2016b. Spatial pattern evolution of Aedes aegyptibreeding sites in an Argentinean city without a dengue vector control programme. Geospatial Healthzz doi.org/10.4081/gh.2016.471.

Estrada-Peña A, Ostfeld RS, Peterson AT, Poulin R, de la Fuente J. 2014. Effects of environmental change on zoonotic disease risk: An ecological primer. Trends in Parasitology 30:205-14.

Fox J. 2015. Applied Regression Analysis and Generalized Linear Models. SAGE Publications, California, USA, 791 pp.

Fuller DO, Troyo A, Calderón-Arguedas O, Beier JC. 2010. Dengue vector (Aedes aegypti) larval habitats in an urban environment of Costa Rica analysed with ASTER and QuickBird imagery. International Journal of Remote Sensing 31 :3-11.

Garelli FM, Espinosa MO, Weinberg D, Coto HD, Gaspe MS, Gürtler RE. 2009. Patterns of *Aedes aegypti* (Diptera: Culicidae) Infestation and Container Productivity Measured Using Pupal and Stegomyia Indices in Northern Argentina. Journal of Medical Entomology 46:1176-86.

Garelli FM, Espinosa MO, Weinberg D, Trinelli MA, Gürtler RE. 2011. Water use practices limit the effectiveness of a temephos-based *Aedes aegypti* larval control program in northern Argentina. PLoS Neglected Tropical Diseases 5:e991.

Gatherer D, Kohl A, 2016. Zika virus: a previously slow pandemic spreads rapidly through the Americas. Journal of General Virology 97:269-73.

Gómez-Bravo A, Germán A, Abril M, Scavuzzo M, Salomón OD. 2017. Spatial population dynamics and temporal analysis of the distribution of Lutzomyia longipalpis (Lutz & Neiva, 1912) (Diptera: Psychodidae: Phlebotominae) in the city of Clorinda, Formosa, Argentina. Parasites & Vectors 10, <u>doi.org/10.1186/s13071-017-2296-0</u>.

González AS. 2001. El impacto de la enfermedad en la organización social y el espacio urbano: el caso de la epidemia de Fiebre Amarilla en la ciudad de Buenos Aires en 1871. Medicina y Sociedad 24: 93-102.

Gottdenker NL, Streicker DG, Faust CL, et al. 2014. Anthropogenic Land Use Change and Infectious Diseases: A Review of the Evidence. EcoHealth 11: 619–632.

Gubler DJ. 1998. Dengue and Dengue hemorrhagic fever. Clinical Microbiology Reviews 11: 480-496.

Guisan A, Zimmermann NE. 2000. Predictive habitat distribution models in ecology. Ecological Modelling 135: 147-186.

Gürtler RE, Garelli FM, Coto HD, 2009. Effects of a Five-Year Citywide Intervention Program to control Aedes aegypti and prevent Dengue outbreaks in northern Argentina. PLoS Neglectged Tropical Diseases 3:e427.

Gwadz RW. 1969. Regulation of blood meal size in the mosquito. Journal of Insect Physiology 15: 2039-2044.

Hartigan JA, Wong MA, 1979. Algorithm AS 136: A K-Means clustering algorithm. Applied Stattistics 28:100-108.

Hayden MH, Uejio CK, Walker K, Ramberg F, Moreno R, Rosales C, Gameros M, Mearns LO, Zielinski-Gutierrez E, Janes CR. 2010. Microclimate and human factors in the divergent ecology of Aedes aegypti along the Arizona, U.S./Sonora, MX Border. EcoHealth 7:64-77.

Hayes RO, Maxwell EL, Mitchell CJ, Woodzick TL. 1985. Detection, identification, and classification of mosquito larval habitats using remote sensing scanners in earth-orbiting satellites. Bulletin of the World Health Organization 63: 361-374.

He Y, Chen G, Potter C, Meentemeyer RK. 2019. Integrating multi-sensor remote sensing and species distribution modeling to map the spread of emerging forest disease and tree mortality. Remote Sensing of Environment, doi.org/10.1016/j.rse.2019.111238.

Herbreteau V, Salem G, Souris M, Hugot JP, Gonzalez JP. Thirty years of use and improvement of remote sensing, applied to epidemiology: from early promises to lasting frustration. Health & Place doi: 10.1016/j.healthplace.2006.03.003

Higgs S, Vanlandingham D, 2015. Chikungunya virus and its mosquito vectors. Vector-Borne and Zoonotic Diseases 15:231-40.

Honório NA, Costa-Silva W, Leite PJ, Goncalves JM, Lounibos LP, Lorenço-de-Olivera R. 2003. Dispersal of *Aedes aegypti* and *Aedes albopictus* (Diptera: Culicidae) in an urban endemic dengue area in the state of Rio de Janeiro, Brasil. Memórias do Instituto Oswaldo Cruz 98: 191-198.

Hothorn T, Bretz F, Westfall P, 2008. Simultaneous inference in general parametric models. Biometrical Journal 50:346-63.

Kamal M, Kenawy MA, Rady MH, Khaled AS, Samy AM. 2018. Mapping the global potential distributions of two arboviral vectors *Aedes aegypti* and *Ae. albopictus* under changing climate PLoS ONE. doi.org/10.1371/journal.pone.0210122.

Kay BH, Ryan PA, Russell BM, Holt JS, Lyons SA, Foley PN. 2000. The importance of subterranean mosquito habitat to arbovirus vector control strategies in North Queensland, Australia. Journal of Medical Entomology 36: 846-853.

Kirby RS, Delmelle E, Eberth JM. 2017. Advances in spatial epidemiology and geographic information systems. Annals of epidemiology 27: 1-9.

LaCon G, Morrison AC, Astete H, Stoddard ST, Paz-Soldan VA, ..., Vazquez-Prokopec G. 2014. Shifting patterns of Aedes aegypti fine scale spatial clustering in Iquitos, Peru. PLoS Neglected Tropical Diseases. 8:e3038.

Landau KI, van Leeuwen WJD. 2012. Fine scale spatial urban land cover factors associated with adult mosquito abundance and risk in Tucson, Arizona. Journal of Vector Ecology 37: 407-18.

Liang G, Gao X, Gould EA, 2015. Factors responsible for the emergence of arboviruses; strategies, challenges and limitations for their control. Emerging Microbes & Infections doi.org/10.1038/emi.2015.18

Liu Y, Liu J, Du S, Shan C, Nie K, ... & Wang P. 2017. Evolutionary enhancement of Zika virus infectivity in *Aedes aegypti* mosquitoes. Nature 545: 482.

Lounibos L. 2002. Invasions by insect vectors of human disease. Annual Review of Entomology 47: 233-266.

MacQueen, J. 1967. Some methods for classification and analysis of multivariate observations.

Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, Volume 1: Statistics, 281-297, University of California Press. Berkeley, California. Maechler M, Rousseeuw P, Struyf A, Hubert M, Hornik K. 2017. Cluster: Cluster analysis basics and extensions. R package version 2.0.6.

Magurran, A.E. (2013). Measuring biological diversity. Wiley & Sons.

Marquardt WC. 2004. Biology of disease vectors. Elsevier Academic Press. San Diego, California. 800pp.

Muturi E, Dunlap C, Ramirez JL, Rooney AP, Kim CH. 2018. Host blood-meal source has a strong impact on gut microbiota of *Aedes aegypti*. FEMS Microbiology Ecology <u>https://doi.org/10.1093/femsec/fiy213</u>.

Nakagawa S. 2004. A farewell to Bonferroni: the problems of low statistical power and publication bias. Behavioral Ecology 15: 1044-1045.

Nelson, M.J. 1986. Aedes aegypti: biología y ecología. Organización Panamericana de la Salud – Programa de enfermedades transmisibles 50 pp.

Obenauer JF, Joyner TA, Harris JB. 2017. The importance of human population characteristics in modeling *Aedes aegypti* distributions and assessing risk of mosquito-borne infectious diseases. Tropical Medicine and Health 45: 38. doi: 10.1186/s41182-017-0078-1.

Organización Mundial de la Salud. 2016. Dengue and severe dengue. Factsheet N°117. Disponible en: <u>http://www.who.int/mediacentre/factsheets/fs117/.</u>

Ostfeld RS, Glass GE, Keesing F, 2005. Spatial epidemiology: An emerging (or re-emerging) discipline. Trends in Ecology and Evolution 20:328-36.

Palaniyandi M, Anand PH, Maniyosai R. 2014. Spatial cognition: a geospatial analysis of vector borne disease transmission and the environment, using remote sensing and GIS. Int. J of Mosquito Research 1: 39-54.

Palaniyandi M, Anand PH, Pavendar T. 2017. Environmental risk factors in relation to occurrence of vector borne disease epidemics: Remote sensing and GIS for rapid assessment, picturesque, and monitoring towards sustainable health. Int. J of Mosquito Research 4: 9-20.

Patz JA, Githeko AK, McCarty JP, Hussein S, Confalonieri U, De Wet N. 2003. Climate change and infectious diseases. Climate change and human health: risks and responses 2: 103-32.

Phillips S, Anderson R, Schapire R. 2006. Maximum entropy modeling of species geographic distributions. Ecological Modelling 190: 231-59.

Porcasi X, Andreo V, Ferral A, Guimarey P, Santini MS, ..., Aguirre E. 2018. Regional scale environmental variables complementing a Risk Model of Chagas Disease vectorial transmission. En: 2018 IEEE Biennial Congress of Argentina (ARGENCON) (pp. 1-5). IEEE.

Powell JR, Tabachnick WJ. 2013. History of domestication and spread of Aedes aegypti - a review. Memorias do Instituto Oswaldo Cruz 108: 11-7.

Reisen WK, 2010. Landscape epidemiology of vector-borne diseases. Annual Reviews of Entomology 55:461-83.

Reiter P. 1996. Oviposition and dispersion of *Aedes aegypti* in an urban enviroment. Bulletin of the Exotic Pathology Society 89: 120-122.

Reiter P, Amador MA, Anderson RA, Clark GG. 1995. Dispersal of *Aedes aegypti* in an urban area after blood feeding as demonstrated by rubidium-marked eggs. American Journal of Tropical Medicine and Hygiene 52: 177-179.

Roiz D, Neteler M, Castellani C, Arnoldi D, Rizzoli A. 2011. Climatic factors driving invasion of the tiger mosquito (Aedes albopictus) into new areas of Trentino, northern Italy. PloS One, 6(4).

Rossi GC, Almirón WR, 2004. Clave ilustrada para la identificación de larvas de mosquitos de interés sanitario encontradas en criaderos artificiales en la Argentina. Buenos Aires: Fundación Mundo Sano.

Rotela C, López L, Frías Céspedes M, Barbás G, Lighezzolo A,..., Gorla DE. 2017. Analytical report of the 2016 dengue outbreak in Córdoba city, Argentina. Geospatial Health 12: doi.org/10.4081/gh.2017.564

Savage HM, Smith GC. 1995. *Aedes albopictus* and *Aedes aegypti* in the new World. Journal of the American Mosquito Control Association 10: 440-442.

Scott TW, Chow E, Strickman D, Kittayapong P, Wirtz RA,..., Edman JD. 1993. Blood-feeding patterns of *Aedes aegypti* (Diptera: Culicidae) collected in a rural Thai village. Journal of Medical Entomology 30: 922-927.

Stefani A, Dusfour I, Correa AP, Cruz MCB, Dessay N, ..., Roux E. 2013. Land cover, land use and malaria in the Amazon: a systematic literature review of studies using remotely sensed data. Malaria Journal 12: 192.

Stein M, Oria GI, Almirón WR. 2002. Principales criaderos para *Aedes aegypti* y culícidos asociados, Argentina. Revista de Saúde Pública 36: 627-630.

Tabachnick WJ. 1991. Evolutionary genetics and arthropod-borne disease. The yellow fever mosquito. American Entomologist 37: 14-24.

Thomson MC, Connor SJ. 2000. Environmental information systems for the control of arthropod vectors of disease. Medical and Veterinary Entomology, 14: 227-244.

Toledo ME, Rodriguez A, Valdés L, Carrión R, Cabrera G.,..., Van der Stuyft P. 2011. Evidence on impact of community-based environmental management on dengue transmission in Santiago de Cuba. Tropical Medicine and International Health 16:744-7.

Tran A, Ponçon N, Toty C, Linard C, Guis H, Ferré JB. 2008. Using remote sensing to map larval and adult populations of Anopheles hyrcanus (Diptera: Culicidae) a potential malaria vector in Southern France. Int J Health Geogr 26: 7-9.

Tran A, Kassié D, Herbreteau V. 2016. Applications of remote sensing to the epidemiology of infectious diseases: some examples. Editores: Nicolas Baghdadi, Mehrez Zribi. Land Surface Remote Sensing, Elsevier. Páginas: 295-315, ISBN 9781785481055.

Vogel KJ, Brown MR, Strand MR. 2015. Ovary ecdysteroidogenic hormone requires a receptor tyrosine kinase to activate egg formation in the mosquito *Aedes aegypti*. PNAS doi.org/10.1073/pnas.1501814112.

Weaver S, Barrett A. 2004. Transmission cycles, host range, evolution and emergence of arboviral disease. Nature Reviews Microbiology 2: 789–801.

Weaver SC. 2013. Urbanization and geographic expansion of zoonotic arboviral diseases: Mechanisms and potential strategies for prevention. Trends in Microbiology 21:360-3.

Wei T y Simko V. 2017. R package "corrplot": Visualization of a Correlation Matrix (Version 0.84). Disponible en <u>https://github.com/taiyun/corrplot</u>

Weinberg D, Porcasi X, Lanfri S, Abril M, Scavuzzo CM. 2018. Spatial analyzes of triatomine infestation indices and their association to the actions of a Chagas disease program and environmental variables during a 5-year intervention period. Acta Tropica 188: 41-49. doi.org/10.1016/j.actatropica.2018.08.025.

Zou L, Miller SN, Schmidtmann ET. 2006. Mosquito larval habitat mapping using remote sensing and GIS: implications of coal bed methane development and West Nile virus. J Med Entomol 43: 1034-41.

Zuur AF, Ieno EN, Smith GM. 2007. Analysing Ecological Data. Springer, NY, USA. 672 pp.

Zuur AF, Ieno EN, Elphick CS. 2010. A protocol for data exploration to avoid common statistical problems. Methods in Ecology and Evolution 1: 3 -14.

ANEXO I – Tabla de diseño de muestreos

Los muestreos seleccionados fueron apareados de acuerdo a las fechas, para poder contar con recorridos comparables por estación. M1: Octubre/2011 a Febrero/2012 (se lo considera verano de "2012" en la codificación de la tabla siguiente). M5: Noviembre 2012 a Marzo 2013 (se lo considera verano de "2013" en la tabla siguiente). M2: Febrero/2012 a Junio/2012 (Otoño de 2012). M6: Marzo/2013 a Mayo/2013 (Otoño de 2013). M4: Agosto/2012 a Noviembre/2012 (Primavera de 2012). M8 y M9: Agosto/2013 a Noviembre/2013 (Primavera de 2013).

	Verano		Otoño		Primavera		
Barrio	M1	M5	M2	M6	M4	M8	M9
1	17/10 - 21/10	8/11 - 9/11	16/2 - 22/2 y 15/3	4/3 – 8/3	5/9 – 10/9	13/8 – 15/8	7/10 – 9/10
2	24/10	12/11	16/3	12/3	10/9 – 11/9	15/8 – 16/8	9/10
3	28/10	12/11-13/11	19/3	15/3 y 25/3	11/9 – 12/9	16/8	16/10
4	15/11-14/12	3/12-18/12	17/5 – 4/6	15/4 – 26/4	12/10 – 19/10	6/9 – 12/9	21/11 – 28/11
5	5/12	4/3	4/6	3/5	5/11	26/9 – 27/9	19/11
6	15/2-16/2	27/2-/1/3	15/5 – 16/5	5/5 – 7/5	22/10 – 24/10	30/9 – 2/10	15/11
7	23/1-25/1	26/2-27/2	27/3 – 29/3	6/4 – 8/4	24/10 – 30/10	23/9 – 25/9	13/11 – 14/11
8	25/1-26/1	10/1-11/1	23/3	5/4	16/10	30/9	18/11
9	26/10-27/10	16/11	20/3 – 29/3	21/3 – 22/3	19/9 – 20/9	20/8	11/10
10	25/10-27/10	15/11	30/3 – 3/4	19/3 – 25/3	14/9 – 18/9	22/8 – 23/8	22/10 y 28/10
11	3/11-4/11	13/11-14/11	13/4 – 16/4	18/3 – 19/3	13/9 – 14/9	23/8 y 27/8	17/10 – 21/10
12	28/10-3/11	19/11-21/11	30/3 – 4/4	22/3 – 26/3	17/9 – 19/9	28/8 – 30/8	17/10 – 23/10
13	12/1	19/12	6/3 – 7/3	2/5	4/9	19/9	23/10 y 28/10
14	6/11	27/12	24/2	7/3	3/9	19/9	7/11
15	7/11-8/11	20/12-21/12	24/2	6/3	3/9	20/9	7/11
16	22/11-24/11	14/12-17/12	28/2 – 5/3	7/3 – 14/3	2/10 – 5/10	2/10 – 4/10	7/11 – 11/11
17	9/11	22/11	19/4	27/3	27/9	30/8	29/10 - 30/10
18	7/11-8/11	27/11	18/4	3/4	26/9	2/9 – 3/9	28/10 – 29/10
19	8/11	22/11	17/4	3/4	27/9	2/9	30/10
20	21/12-22/12	30/11	7/3	11/4	4/9 – 5/9	5/9	4/11
21	19/12-20/12	12/12	8/3 – 13/3	11/4 y 3/5	20/9 – 21/9	5/9	4/11
22	13/12-15/12	30/11	12/3	10/4	25/9	4/9	1/11
23	6/12-7/12	27/11-28/11	24/4 – 2/5	3/4 – 9/4	28/9 – 1/10	21/8 y 3/9	31/10 – 1/11
24	26/12	7/1	23/4	29/4	11/10	13/9	4/11/ - 5/11
25	27/12-28/12	2/1-4/1	3/5 – 4/5	29/4 – 30/4	9/10 - 10/10	17/9 – 18/9	5/11
26	29/12	4/1	4/5	30/4	11/10	13/9	5/11
27	26/12-27/12	2/1	12/3 – 13/3	10/4	5/10 – 9/10	4/9	1/11
28	10/1	8/1-9/1	4/5 – 6/5	3/5	11/10	12/9 – 13/9	6/11
29	20/1-23/1	1/3	11/5	No se muestreó	1/11 – 5/11	20/9 – 23/9	15/11 – 18/11
30	30/1-31/1	1/3	14/5 -15/5	7/5	31/10	25/9	13/11
31	26/1	10/1	16/5	2/5	31/10	27/9	11/11
32	28/10-31/10	13/11	29/3	15/3	12/9	16/8 – 17/8	9/10

ANEXO II – Tabla de datos usados en el modelo lineal generalizado

Barrio	Casas positivas	Total	Cluster (k)	Estación	Año
1	29	121	4	verano	2012
2	1	45	Д	verano	2012
2	6	24	-т И	verano	2012
3	109	24	4	verano	2012
4	100	25	1	verano	2012
5	2	20	1	verano	2012
6	13	86	3	verano	2012
7	15	86	3	verano	2012
8	12	60	2	verano	2012
9	10	89	2	verano	2012
10	8	53	2	verano	2012
11	6	56	1	verano	2012
12	38	95	2	verano	2012
13	0	30	2	verano	2012
14	1	19	1	verano	2012
15	0	26	1	verano	2012
16	37	198	2	verano	2012
17	5	26	4	verano	2012
18	8	48	2	verano	2012
19	0	20	4	verano	2012
20	4	46	2	verano	2012
21	3	67	1	verano	2012
22	6	50	2	verano	2012
23	a	105	_ ⊿	verano	2012
20	6	30	т 2	Verano	2012
24	0	100	5	Verene	2012
20	3	102	2	verano	2012

26	0	16	3	verano	2012
27	0	23	2	verano	2012
28	4	39	2	verano	2012
32	2	33	2	verano	2012
33	2	53	1	verano	2012
35	1	23	3	verano	2012
37	7	23	4	verano	2012
1	33	139	4	verano	2013
2	2	43	4	verano	2013
3	5	39	4	verano	2013
4	47	180	1	verano	2013
5	42	43	1	verano	2013
6	13	66	3	verano	2013
7	28	75	3	verano	2013
8	8	73	2	verano	2013
9	16	81	2	verano	2013
10	14	53	2	verano	2013
11	53	64	1	verano	2013
12	13	90	2	verano	2013
13	16	34	2	verano	2013
14	19	23	1	verano	2013
15	41	43	1	verano	2013
16	52	173	2	verano	2013
17	3	27	4	verano	2013
18	9	37	2	verano	2013
19	3	13	4	verano	2013
20	4	19	2	verano	2013
21	62	64	1	verano	2013
22	6	49	2	verano	2013
23	32	86	4	verano	2013
24	4	21	3	verano	2013
25	15	64	2	verano	2013

26	0	17	3	verano	2013
27	4	21	2	verano	2013
28	5	49	2	verano	2013
32	4	25	2	verano	2013
33	23	25	1	verano	2013
35	1	23	3	verano	2013
37	3	30	4	verano	2013
1	20	127	4	otoño	2012
2	0	36	4	otoño	2012
3	0	39	4	otoño	2012
7	38	75	3	otoño	2012
8	10	62	2	otoño	2012
9	23	87	2	otoño	2012
10	28	57	2	otoño	2012
12	38	117	2	otoño	2012
13	10	34	2	otoño	2012
14	2	25	1	otoño	2012
15	5	27	1	otoño	2012
16	31	206	2	otoño	2012
20	3	29	2	otoño	2012
21	3	71	1	otoño	2012
22	5	36	2	otoño	2012
27	8	41	2	otoño	2012
37	7	24	4	otoño	2012
1	33	142	4	otoño	2013
2	7	47	4	otoño	2013
3	9	48	4	otoño	2013
7	17	98	3	otoño	2013
8	22	56	2	otoño	2013
9	28	84	2	otoño	2013
10	14	55	2	otoño	2013
12	26	166	2	otoño	2013

13	5	32	2	otoño	2013
14	23	25	1	otoño	2013
15	38	40	1	otoño	2013
16	41	174	2	otoño	2013
20	17	28	2	otoño	2013
21	58	62	1	otoño	2013
22	2	22	2	otoño	2013
27	4	28	2	otoño	2013
37	5	25	4	otoño	2013
1	2	110	4	primavera	2012
2	2	40	4	primavera	2012
3	1	38	4	primavera	2012
4	13	340	1	primavera	2012
5	1	55	1	primavera	2012
6	20	59	3	primavera	2012
7	11	98	3	primavera	2012
8	1	40	2	primavera	2012
9	2	66	2	primavera	2012
10	9	45	2	primavera	2012
11	3	77	1	primavera	2012
12	5	88	2	primavera	2012
13	0	13	2	primavera	2012
14	0	16	1	primavera	2012
15	0	40	1	primavera	2012
16	16	187	2	primavera	2012
17	1	22	4	primavera	2012
18	1	32	2	primavera	2012
19	2	12	4	primavera	2012
20	5	43	2	primavera	2012
21	5	62	1	primavera	2012
22	0	41	2	primavera	2012
23	5	90	4	primavera	2012

24	0	24	3	primavera	2012
25	4	76	2	primavera	2012
26	1	30	3	primavera	2012
27	3	17	2	primavera	2012
28	1	25	2	primavera	2012
32	3	38	2	primavera	2012
33	7	47	1	primavera	2012
35	3	14	3	primavera	2012
37	1	22	4	primavera	2012
1	6	191	4	primavera	2013
2	0	65	4	primavera	2013
3	1	51	4	primavera	2013
4	40	448	1	primavera	2013
5	21	77	1	primavera	2013
6	3	100	3	primavera	2013
7	8	134	3	primavera	2013
8	7	121	2	primavera	2013
9	3	112	2	primavera	2013
10	1	64	2	primavera	2013
11	141	154	1	primavera	2013
12	14	155	2	primavera	2013
13	4	54	2	primavera	2013
14	29	32	1	primavera	2013
15	44	46	1	primavera	2013
16	18	345	2	primavera	2013
17	4	65	4	primavera	2013
18	5	80	2	primavera	2013
19	1	18	4	primavera	2013
20	4	36	2	primavera	2013
21	53	61	1	primavera	2013
22	13	48	2	primavera	2013
23	7	127	4	primavera	2013

24	4	40	3	primavera	2013
25	7	141	2	primavera	2013
26	5	33	3	primavera	2013
27	2	69	2	primavera	2013
28	0	28	2	primavera	2013
32	2	72	2	primavera	2013
33	20	21	1	primavera	2013
35	0	35	3	primavera	2013
37	1	48	4	primavera	2013

ANEXO III - Tak	bla de valores	medios de variables	ambientales p	or barrio
-----------------	----------------	---------------------	---------------	-----------

id	DCA	DH	DCULT	DSD	DAU	DVA	DVB	%CA	%H	%CULT	%SD	%AU	%VA	%VB	Sh
1	17,42318738	7,729876118	1,309428709	12,6243389	5,723680503	6,437932995	1,858996012	0,0656	0,07618	0,31477	0,02811	0,14781	0,08115	0,28638	1,638
2	8,924973422	2,549363945	1,117399716	13,31582765	5,236621109	2,124546631	1,028969625	0,04742	0,11407	0,2798	0,00609	0,06115	0,18022	0,31125	1,617
3	4,656979183	1,167109578	1,652163286	18,58871204	9,280243422	2,132052372	1,996018336	0,12361	0,2986	0,14237	0,01037	0,03509	0,19874	0,19123	1,703
4	74,71421948	29,86053317	1,350017288	1,92201297	0,444649273	13,44043576	1,467753552	0,00016	0,00268	0,18506	0,14337	0,44501	0,00587	0,21785	1,318
5	75,92458052	24,70169833	0,383917814	4,906889525	2,098570772	32,37723936	2,701535834	0	0,01234	0,48884	0,07288	0,18918	0,00027	0,23649	1,109
6	23,68370326	34,09587516	2,087270811	1,842743901	0,181487773	14,85228351	1,398176786	0,00038	0,00035	0,1409	0,16236	0,5585	0,00372	0,13379	1,088
7	43,17211538	31,69688815	3,568595919	0,901488722	0,238051441	23,41183044	2,660211625	0,00004	0,00012	0,07485	0,23132	0,60754	0,00037	0,08576	0,972
8	43,833587	20,05911933	1,996767665	1,770502889	0,616599295	12,57070585	2,439885654	0,00001	0,00977	0,20073	0,17497	0,53894	0,00143	0,07415	1,2
9	32,21848555	14,84769331	1,300770103	1,802021429	0,501088826	8,376397596	1,646382421	0,00584	0,02241	0,19648	0,14761	0,4314	0,02341	0,17285	1,396
10	39,83439734	15,47885001	1,240256137	1,909280541	0,62493627	9,978587419	1,214105208	0,01532	0,03259	0,18732	0,12066	0,40901	0,01536	0,21973	1,488
11	82,22153656	12,62367722	1,044280927	2,281509611	0,674726129	20,44288811	3,102573516	0	0,00866	0,3042	0,12251	0,44219	0,00481	0,11764	1,299
12	30,29447538	10,53878916	1,268028292	3,083705159	1,311276474	5,028237495	1,008016633	0,00867	0,04108	0,1848	0,0905	0,31791	0,0684	0,28865	1,575
13	37,32438071	13,61623744	1,294909106	2,886256778	0,854753574	14,2521519	3,458195916	0,00007	0,00365	0,3206	0,16215	0,40617	0,00948	0,09788	1,301
14	100,7143091	15,7845647	2,903521143	0,714792102	0,368133254	17,95836615	4,053850342	0	0,0043	0,12736	0,30014	0,47711	0,00149	0,0896	1,125
15	97,40469146	8,722264311	1,488226702	1,568993942	0,533210913	14,81574152	1,802581158	0	0,01338	0,16931	0,21008	0,46125	0,00627	0,13971	1,378
16	54,92876635	15,80303705	2,018488702	1,495463347	0,392428932	12,53769078	1,501897189	0,00257	0,00661	0,12032	0,194	0,49768	0,00576	0,17306	1,29
17	14,18004164	4,697616484	1,133198534	3,661486893	1,443050643	3,74087814	0,933719223	0,01925	0,08504	0,18274	0,10428	0,24667	0,1004	0,26163	1,734
18	41,27678607	14,51776868	1,245766003	1,861331703	0,613393403	5,62353325	1,096682744	0,00019	0,009	0,17887	0,11615	0,38328	0,03191	0,2806	1,387
19	14,91114892	3,580362444	2,481543616	13,84365417	6,607763777	1,424382112	0,787789788	0,03136	0,08992	0,16083	0,01323	0,07519	0,30743	0,32205	1,569
20	47,06714616	14,81030302	1,398695166	1,67800823	0,496326236	13,11552519	1,58314228	0,00005	0,00359	0,19491	0,18047	0,4605	0,00378	0,1567	1,314
21	74,584312	14,98304593	1,369617907	1,669139455	0,425252946	11,1210987	1,677216823	0	0,00332	0,17056	0,17489	0,48959	0,00507	0,15656	1,306
22	41,49764914	11,20077748	1,481783592	2,295505144	1,067735861	13,28511899	1,778980532	0,00008	0,05625	0,19232	0,15713	0,41755	0,00675	0,16993	1,488
23	29,57318477	11,27236749	3,517959387	18,51065282	10,05634713	3,22308991	0,706268553	0,01275	0,02943	0,13662	0,04586	0,13399	0,21185	0,42949	1,497
24	30,31292507	36,07394821	1,014301924	3,09455217	0,626201116	19,19516019	1,931165032	0,00133	0,00003	0,27604	0,10006	0,4721	0,00184	0,1486	1,207

25	40,51763153	19,05977326	0,947743659	2,558171422	0,706988206	13,66787682	2,277647663	0,02134	0,00226	0,31503	0,11052	0,41762	0,00315	0,13008	1,253
26	42,21592995	26,15178146	1,343248956	1,468187609	0,275086746	17,11404867	1,72711316	0	0,00004	0,21773	0,12489	0,46398	0,00367	0,18969	1,205
27	65,86951855	9,74566322	0,807671406	5,242843118	2,356092706	7,007408426	1,570061964	0	0,05678	0,33899	0,08319	0,24351	0,03469	0,24284	1,515
28	29,09435342	18,5165023	1,785019387	10,92019867	5,912244673	6,010436756	0,582356676	0,0057	0,00781	0,24182	0,01956	0,1245	0,09926	0,50134	1,256
29	44,17152051	12,44959156	0,842525076	2,507573111	0,884760414	14,17748691	2,881715292	0	0,00444	0,37233	0,13564	0,40101	0,00699	0,0796	1,267
30	66,88650227	13,67346609	1,590223125	1,467419519	0,248320672	26,98971592	2,46197152	0	0,01939	0,18785	0,17759	0,54382	0,00023	0,07111	1,152
31	25,91908025	16,12270642	3,521727353	0,864980312	0,260241555	12,35805303	4,86946846	0,00251	0,008	0,29144	0,20249	0,37825	0,00527	0,11203	1,048
32	6,835379478	3,22910248	1,2350057	17,18333912	7,251225509	3,096754393	1,01446531	0,02501	0,14877	0,22074	0,02036	0,11244	0,20984	0,26285	1,707

_		DCA	DH	DCULT	DSD	DAU	DVA	DVB	%CA	%Н	%CULT	%SD	%AU	%VA	%VB	Sh
	DCA	1	0.23	-0.1	-0.55	-0.52	0.6	0.25	-0.54	-0.52	0.08	0.59	0.49	-0.59	-0.38	-0.42
	DH	0.23	1	0.15	-0.55	-0.54	0.64	0.28	-0.52	-0.64	-0.05	0.44	0.65	-0.62	-0.33	-0.79
	DCULT	-0.1	0.15	1	0.07	0.15	-0.03	0.3	-0.06	-0.1	-0.59	0.36	0.15	0.16	-0.04	-0.35
	DSD	-0.55	-0.55	0.07	1	0.99	-0.61	-0.43	0.72	0.74	0.03	-0.82	-0.88	0.9	0.67	0.64
	DAU	-0.52	-0.54	0.15	0.99	1	-0.61	-0.4	0.69	0.71	-0.01	-0.78	-0.87	0.89	0.69	0.6
	DVA	0.6	0.64	-0.03	-0.61	-0.61	1	0.55	-0.51	-0.55	0.23	0.57	0.63	-0.71	-0.62	-0.78
	DVB	0.25	0.28	0.3	-0.43	-0.4	0.55	1	-0.2	-0.27	0.14	0.6	0.44	-0.47	-0.68	-0.65
	%CA	-0.54	-0.52	-0.06	0.72	0.69	-0.51	-0.2	1	0.89	-0.04	-0.6	-0.66	0.62	0.27	0.6
	%Н	-0.52	-0.64	-0.1	0.74	0.71	-0.55	-0.27	0.89	1	-0.09	-0.6	-0.7	0.7	0.28	0.71
	%CULT	0.08	-0.05	-0.59	0.03	-0.01	0.23	0.14	-0.04	-0.09	1	-0.34	-0.29	-0.15	0.03	0.01
	%SD	0.59	0.44	0.36	-0.82	-0.78	0.57	0.6	-0.6	-0.6	-0.34	1	0.88	-0.75	-0.75	-0.69
	%AU	0.49	0.65	0.15	-0.88	-0.87	0.63	0.44	-0.66	-0.7	-0.29	0.88	1	-0.84	-0.76	-0.72
	%VA	-0.59	-0.62	0.16	0.9	0.89	-0.71	-0.47	0.62	0.7	-0.15	-0.75	-0.84	1	0.66	0.68
	%VB	-0.38	-0.33	-0.04	0.67	0.69	-0.62	-0.68	0.27	0.28	0.03	-0.75	-0.76	0.66	1	0.55
	Sh	-0.42	-0.79	-0.35	0.64	0.6	-0.78	-0.65	0.6	0.71	0.01	-0.69	-0.72	0.68	0.55	1

ANEXO IV – Matriz de correlación de variables ambientales por barrio

	DCA	DH	DCULT	DVA	DVB	%CA	%CULT	%VB
DCA	1	0.23	-0.1	0.6	0.25	-0.54	0.08	-0.38
DH	0.23	1	0.15	0.64	0.28	-0.52	-0.05	-0.33
DCULT	-0.1	0.15	1	-0.03	0.3	-0.06	-0.59	-0.04
DVA	0.6	0.64	-0.03	1	0.55	-0.51	0.23	-0.62
СУВ	0.25	0.28	0.3	0.55	1	-0.2	0.14	-0.68
%CA	-0.54	-0.52	-0.06	-0.51	-0.2	1	-0.04	0.27
%CULT	0.08	-0.05	-0.59	0.23	0.14	-0.04	1	0.03
%VB	-0.38	-0.33	-0.04	-0.62	-0.68	0.27	0.03	1

ANEXO V – Matriz de correlación de variables ambientales por barrio, seleccionadas con r<0.7
Id	DCA	DH	DCULT	DVA	DVB	%CA	%CULT	%VB
1	-1,00722391	-0,87525379	-0,41807095	-0,76756067	-0,14354721	2,13997483	1,09268798	0,89168304
2	-1,34576654	-1,44956875	-0,65988496	-1,34915309	-0,85423500	1,41685464	0,69216680	1,13560695
3	-1,51579026	-1,60280637	0,01352023	-1,34814106	-0,02622553	4,44735666	-0,88185740	-0,04154413
4	1,27507407	1,57816531	-0,36695947	0,17661688	-0,47853806	-0,46293963	-0,39291690	0,21954370
5	1,32329112	1,00625351	-1,58352775	2,72994750	0,57785487	-0,46930372	3,08636034	0,40236403
6	-0,75782427	2,04769808	0,56143276	0,36698238	-0,53811129	-0,45418900	-0,89869372	-0,60491322
7	0,01853395	1,78174482	2,42680293	1,52110288	0,54247212	-0,46771270	-1,65518283	-1,07598943
8	0,04488493	0,49157417	0,44746597	0,05934746	0,35382393	-0,46890597	-0,21344401	-1,18985982
9	-0,41782488	-0,08616793	-0,42897437	-0,50618915	-0,32559189	-0,23701429	-0,26212045	-0,22181439
10	-0,11443044	-0,01619748	-0,50517705	-0,29015901	-0,69571763	0,14005828	-0,36703249	0,23798266
11	1,57414242	-0,33272380	-0,75196036	1,12078762	0,92123260	-0,46930372	0,97162682	-0,76331174
12	-0,49447150	-0,56385586	-0,47020475	-0,95763597	-0,87217544	-0,12444938	-0,39589475	0,91394711
13	-0,21442176	-0,22268793	-0,43635488	0,28606405	1,22572473	-0,46651943	1,15946060	-0,95711698
14	2,31083743	0,01769425	1,58930234	0,78578885	1,73573782	-0,46930372	-1,05377113	-1,03832687
15	2,17899247	-0,76523699	-0,19291819	0,36205526	-0,19185092	-0,46930372	-0,57330605	-0,54685011
16	0,48688265	0,01974210	0,47481836	0,05489589	-0,44930348	-0,36708046	-1,13440221	-0,21975472
17	-1,13642083	-1,21141207	-0,63999017	-1,13121617	-0,93579060	0,29637634	-0,41948851	0,64893609
18	-0,05697013	-0,12274359	-0,49823871	-0,87736969	-0,79625746	-0,46174636	-0,46381270	0,83499305
19	-1,10729577	-1,33527159	1,05792386	-1,44355928	-1,06073876	0,77805871	-0,67042986	1,24153289
20	0,17369996	-0,09031304	-0,30566147	0,13280779	-0,37973956	-0,46731494	-0,28010210	-0,38021291
21	1,26989895	-0,07116265	-0,34227724	-0,13610931	-0,29919076	-0,46930372	-0,55898945	-0,38158602
22	-0,04817162	-0,49046740	-0,20103173	0,15567485	-0,21205832	-0,46612167	-0,30976609	-0,25045363
23	-0,52320550	-0,48253088	2,36303848	-1,20103178	-1,13053912	0,03783502	-0,94771376	2,29529987

ANEXO VI – Tabla de variables ambientales seleccionadas por barrio, centradas y estandarizadas

24	-0,49373653	2,26698855	-0,78971165	0,95255111	-0,08175443	-0,41640219	0,64910247	-0,45965737
25	-0,08721250	0,38078601	-0,87352567	0,20728373	0,21491201	0,37950732	1,09566583	-0,64130074
26	-0,01955753	1,16701067	-0,37548255	0,67194590	-0,25646836	-0,46930372	-0,01873826	-0,05664839
27	0,92272847	-0,65178231	-1,04991274	-0,69077585	-0,39093923	-0,46930372	1,37008641	0,46464456
28	-0,54228066	0,32055863	0,18082031	-0,82520182	-1,23663534	-0,24258287	0,25717124	3,00000163
29	0,05834716	-0,35202305	-1,00602299	0,27599665	0,73212868	-0,46930372	1,75193875	-1,13640645
30	0,96324196	-0,21634352	-0,06447840	2,00352456	0,37273436	-0,46930372	-0,36096225	-1,21967601
31	-0,66877375	0,05518086	2,36778332	0,03067456	2,43408891	-0,36946700	0,82548297	-0,81833438
32	-1,42900952	-1,37421249	-0,51178871	-1,21806614	-0,86665393	0,52548373	0,01573611	0,66090180

ANEXO VII – *Script* de R para los diferentes pasos de la metodología

El símbolo # indica comentarios que no representan funciones ni operaciones en la consola de RStudio.

```
1 - Centrado y escalado de variables ambientales
```

Se seleccionan las columnas con valores a centrar y escalar tabla_variables<-scale(nombre_tabla[,c(3:19)],center=TRUE, scale=TRUE) rownames(varc)=vclmeans\$X

2 - Análisis de correlación y reducción de tabla de variables según criterio r<0.7

```
# Se calculan las correlaciones y se muestran en un plot
library(corrplot)
correlacion<-round(cor(tabla_variables[,c(1:17)]),2)
corrplot(correlacion, method="circle")
install.packages("caret")
library(caret)
# A continuación se elige el criterio de corte/cutoff para selección de variables
hc= findCorrelation(correlacion, cutoff=0.7)
hc = sort(hc)
# Y se crea una nueva tabla de variables seleccionadas
variables_reducida = varc[,-c(hc)]
```

3 - K-means para agrupamiento de barrios usando tabla "variables_reducida"

Cálculo de Suma de cuadrados y ploteo de gráfico de sedimentación library(vegan) par(mar=c(4,4,2,2)) wss <- (nrow(variables_reducida)-1)*sum(apply(variables_reducida,2,var)) for (i in 1:32) wss[i] <- sum(kmeans(variables_reducida, centers=i)\$withinss) plot(1:32, wss, type="b", xlab="Número de Clusters", ylab="Suma de cuadrados intragrupo") # para recuperar el id de cada fila: rownames(varedu)=vclmeans\$name

k-means del conjunto de variables, usando k=4 (asignado a un fit)

Kmeans4 <- kmeans(variables_reducida, 4)

Cálculo de medias/centroids de cada cluster

Centroides_k4<-aggregate(varedu,by=list(fit_4\$cluster),FUN=mean)

grafica el ordenamiento de barrios agrupados

library(cluster)

clusplot(variables_reducida, fit_4\$cluster, color=TRUE, shade=TRUE,

labels=2, lines=0,main="k-means varedu (k=4)")

extracción del valor de k para cada barrio

k4<-fit_4\$cluster

4 - Modelo Lineal Generalizado / GLM

Para familia Poisson se usó el mismo código pero con la función glm(), usando Family=Poisson # dentro de los argumentos.

#GLM con familia de distribución Binomial Negativa, usando la columna de conteo de casas #positivas como variable respuesta de la interacción *cluster:*año, aplicando offset de "corrección" #en relación al total de casas muestreadas por barrio.

library(MASS)

glm_bn<-glm.nb(casas_positivas~año*k4+offset(log(total)),data=tabla_clorinda)

para detectar la presencia de interacciones significativas:

library(car)

Anova(glm_bn, type=2)

para observar el resumen de la salida del GLM

summary(glm_bn)

Gráficos de dispersión de los residuos
par(mar=c(4,4,2,2))
plot(glm_bn, which=1:4)
par(mfrow=c(1,1))

5 - Contrastes a posteriori

library(car) library(sjPlot) library(sjmisc) library(multcomp) library(factorplot)

Se crea nueva variable year.k (interacion año*k), en una nueva columna "interaccion"

tabla_clorinda\$interaccion<-with(clototalk4,interaction(year,k4,sep="_"))

Se vuelve a correr el modelo, pero la interacción year*k4 se incorpora ahora como la # nueva variable creada.

glminter<-glm.nb(positivas~interaccion+offset(log(total)),

data=tabla_clorinda)

Se realizan los contrastes con la función glht (General linear hypotheses and multiple # comparisons)

glminter.apost=glht(glminter, linfct=mcp(interaccion="Tukey"))

Aposteriori=factorplot(glminter.apost, pval=0.05,

adjust.method="Bonferroni",

factor.variable = "interaccion", data=tabla_clorinda)

Se obtiene el gráfico de contrastes de medias, con la magnitud de los efectos y el desvío # estándar

plot(Aposteriori,trans="exp")